DAAD Basics of Radio Astronomy

2022

Contents

1 Introduction

1.1 History of Radio Astronomy

Everyone can immediately name the brightest astronomical object in our sky, our Sun. It shines with such a high intensity that it provides about 10,000 times more energy than the entire human race needs today. It is dazzlingly bright in visible light. Our eyes detect light logarithmically. This means that if we have two objects, one of whom is ten times brighter than the other, to our eyes it will only appear twice as bright. Despite this, it is not possible for us to look at the Sun without pain. Its radiation is too intense. Therefore, the Sun is the object in the sky which intuitively should be the brightest over the entire electromagnetic wavelength range. If the detection of the Sun does not succeed in a wavelength range, it is probably not scientifically relevant. It is this train of thought that will guide us in the discovery of radio astronomy to begin with.

The Sun is invisible at radio waves

At the beginning of the 20th century, astronomers already knew that the Sun and stars had very comparable physical properties. The continuous part of the solar spectrum can be described almost perfectly by the established Planck black–body spectrum. The unique feature of a black–body spectrum is that it can be completely described by a single variable, its (effective) temperature, which is measurable quantity (solid line in Fig. 1.1.1).

When sunlight is refracted into a rainbow, the various colours come from the black–body spectrum emitted by the Sun. When this spectrum is more closely inspected with a spectrometer, it can be seen that there are fine dark lines superimposed on the rainbow's continuous spectrum (see Fig.1.1.1 yellow filled area). These lines are darker than their immediate vicinity in wavelength. This is why they are called absorption lines. They are caused by neutral atoms re-configuring the population number of their shells. The light is being absorbed, removed from the Sun's interior black–body spectrum, by neutral gas atoms located in the outermost parts of the Sun's stellar body, the so–called photosphere. These absorption lines were first discovered by the German physicist, Joseph von Fraunhofer, using state of the art optical instruments. These dark lines can be used to determine the chemical elements that make up the Sun. These elements represent the chemical composition of our solar system. They are the basis for all the elements from which the planets and ultimately humanity were formed.

At the beginning of the 20th century, photography as well as spectroscopy was technologically advanced enough to deduce the spectral type, mass, lifetime and future evolution of a star from its colour and absorption lines. The colour of stars depends on the temperature T_{eff} (Eq. 1) and age of the star, and can range from hot and young blue stars, to older and cooler red stars. By measuring the colour of the star (Wiens displacement law, Fig.1.1.3), its effective temperature of the star, can be calculated. Typical values are between 30,000 K for blue (O-stars) and 3,000 K for red stars (M-stars).

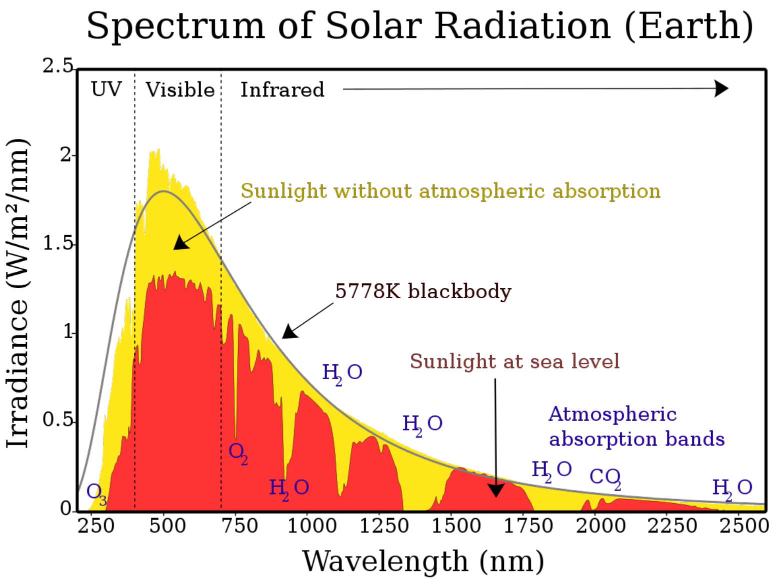

Fig. 1.1.1: Spectrum of solar radiation at sea level (red) and out of the atmosphere (yellow) from [8]. The black–body spectrum is marked by the thin solid line. The Fraunhofer lines from the Sun's photosphere are visible by isolated dips in the yellow intensity distribution. At sea level the absorption of the Earth's atmospheric molecules is easily visible by the intensity distribution of the red spectrum. These lines are broad because here the rotation of the molecule absorbs the electromagnetic radiation in a dense and cool environment close to the ground. The high number of broad water vapor absorption lines underlines the role of water vapor as most abundant green–house gas. Making Earth habitable for human kind.

1.1.1 Long-wavelength radio astronomy sees no stars

Planck's radiation law (Eq. 1) describes the spectral-energy distribution of radiation emitted by a black–body. Note, the whole spectrum is determined entirely by a single variable, the effective temperature T_{eff}.

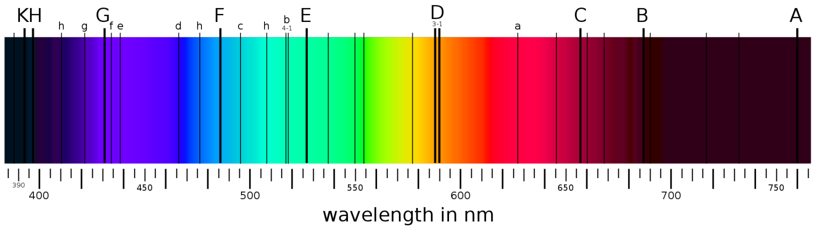

Fig. 1.1.2: The Fraunhofer absorption lines appear as dark lines against the continuous black body frequency spectrum emerging from the Sun's interior. The lines mark the different neutral and partly ionised elements the Sun is consisting of. The position of the absorption lines in the frequency spectrum depends thus on the chemical composition of the Sun as a stellar body and the effective temperature, determining the population of the different electronic levels of the atoms and ions. The width and the depth of the absorption lines contains information on the chemical abundance of the different elements.

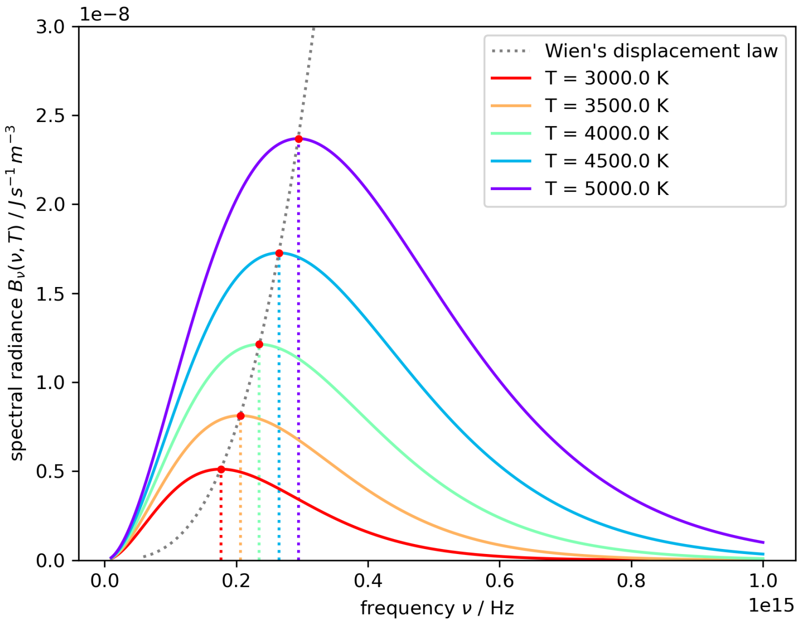

Fig. 1.1.3: Planck's law for different temperatures. Note how the maxima are evenly spaced in frequency. This is called "Wien's displacement law". You can find the code used to create this plot in Sec. **??**.

$$B_\nu = \mathrm{c} \cdot \frac{1}{exp\left(\frac{\mathrm{h}\cdot\nu}{\mathrm{k_B}\cdot T_{\mathrm{eff}}}\right) - 1} \tag{1}$$

Planck's law describes how much energy is given out at each wavelength from a black–body. Black–body denotes the fact that its surface is absolutely opaque. All radiation received stems form the surface. From Planck's law can be used to calculate the number of photons emitted per second (radiation power) from the Sun for specific wavelength ranges. However, for long wavelengths, such as radio waves, the results of such a calculation are discouraging. Inserting the numbers for the Planck h, Boltzmann constant $\mathrm{k_B}$, the radio frequency and the effective temperature we find a ratio value around 10^{-6} or even lower. Implying the a tiny little fraction of the total black–body energy is emitted at radio wavelength.

The radiant power in the radio wavelength range is thus multiple orders of magnitudes fainter than in the visible light wavelength range. This meant that in the early 20th century, radio waves from stars were not expected to be detected, not even from the Sun. Perhaps the Sun might be detectable but with a very sensitive equipment, as the scientist in this epoch expected, but nothing from beyond the solar system.

Even more discouraging was the discovery that electromagnetic radiation, with a wavelengths on the order of a few hundred meters, are reflected back by the Earth's atmosphere (Fig. 1.1.4). If a signal consisting of photons in this wavelength range is sent into the sky, it is bounced back down to the ground, like a ball that is shot against a wall. It was quickly discovered that this effect could be used for radio communication. In 1901, Guglielmo Marconi achieved the first intercontinental communication using radiation with wavelengths of approximately 350 meters. This also works in the other direction, where radio waves originating from space are bounced back from the atmosphere and never reach the ground.

The reason for this bouncing of radiation from the uppermost layers of the atmosphere is the ionized outermost layers called the ionosphere. The high energy radiation, such as ultraviolet and X-rays, from the Sun separates the electrons from the atoms and molecules which make up this layer. Once separated from their hosts, these free electrons can move freely. These electrons interact with the incoming radio waves, vibrate and then remit the energy at the same frequency. This effectively bounces the radio waves back, making the ionosphere act like a mirror. The incoming radiation, whether from the ground or from space, is reflected, according to the simple laws of geometric optics. Such long radio waves cannot be used for astronomy as they are simply reflected back into space and never reach the ground.

However, this is only true for very long wavelengths, longer than about 20 m. For shorter wavelengths, the Earth's atmosphere is transparent and does not absorb or reflect signals, which is why they are widely used today. For example, they are used

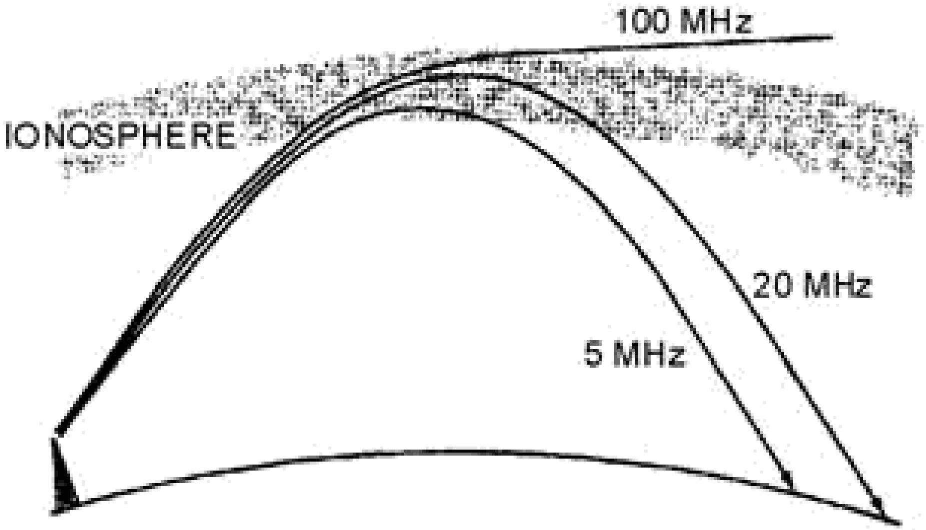

Fig. 1.1.4: Image from rfcafe. Free electrons in the Earth's ionosphere act as a mirror for long wavelength or low frequencies. The ionisation fraction varies with the Sun's irradiation. Also a day and night cycle is observable at low frequencies. Frequencies higher 30 MHz can penetrate the ionosphere. Frequencies below 80 MHz suffer by scintillation, comparable what we see as flickering light of stars in the visible, both simply refraction. Displayed here is the reflection of low frequencies back to the Earth's surface. Allowing to use these frequencies for terrestrial communication. High frequencies are accordingly necessary to communicate with satellites.

to receive television signals from geostationary satellites. Only during heavy rain do these satellite signals disappear, but this is caused the signal being absorbed by the weather layer of the atmosphere (the troposphere) and not due to reflection from the ionosphere.

For a long time it was believed that there were no good reasons to pursue long wavelength radio astronomy due to the combined findings that stars are weak radio emitters, according to Planck's law, and that long wavelength radio waves are reflected by the ionosphere of the upper Earth's atmosphere.

year.

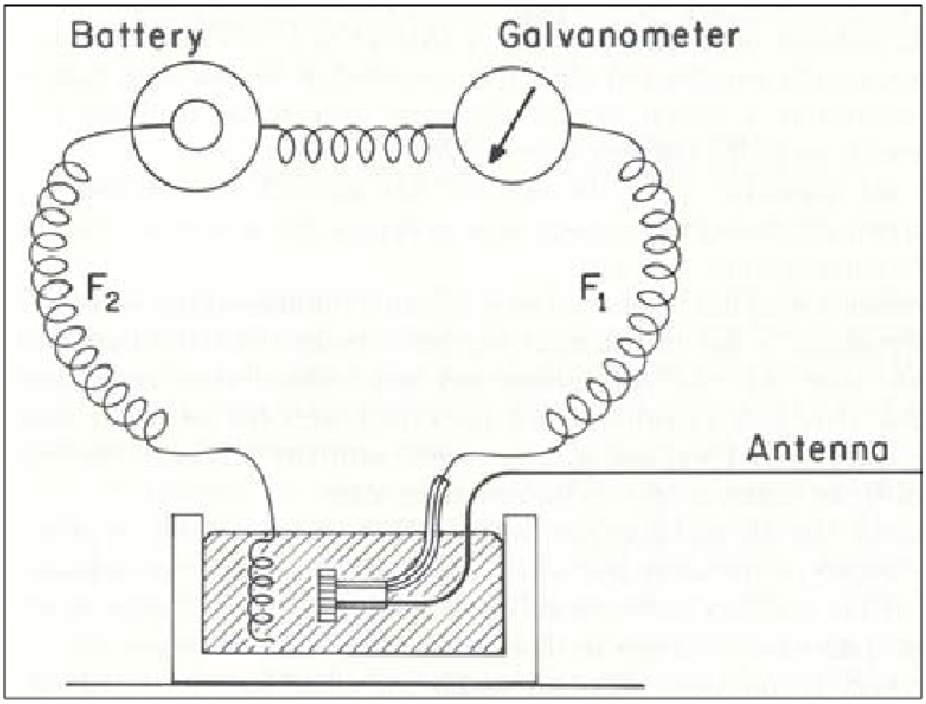

Figure 4: The radio receiver used by Nordmann and Haberkorn at the Bossons Glacier, Mont Blanc, in September 1901 (after Sullivan, 1982: 159; cf. Nordmann, 1902a: 274).

Fig. 1.1.5: Image from Debarbat, Lequeux & Orchiston 2007. The antenna used has a certain resonance frequency. In the case that the Sun emits at that frequency the Sun's emission will induce an additional current changing the resistivity of the system while the voltage remains constant due to the battery. Simply Ohm's law.

In fact, in 1896 and 1901 this was tested by two research attempts. Johannes Wilsing and Julius Scheiner, from the University of Potsdam in Germany, and independently by Charles Nordmann, from the Nice Observatory in France [1].

In 1896 Wilsing and Scheiner used a mirror to focus the Sun's radiation onto an electric circuit, searching for changes in the resistivity of the material illuminated. They failed to detect the Sun at long radio waves. In a very detailed description they point out their instrumental limitation of their attempt [10].

Charles Nordmann climbed with a radio detector to the Aiguile du Midi, a mountain in the French Alps in an attempt to detect also m-radio waves from the Sun. He used the ice of the glacier at 3000 m to build up an antenna with a large collecting area. The antenna consisted of a single long cable of 175 m length. As well as Wilsing and Scheiner Nordmann failed to detect the Sun. In retrospective both attempts were unlucky. The Sun cycles through periods of higher and lower activity over a timescale of approximately 11 years. During the time period of their attempts, the Sun was undergoing a period of lower activity. Nordmann himself expected that the Sun's activity affects its luminosity in the radio band but he did not repeat his attempt. If both teams had observed during a strong solar maximum, their apparatuses would probably have detected the Sun (Fig.: 1.1.5).

Karl Jansky and non-thermal radio radiation

While radio waves were excessively used since the beginning to the 20th century for radio broadcast and communication, radio astronomy was not of any interest. Perhaps also due to the multitude of important discoveries during this period of time driven by astrophotography and spectroscopy in visible light. Using photography the women of the Harvard University made ground breaking discoveries on the nature of stars, their distances – i.e. the discovering the period –luminosity relation of Cepheids –. Harlow Shapley determined the size and the place of the Earth within the Milky Way Galaxy, revolutionizing the picture of the universe be stretching the distances to an extreme. Finally Vesto Slipher and Edwin Hubble discovered the cosmological expansion of the universe as a whole.

So it lasts 30 years after the first attempts of the radio astronomy pioneers to detect eventually radio waves from space accidentally. Karl Jansky discovered something profoundly amazing. His employer, the Bell Telephone Laboratories in New Jersey, gave him the assignment of investigating atmospheric effects to find out why the quality of telephone calls was not up to the standard they technically should be. It was clear that interference was coming from outside the telecommunications system. It was not the system's own hardware causing the issue, but interference from the outside environment.

Jansky built an antenna to receive a signal of 20 MHz or a wavelength of about 14.6m. Recall from what was discussed previously, this wavelength is just short enough to not be completely reflected by the ionosphere (Fig. 1.1.4). The antenna was mounted on four wheels, which actually belonged to a Ford T-model car, so that it could

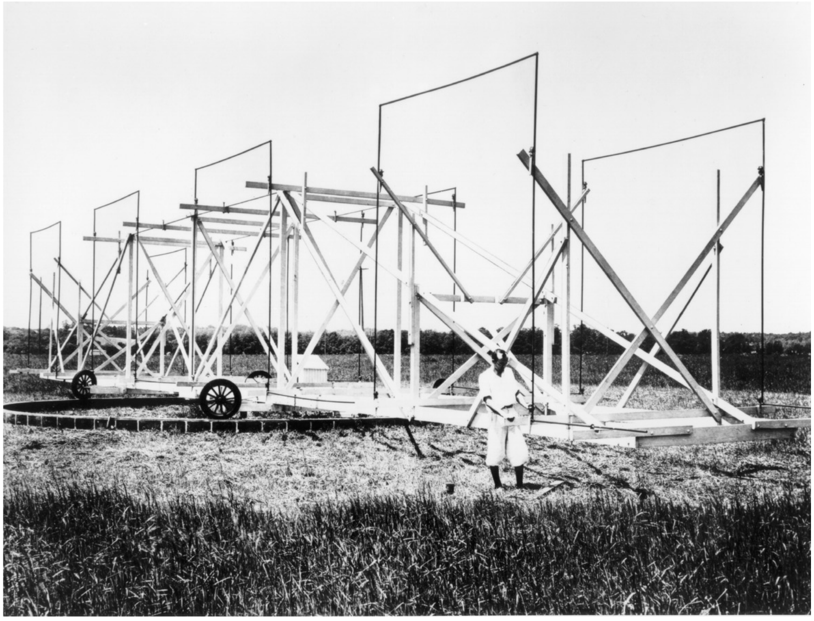

Fig. 1.1.6: Photograph of Jansky with his telescope in the early 1930's from [4]. Image courtesy of NRAO/AUI.

rotate. This allowed Jansky to point the antenna in different directions, similar to putting a hand around one's ear, to determine the direction from which a sound comes from. The directionality of Jansky's antenna was not great but sufficient to detect that the short term interference in the call was due to near and far flashes in the atmosphere caused by thunderstorms. However there was another kind of interference also present, which could not be explained by weather patterns. Janksy referred to this as "hiss". This interference grew to a maximum every 23 hours and 56 minutes. Unmistakable for all astronomers, this was a signal related to the rotation period of the Earth around its own axis. The duration of a stellar day. Jansky managed to fix the direction so precisely that he aligned the radiation maximum to the position of the constellation Sagittarius. From his location in New Jersey, the center of the Milky Way could not be seen directly. He was too far north. However, the location of the radiation maximum was located towards the Galactic center. His results made front page news in the New York Times on May 5, 1933, just as Edwin Hubble's discovery of the redshift of galaxies had done a few years previously.

This discovery was certainly worthy of a Nobel Prize, but was not awarded one. There was still too much bias in the scientific community against the possible detection of radio waves. Also Jansky did not succeed in detecting radio radiation from the Sun, which was believed at the time should be the largest radio wave emitter purely because of its proximity. A black–body emitter could not be the source of this new, unknown radio radiation. The necessary effective temperatures would be far to high, with temperatures of several million Kelvin, which is "unphysically high". The term non-thermal radiation was used to describe this source of radio waves as they must originate from processes not associated with thermal radiation. The bias amongst the scientific community at the time against the detection of radio waves originating from outside of the Earth was so strong that it delayed the investigation of Janky's unexplained detection of radiation.

Grote Reber

It took another mind, a more practical one, to make finally progress. Grote Reber was the one who had it. He was a technician with an interest in astronomy and above all technical know-how. During the summer of 1937, he built a prototype for one of the first modern radio telescopes in his garden, all at his own expense. It was a large parabolic mirror with a 10m diameter and a receiver in the primary focus. He began to look for Jansky's radiation at short wavelengths.

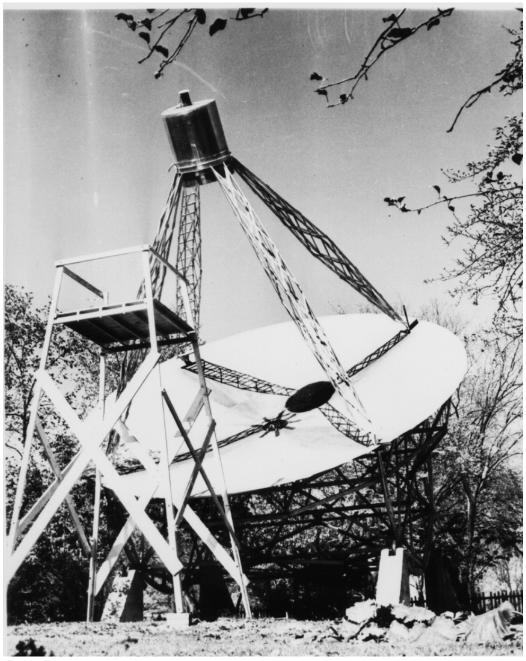

Fig. 1.1.7: The original 9-m telescope built by Grote Reber in his backyard in 1937 (left, from [6]), and the reconstruction located at the National Radio Astronomy Observatory in Green Banks, West Virginia (right, from [9])

Also Reber's thoughts were in-line with the classical physical approach of thermal emitters. He knew that the Sun and thus the stars are radiating like black bodies. Inspecting Eq. 1 and realising the exponent is something to 10^{-6} one derives the

Rayleigh–Jeans approximation of Planck's law by a series expansion:

$$B_\nu \simeq \nu^2 \cdot T_{\text{eff}} \tag{2}$$

This is called Rayleigh-Jeans approximation of the Planck radiator for long wavelengths. It gives the maximum radiation at high frequencies (or equivalently at short wavelengths). So, Reber started his investigations at the highest frequencies he could technical archive. Most likely he had the picture of radio stars in his mind. Radio stars were thought to be dim at the optical but luminous at the radio wave band. Applying Wien's displacement law make these radio stars to the coldest objects in the Universe. But Reber observed nothing. Successively he developed new receivers for longer wavelength, moving in frequency towards Jansky's wave band of detection. It was only when he approached 160 MHz or 2 m wavelength, that he began to detect the Jansky radiation. Reber began to map this radiation, and was the first to map the Milky Way in the radio wave range ([7]). Reber measured that the distribution of the radiation was similar to the distribution of stars in the Milky Way. This did not help to determine the origin of this radiation. While he did not find the source of the radiation, his work work helped to confirm that the Jansky/Reber radiation at long wavelengths could no longer be ignored by scientific community.

In his book "Blick ins kalte Universum", Peter G. Mezger wrote down the following story which reflects the attitudes of members of the scientific community at the time when Grote Reber tried to publish his results scientifically. "My friend Bart Bok of Harvard College Observatory, always with a whimsical story at hand, remembered this publication. Reber had submitted his paper to the Astrophysical Journal, the prestigious journal of the American Astronomical Society. No one there knew Grote Reber, no one dared to judge whether it was a serious paper or the work of a charlatan. Bart Bok, as co-editor of the journal, was asked to visit Grote Reber at his home in a Chicago suburb and see the radio telescope in person. Bok went there, but met only Reber's mother. After he made his request, the old lady said, "Yeah, just go out in the yard, that's where the damn thing is. Ever since Grote put it up there, I don't know where to hang my laundry." Bok was thus at least able to convince himself of the existence of the radio telescope." " Further inquiries into Reber only revealed that he was an electrical engineer by profession, a member of the Institute of Electrical Engineers, and that he always paid his membership dues. This was enough to get his article published in the Astrophysical Journal in 1944.

Radio Engineers to Radio Astronomers

As World War 2 came to a close in 1945, radar technology and radio astronomy used very similar technologies. Radar engineers began the transition to becoming radio astronomers. In the Netherlands, at Leiden Observatory, Jan Hendrik Oort was particularly noteworthy. He approached the new astronomical research field as an

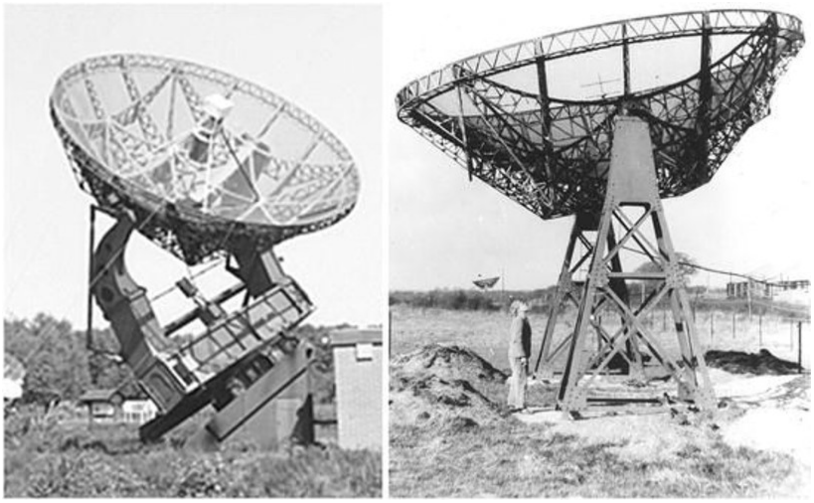

Fig. 1.1.8: This giant WWII radar dish was called "Würzburg Riese". With 7.5 m diameter it was one of the technological developments of Nazi-Deutschland to identify allied bomber formations. The antenna was operated at 560 MHz. Recall Fig. 1.1.4, above the frequency where the ionosphere blocks the microwaves. Thus, these instruments detected already the Sun and eruptions at its surface during WWII. Hundreds of these Instruments were after the war distributed across Europe. On the left hand side an asto-modified is displayed, using an equatorial mount to compensate the Earth rotation.

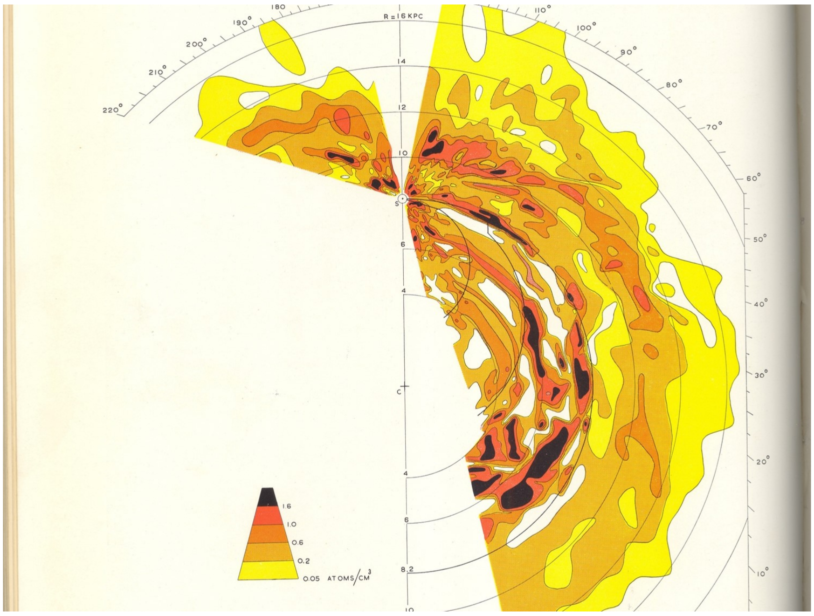

Fig. 1.1.9: Jan Hinrich Oort used an modified Würzburg Riese to explore the Struktur of the Milky Way Galaxy. He was aware, that interstellar dust blocks effectively visible light from distant portions of the Milky Way. He and Hendrik van der Hulst identified the 21-cm line of neutral atomic hydrogen as a tool to explore even the distant outskirts of the Galaxy. Displayed here is the map showing the spiral structure of the Milky Way Galaxy. Note that this map is made without any digital equipment.

astrophysicist. Already in 1944 Henk van der Hulst predicted that the 21-cm line of neutral atomic hydrogen, the most abundant element in space, is detectable in emission. The "forbidden" line transition has a life-time of several million years, in consequence it is extraordinarily narrow in line width. So, Doppler motion guarantees that resonant absorption along the line of side is extremely rare. This "optical thin" line emission is able to disclose the structure of the Milky Way Galaxy.

Fig. 1.1.10: Taken from CSIRO Gordon Stanley is seen using a theodolite to pin-point the position of the antenna mast. This radio telescope was used to survey the sky for radio sources, and for a detailed study of a strong radio source that was detected in the constellation of Sagittarius. The radio astronomers realised that this source (called Sagittarius A) was located at the very centre of our Galaxy! These observations gained further fame for Australia and Dover Heights when in 1958 the International Astronomical Union decided to adopt the position of Sagittarius A as the coordinate centre for the system of galactic ‘latitude’ and ‘longitude’ that is still used today by all astronomers.

In Australia the Dover Height Team, John Bolton, Bruce Slee, Gordon Stanley, Kevin Westfold and Dick McGee used the radar equipment to explore from an ingeneurs perspective the space. They did extremely well. Their location at the southern sky allowed them to perform unique observations. For a Milky Way astronomer most remarkable is the discovery of Sgr A*. While the optical astronomer already discover a glimpse of the Galactic center region towards the Baade Window, the radio astronomer identified the location of the very center. Using radio continuum emission at 400 MHz they determined the location of the central source that well that 1958 the internation astronomical union (IAU) adopted officially that source as the center of Milky Way

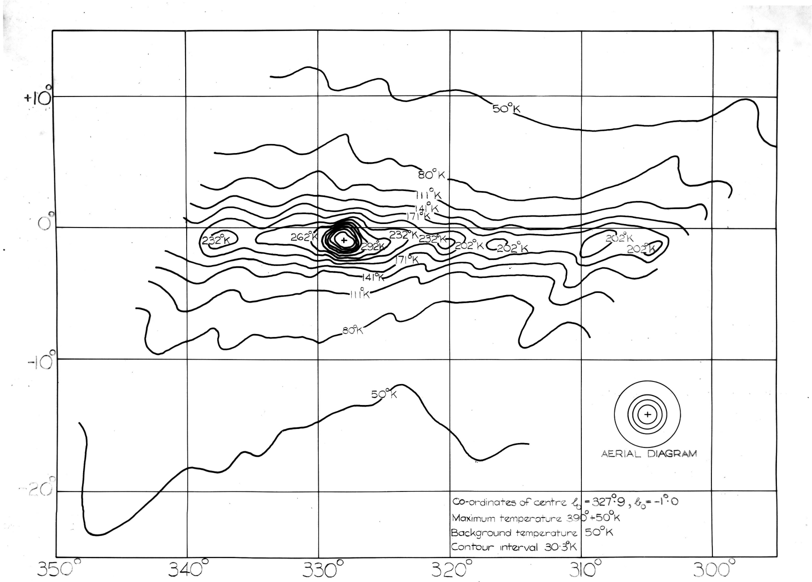

FIG. I

CONTOURS OF AERIAL TEMPERATURE OBSERVED AT 400 Mc/s.

Fig. 1.1.11: Taken from CSIRO This diagram was obtained from observations taken with the Hole-in-the ground antenna in 1954. It shows contours of radio emission detected at a frequency of 400 MHz, in the region of the Galactic Centre. The strong source near the centre is Sgr A. Note that the position of Sgr A does not coincide with the centre of the coordinate reference frame at 360 degrees longitude and 0 degrees latitude. The coordinate frame was first determined by optical astronomers, but the Dover Heights investigations revealed the true position of the galactic centre. This new position was subsequently adopted by the International Astronomical Union and the reference frame for all astronomical observations of our Galaxy was adjusted to reflect this. In 1954 the focus of the Division of Radiophysics shifted to the Fleurs field station, and the Dover Heights facility was closed down, bringing to an end a remarkable decade of scientific breakthroughs and achievement. More than any other field station, Dover Heights helped establish Australia's reputation as the world's leading nation in the emerging field of radio astronomy

Galaxy (Fig. 1.1.11). Note the displacement between the galactic coordinate origin and the location of Sgr A*. This displacement is because of the inability of optical light to travel long distances through the Galactic plane because of dust extinction. Moreover, the team discovered hundreds of distant sources and compiled first catalogs of extragalactic radio sources with optical counterparts. Finally, they identify by chance the radio interferometry, see below (Sect. 1.2).

Using a modified Würzburg Riese Oort deduced for the very first time the spiral structure of the Milky Way Galaxy (1.1.9). By that he identified also the flat rotation curve of the Galaxy. He and his Dutch collaborators also were one of the first who discovered the emission line of the neutral atomic hydrogen from space ([3])

In Germany, Albrecht Unsöld at the University of Kiel, Ludwig Biermann, Karl-Otto Kiepenheuer and Heinrich Siedentopf also looked at the results of the radio engineers with the eyes of the astrophysicists. The Sun once again took on a special role. In 1948, using a German V2 rocket from the war, US astronomers were able to carry out X-ray observations of the Sun from outer space. This led to the surprising discovery that the solar corona, the outermost layer of the Sun's atmosphere was extremely hot at millions of degrees Celsius, while further down, the photosphere was about 5900°C. When this new higher temperature is inserted into Planck's radiation law, the resulting black–body spectrum indicates that the solar corona should emit detectable amounts of thermal radio radiation.

1.2 Importance of Radio Astronomy IMAGES FORMATTING

The technical foundations of radio astronomy are the basis of communication in our society. Whether cell phones, television reception via satellites or radio signals, they are all transmitted by means of electromagnetic waves in the radio wavelength range. The signal processing of today's radio telescopes is done digitally. Much of the technology, methods and techniques that are used in the data acquisition and analysis of radio telescope data are also used commercially in industry, especially with modern developments such as "Industry 4.0" and the Internet of Things.

In academia, radio astronomy has also become indispensable for the study of the Universe. This is because radio radiation can travel across vast distances almost unabated due to its long long wavelengths. Not a single photon in the optical wavelength range reaches astronomers from the center of the Milky Way, because of large dust and gas clouds which obscure the view. On the other hand, radio waves can travel from the Galactic center easily and can be studied in great detail. The radiation from the Galactic center, its discovery and correct interpretation by Karl Jansky marked the beginning of radio astronomy. Using the light of neutral atomic hydrogen, at 21-cm wavelength (which will be explained in much detail later), the most distant regions of the Milky Way, and even further afar in extragalactic space have become explorable. To optical astronomy, the regions behind the Milky Way is known as the ”Zone of Avoidance” because the interstellar and intergalactic gas blocks the view of everything beyond, except in radio astronomy.

Radio waves are in a wavelength range that is beyond the sensitivity of human senses. The development of modern science, especially Maxwell's description of electromagnetic radiation, and the successful generation and transmission of electromagnetic signals by are closely related to major technological advances in the radio wave frequency range in the very beginning of the 20th century. Today, the frequencies below 300 GHz, or wavelengths greater than 1mm are referred to as radio waves. the relation bewteen frequency and wavelength is given by the equation $c = \lambda \cdot \nu$, where c is the speed of light. Therefore, large wavelengths imply low frequencies. The equation $E = \mathrm{h} \cdot \nu$ gives the energy of a photon of electromagnetic radiation, where ν is the frequency of the photon, and h is Planck's constant. This results in not only only hot objects emit radiation, such as the hot photospheres of stars which give out high energy radiation, but also cold objects such as Cold Molecular Gas which emit low energy radiation. This is because the gas requires little energy to produce a photon in the radio wave range. Even the cosmic microwave background, whose radiation can be perfectly described by a black–body with a temperature of only 3 K, the coldest signal we can observe from the Universe at present, becomes observable with the techniques of radio astronomy. It is only with radio astronomy that the formation of the first galaxies and stars in the Universe by electromagnetic radiation can be studied. Using gravitational waves, astrophysicists will also begin to be able to look for these early signals in a little more than a decade. The next large radio telescope is the Square Kilometer Array (SKA), set to be complete in the late 2020's. It is currently under construction at two sites, in Australia and South Africa. Its huge collecting area and its angular resolution of 1 arcsec at 21-cm wavelength will allow observations to take place that will significantly broaden and deepen our understanding of the early Universe in the coming decade.

Radio Interferometry

[note [JB]: Maybe include image of ALMA or other telescope array] Radio interferometry is the technique of combining the signal from multiple separate radio telescopes in order to act like a single telescope, which provides much higher resolution than single dish radio telescopes. Radio waves, like all radiation we receive from objects outside the solar system, is coherent radiation. That is, the phase and amplitude position of electromagnetic waves are fixed and not a function of time. In simple terms this means that the radio waves travel in parallel lines, and the location of the peaks and troughs of the waves are aligned. Because of this property, which is not altered in the radio wave range by the passage of the signal through the terrestrial

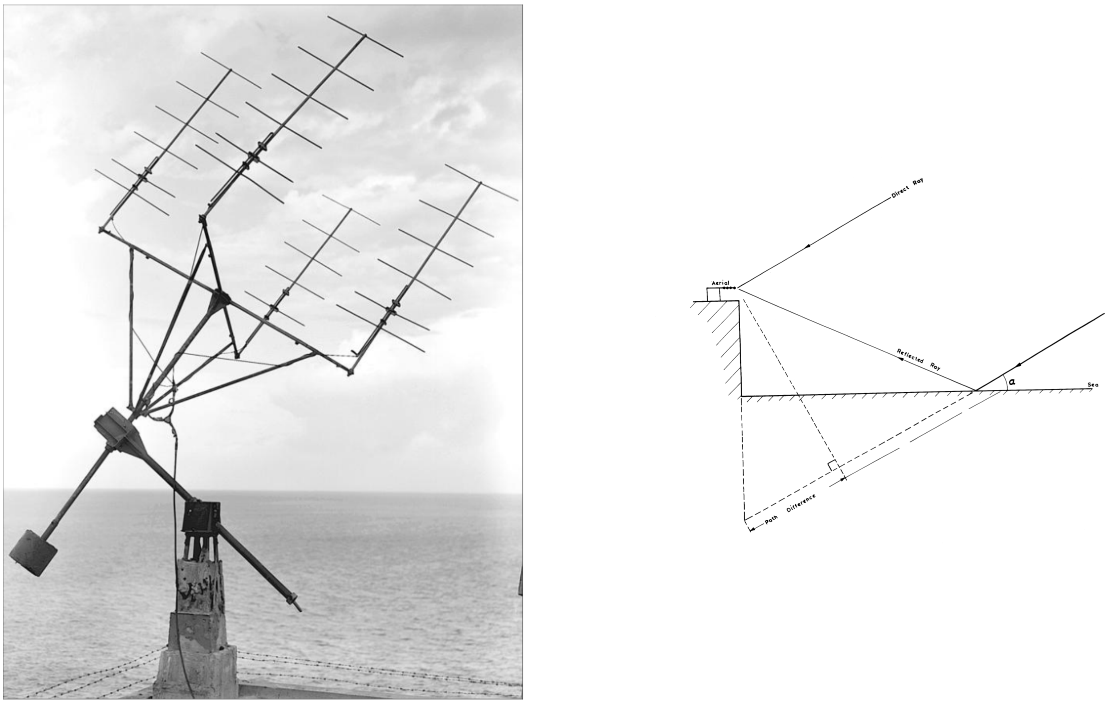

Fig. 1.2.1: Left: The Yagi antenne at the Dover Heights in Australia. This 4-element Yagi antenna was used until early 1950th at Dover Heights at a frequency of about 200 MHz. Its location allows a spectacular discovery, the radio interferometry. Right: The optical geometry of the Yagi antenna. The radio emission of the cosmic sources reaches the antenna from two different direction. a) directly from space and b) reflected by the ocean's surface. The path-length of both lines is apparently different. So the different wave fronts interfere constructively and destructively with each other, creating the temporal signature of an interference pattern at the receiver. from CSIRO)

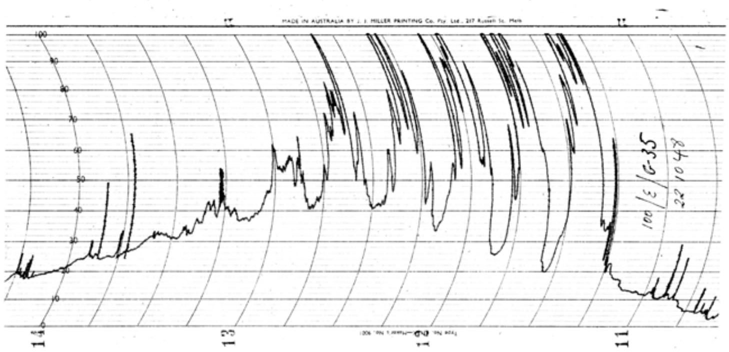

Fig. 1.2.2: Taken from CSIRO This photograph shows a sea interferometer chart recording of the strong extragalactic radio source, Cygnus A at 200 MHz with the sea interferometer. It was possible to calculate that the angular size of Cygnus A was less than eight arcminutes. Much smaller than the angular resolution limit of the Yagi four element antenna. The interference fringes from Cyg A seen here show unusual intensity fluctuations. Even more important is the fact, that radiation from space is coherent. This is not limited to the radio domain, but also true for high energy radiation and gravitational waves. The huge distances warrant this remarkable nature of radiation received form space. Allowing nearly unlimited angular resolution by using interferometry.

atmosphere, radio interferometry is a common technique used by radio astronomers. Signals received by individual radio telescopes can be combined across large distances or even across entire continents. They give an image of an astronomical source with an incredible level of detail. The angular resolution of a telescope is given approximately by λ/D, where D is the diameter of the lens. In the case of interferometry, the value D is the physical distance between the telescopes. Today, radio interferometric observations are performed regularly even using satellites orbiting Earth. **[note [JB]: the importance of radio interferometry is not really explained. Maybe list some of the big discoveries made with it and why they are important]**

The study of matter in strong gravitational fields

[note [JB]: what exactly does this section have to do with radio astronomy. Probably don't need to go into as much detail. I don't think equations add to understanding.] Radio astronomy allows the study of matter in strong gravitational fields, for example neutron stars. These are known in radio astronomy as pulsars since about fifty years. Let us compare the gravitational acceleration at the surface of the heaviest body of the solar system with that of a neutron star.

$$\begin{aligned}
\frac{g_{\mathrm{neutronStar}}}{g_{\odot}} &\propto \frac{M_{\mathrm{neutronStar}}}{M_{\odot}} \cdot \left(\frac{r_{\odot}}{r_{\mathrm{neutronStar}}}\right)^2 \\
&\simeq 1.5 \cdot \left(\frac{r_{\odot}}{r_{\mathrm{neutronStar}}}\right)^2 \\
&\simeq 1.5 \cdot \left(\frac{700 \cdot 10^6 \mathrm{m}}{15 \cdot 10^3 \mathrm{m}}\right)^2 \\
&\simeq 3 \cdot 10^9
\end{aligned}$$

then it becomes clear that only by observing pulsars can the physics of matter in strong gravitational fields be probed. Hulse and Taylor received the Nobel Prize in Physics in 1993 for their study of the pulsar 1913 + 16, because they succeeded in detecting for the first time the gravitational waves predicted by Einstein by observing the change in the time interval of the pulses. Only 23 years later, the direct detection of gravitational waves with LIGO, which resulted from the merger of two black holes, showed them directly for the first time. In 2017, this resulted in the second Nobel Prize for gravitational wave research and another for astrophysics.

2 Antennas, Waves and Diffraction CONSISTENCY

2.1 Angles and distances: properties of waves from space IMAGES FORMATTING CONSISTENCY

The electromagnetic radiation reaching us from cosmic objects is coherent, meaning that they reach us in the form of plane waves. This coherence is essential for the exploration of the cosmos by means of electromagnetic radiation (this is also valid for gravitational waves, and in general for all waves reaching us from sources at large distances). These plane waves are distorted by the the collecting surface of a telescope, for example the promary mirror or lens, called the telescope's aperture. this distortion is referred to as diffraction and forms distinct patterns, called the Airy disk in optical astronomy. The Airy disk is simply a diffraction pattern of coherent waves at a circular aperture. The coherent property of the waves from space allows us, independent of the observed wavelength, to combine multiple telescopes in large arrays called interferometers. **[note [JB]: include a diagram of diffraction and airy disk pattern]**

Scientists use telescopes to magnify the source of interest, collect more photons, extend the explore-able wavelength range beyond the visible window etc. We generally do not have *a priori* knowledge about the size of an object of interest on the sky. For example, the Sun and the Moon appear to be of the same size on the sky, while in reality their physical sizes **[note [JS]: physical size, radial extent?]** differ by many orders of magnitude. This can be expressed via:

$$\theta = \frac{l}{d} \tag{3}$$

θ is the angular extent of the object, commonly measured in radians or degrees. The natural unit is radians, as will be seen below. If the distance d between the observer and the source, and the linear size of the source l is known, the angular extent θ can be calculated.

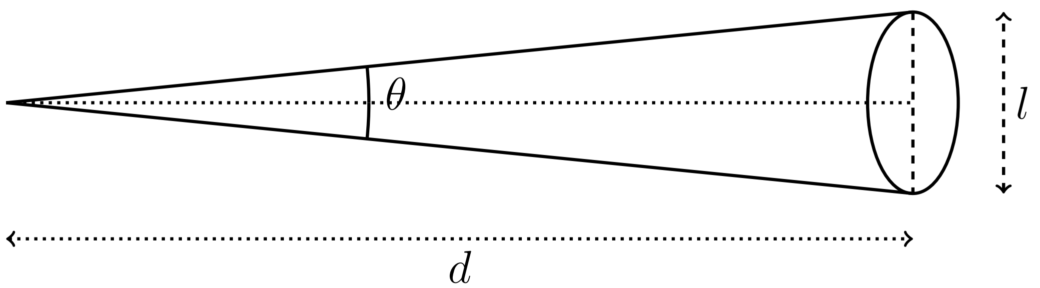

Fig. 2.1.1: Illustration of the angular extent θ of a source of size l at distance d

For a telescope another relevant angle is the angular resolution θ'. This is the smallest angular distance between two sources that can be resolved by the telescope. Assume that an instrument has an aperture size D. The aperture size is the diameter of the light collecting area such as lens or mirror (In case of an interferometer this corresponds to the separation between the individual telescopes.) The angular resolution θ' depends on the wavelength at which the observations take place λ and the aperture size D according to the relation:

$$\theta' \simeq \frac{\lambda}{D} \tag{4}$$

This is known from the diffraction pattern of a single or double slit experiment. Basically verifying the wave nature of electromagnetic radiation. In case of a circular aperture we need to correct this formula slightly by a constant factor of 1.22. This factor is calculated from solving Bessel functions. This is rather complicated and beyond the scope of this textbook.

$$\theta' = 1.22 \cdot \frac{\lambda}{D}$$

In principal there is no limit on the angular resolution θ', as long as we are able to increase the aperture size D. The coherence of cosmic waves allows us to construct telescopes that are in theory unlimited in their ability to observe details of the sources structure. It is indeed theoretically possible to image a beach on a distant extra-solar planet. However, we are limited by our ability to construct telescopes that are large enough.

But what is the relation between the angular resolution and the angular extent? Are they not the same? If this hypothesis was correct it would be possible to derive

$$l \simeq d \cdot \frac{\lambda}{D}$$

This tells us, the shorter the wavelength the finer the disclosed source detail. Which is essentially correct. For a given telescope of aperture D the only way to disclose finer details is to go to shorter wavelengths.

But θ is not equal to θ'. θ corresponds to the linear dimension of the source at distance d, while θ' corresponds to the linear dimension of the source's structure resolvable by the telescope used. For example, take the moon. θ corresponds to the Moon's diameter while θ' is the size of the smallest Moon crater resolvable by the telescope. The angular extent,θ, is a property of the object being observed while the angular resolution θ' is a property of the telescope.

Interferometers are instruments consisting of multiple individual telescopes. Their individual receivers/detectors might have very different instrumental properties but operate at the same wavelength. For an interferometer it is essential that all telescopes observe the same source of interest at precisely the same time. The independently received signals of the source of interest can then be be correlated, or combined. A correlator is (today) a computer that exactly superimposes the amplitude and phases of the signals received by the individual telescopes to a common wave train. Due to the partly large spatial separation of the telescopes, an extremely very high angular resolution (small θ') can be achieved.

Coherence

[note [JS]: I would move this discussion up to the first mention of coherence] The term coherence is derived from Latin *cohaerere*. If we describe the electromagnetic radiation in the wave domain (instead of this you might use the particle domain and speak about photons), coherence means that the phases and amplitudes of different wave trains modulate uniformly in time and space. This behavior is not limited to a narrowly defined range in wavelengths, as in case of a typical LASER (*Light Amplification by Stimulated Emission of Radiation*), but is valid for the whole wavelength spectrum of the source. From the high energy γ-ray radiation to the longest radio waves emitted, all these different kinds of radiation show up with a synchronized modulation of amplitudes and phases.

"...twinkle little star..."

The flickering of the stars visible light is a signature that the coherence of the optical light is disturbed by a very local effect. Carefully inspected however, by a high speed film for instance, one can see that the flickering of a star is nothing else than a diffraction pattern. The Earth's atmosphere destroys the coherence of visible light on its way through the stratos– and troposphere. Pressure and density variations of Earth's atmosphere cause slight statistical variations in the refractive index of the air. Together with strong high altitude winds, the refractive index varies so severely that the visible starlight is deflected in very different directions and thus flickers when observed from the ground. To overcome this, large ground-based telescopes use adaptive optics, which deform mirrors in the telescope set up to account for this. These atmospheric effects are not as strong for longer wavelengths and therefore, beyond the near infrared regime techniques such as adaptive optics are not necessary. In the radio wavelength regime the amplitudes and phases are fixed in time until frequencies of a few tens of MHz. Then, electron density variations in the ionosphere cause statistical variations of the refractive index, similar to the situation in the visible range. In the wavelength range in between however, we observe the sources light limited by the diffraction at the telescope's aperture only.

3 Electromagnetic waves IMAGES

[note [JS]: add illustrations for wave propagation and LC circuit, Antenna shapes]

3.1 Oscillating Circuits to Antennas

The fundamental physics concerned with radio astronomy that of electromagnetic waves. Perhaps you can recall that electromagnetic waves can be precisely described by harmonic functions like sine and cosine and super-positions of both, or Euler's e-function $e^{-i\cdot\omega\cdot t}$. This is great, because you remember the very basic concept necessary to step into this chapter.

The wavelength λ and frequency ν of an electromagnetic wave are connected via the relation $\mathrm{c} = \lambda \cdot \nu$, where c is the speed of light in a vacuum. The energy transferred by a wave is a function of the frequency ν of the wave and is given by $E = \mathrm{h}\,\nu$, where h is Planck's constant. However to fully understand electromagnetic waves it is useful to start from the very basics, mechanical oscillations.

In this section we will set up the wave equation, that is from the mathematical perspective, a differential equation, and investigate this equation and its solution. We will also modify the equation for damped and forced oscillations.

The reception of astronomical signals by a telescope antenna is nothing more than the radiation from an astronomical source causing a forced oscillation in the telescope antenna. The telescope is optimised to receive the electromagnetic radiation from a source of interest. This is done by constructing an oscillating system and setting it up to be optimally resonantly oscillating with the signal emitted by the astronomical source. A system is in resonance when the natural frequency of the receiver is equal to the frequency emitted by the source of interest.

3.2 Wave Production

Waves are a function of both time and position. To generate a wave, energy must be supplied to a system. For a mechanical wave this can be done by displacing a mass from its rest position, such as stretching a spring with a mass at the end of it. For an electromagnetic wave, this is done by accelerating a charged particle. In the ideal case where energy is conserved, the system will oscillate at its natural frequency ω_0. However, if the energy is not conserved, then the oscillation is damped. This damping is described by a *damping constant* δ. The larger the damping the more of the initial energy is transferred into heat. The quicker the system comes to rest. If a periodic oscillating force is applied to the system from outside, the receiving system's motion is described according by a *forced oscillation*.

In the real world losses always occur in systems. Energy is converted into heat. Therefore, all oscillations encountered in nature are forced and damped oscillations.

This relatively simple simple explanation of damped and forced oscillations can nearly explain everything from the telescope dish to the backend of the receiver chain. The reflection of an electromagnetic wave at the mirror's surface is a forced damped oscillation for the electrons in the metallic surface. Cables, which transfer information from the receiver to the computer can be approximated by resistors R, capacitors C and inductors L. The combination of these three components is known as an oscillating circuit. While the capacitor and the coil would oscillate *ad infinitum* the resistor releases energy into the environment via heat conduction and radiation. Also in space the same physical processes are seen at work. For example when an electromagnetic wave such as radiation from a star passes through a gaseous cloud, the processes happening are physically described by a forced damped oscillations, where the radiation causes the gas molecules to oscillate.

3.3 Undamped oscillation

Consider the most simple harmonic oscillation, a simple pendulum, made of a weight on a string. Consider the forces acting on the pendulum. Primarily it is the force of gravity which acts on the pendulum here. Therefore the gravitational acceleration g can be substituted for the acceleration in Newton's second law $F = m\,a$. If the pendulum is displaced from its rest position by an angle ψ then the force acting on the pendulum is

$$F_{\mathrm{T}} = -m\,g\,sin\psi \tag{5}$$

Consider the angle of deflection ψ. Even at $\psi \simeq 60\,\mathrm{deg}$, $\sin(\psi) = 0.866 \simeq 1$ is still sufficiently comparable to 1. This is the basic idea beyond the small–angle approximation. If something is sufficiently close to unity, it can be approximated by a series expansion. Here, we have to inspect the series expansion of the sine function.

$$\sin\psi = \psi - \frac{\psi^3}{3!} + \frac{\psi^5}{5!}... \tag{6}$$

Since 3!=6, this means the second term is only a sixth of the first one. 1/6 is 0.166 and according to the above we make a mistake of only 17% by inserting ψ instead of inserting $\sin(\psi)$. this is a good approximation for most situations. Therefore, Eq. 5 can be rewritten as:

$$F_{\mathrm{T}} = -m\,g\,\psi \tag{7}$$

Now, consider more closely the left hand side of the equation and turn it into an expression describing the force that occurs due to the deflection. The deflection is a linear measure along the arc of the oscillation $l \cdot \psi$, where l is the length of the pendulum. By calculating the derivative of the displacement the velocity of the pendulum $l\frac{d\psi}{dt}$ as well as the acceleration $l\frac{d\psi^2}{dt^2}$ can be found. Inserting the acceleration it follows for the left hand side that

$$\frac{d\psi^2}{dt^2} = -\frac{g}{l}\,\psi \tag{8}$$

$$= -\omega_0\,\psi \rightarrow \tag{9}$$

$$0 = \frac{d\psi^2}{dt^2} + \omega_0\,\psi \tag{10}$$

The last equation is called a second order homogeneous differential equation. This means that it describes the second derivative of a variable as a function of the variable. In this case the variable is displacement, so the second derivative is the acceleration. All undamped harmonic oscillations can be simplified to this form. The Eigenfrequncy ω_0 depends on two constants, namely the length l of the pendulum and the gravitational acceleration g. So, if we like to change the oscillation frequency we need to change the length of the pendulum. Later we change the capacity, inductivity and resistivity to optimize the Eigenfrequency of our receiving system for our purposes, when dealing with electromagnetic oscillations.

The solution of this equation is a harmonic function:

$$\psi(t) = A\,cos(\omega_0\,t + \phi) \tag{11}$$

Where A is the amplitude of the oscillation, ω_0 is the natural frequency, and ϕ is the phase.

3.4 Damped oscillation

As described previously, real systems are subject to energy losses. Mostly this energy is lost through being converted to heat. Therefore, a damped oscillation is one that requires an additional force term, a friction force or damping force F_{D}. Damping is proportional to the velocity $v = l\,\dot{\psi}$. It is given by:

$$F_{\mathrm{D}} = -c\,v = -c\,l\,\dot{\psi} \tag{12}$$

where c is the velocity of the wave through the medium. The higher the velocity of the pendulum, the more frequent the pendulum interacts with its environment. For example air molecules colliding inelastically with the pendulum. The homogeneous differential equation (Eq. 10) can be rewritten as:

$$0 = -m\,l\,\frac{d\psi^2}{dt^2} - c\,l\,\frac{d\psi}{dt} - m\,g\,\psi \tag{13}$$

This can be modified using the following: **[note [jb]: what are these values?]**

$$\frac{c}{m} = 2\,d$$
$$\frac{g}{l} = \omega_0^2$$

then it follows:

$$0 = \frac{d\psi^2}{dt^2} + 2\,d\,\frac{d\psi}{dt} + \omega_0^2\,\psi \tag{14}$$

A solution of this equation has the form $\psi(t) = A\,e^{\lambda\,t}$. Where A is the amplitude of the oscillation and λ is the oscillating frequency of the system. It is different from the Eigenfrequency ω_0 because of the damping, but close to its value. Inserting the solution into the equation gives:

$$0 = \lambda^2 - 2\,d\,\lambda + \omega_0^2 \tag{15}$$

This has the following solutions:

$$\lambda = d \pm \sqrt{-d^2 - \omega_0^2} \tag{16}$$

$$= d \pm \omega_0\,\sqrt{\left(\frac{-d}{\omega_0}\right)^2 - 1} \tag{17}$$

$$\tag{18}$$

The quantity $\frac{d}{\omega_0} = D$ is called the *damping ratio* of the system. The second line of the equation above is introduced to show that the oscillation frequency of the system is de-tuned from the Eigenfrequency ω_0. It makes clear that this is due to the dampling D. Inserting this into the solution:

$$\lambda(t) = A\,e^{\lambda\,t}$$

$$= A\,e^{\left(-d \pm \omega_0\,\sqrt{\left(\frac{d}{\omega_0}\right)^2 - 1}\,t\right.}$$

$$= A\,e^{-d\,t}\,e^{-\omega_0\,\sqrt{D^2-1}\,t}$$

The first e-function only describes the damping of the amplitude. The second e-function represents the oscillation. The first one is an envelope to the oscillation and is also described as *amplitude function*.

If the damping ratio D is small, i.e. less than 1, the system oscillates, this is known as the *oscillation case*. However, this oscillation is not equal in frequency to the natural frequency ω_0 but slightly lower, due to the damping. The magnitude of this frequency shift is given by the expression in the root of the equation above. If $D < 1$, then the root becomes negative, this we can describe mathematically simply by changing to the complex representation of numbers.

$$\begin{aligned}\lambda(t) &= A\, e^{-d\,t}\, e^{-\omega_0\,\sqrt{D^2-1}\,t} \\ &= A\, e^{-d\,t}\, e^{-\omega_0\,\sqrt{-1(1-D^2)}\,t} \\ &= A\, e^{-d\,t}\, e^{-i\omega_0\,\sqrt{1-D^2}\,t}\end{aligned}$$

In the case of strong damping $D \geq 1$ the solution is real and depending on its magnitude the amplitude creeps to the rest position without oscillation. This is called *creep case*.

So, on our tour to bring a radio telescope to the sky, we already build a receiver. It has some losses, due to physics, but it is able to receive radiation. Now let us walk outside and direct our telescope towards a celestial object.

3.5 Forced Oscillations

In the case of forced oscillation, an external force acts on the oscillating system. Here, the celestial source emits electromagnetic waves and accelerate the electrons in the dipole of the telescope antenna. Therefore, the differential equation of a damped oscillation Eq. 14 is used and insert an exciter or force. This is the electromagnetic wave from space.

$$F\, e^{-i\,\omega\,t} = \frac{d\Psi^2}{dt^2} + 2\,d\,\frac{d\Psi}{dt} + \omega_0^2\,\Psi \tag{19}$$

The solution for such an oscillation is

$$\Psi(t) = A\, e^{-i(\,\omega\,t\,+\,\phi)} \tag{20}$$

Where A again describes the amplitude of the oscillation, ω the incident frequency of the source and ϕ the phase shift. Phase shift is the difference in the position of a wave at a given point in time compared to the position of the same wave at the same point in time in another location. Combined they describe the delay between the excitation of the system and its response to the excitation. The phase angle varies between $0 \leq \phi \leq \pi$. t describes the time. Inserting the solution for the undamped oscillation gives:

$$
\begin{aligned}
F &= -A\,\omega^2\,e^{-i\,\phi} + 2\,d\,i\,\omega\,e^{-i\,\phi} + \omega_0^2\,A\,e^{-i\,\phi} \\
&= A\,e^{-i\,\phi}\left(\omega_0^2 - \omega^2 + i\,2\,d\,\omega\right)
\end{aligned}
$$

The equation can now be solved for the amplitude:

$$
A = \frac{F\,e^{i\,\phi}}{\omega_0^2 - \omega^2 + i\,2\,d\,\omega} \tag{21}
$$

This can still be split and put into the normal form:

$$
A\,e^{-i\,\phi} = \frac{F}{(\omega_0^2 - \omega^2)^2 + (2\,d\,\omega)^2}\,e^{i\,\arctan\left(\frac{2\,d\,\omega}{\omega_0^2 - \omega^2}\right)} \tag{22}
$$

The magnitude of this complex number represents the amplitude of our forced oscillation, while the exponential part is the phase position. At low damping, the largest response of the oscillatory system is close to the Eigenfrequency $\omega = \omega_0$, then the phase position is shifted by $\frac{\pi}{2}$. In case of strong damping, the largest response is already reached below a phase shift of $\frac{\pi}{2}$.

So what did we learn? To build a radio telescope we need to build an oscillating system. Because of losses, it will not operate perfectly. For our science, we need to minimize the losses/damping and to tune the system's Eigenfrequency to those of the source of interest. The terminus “bandwidth” becomes important. The larger the bandwidth the more energy is collected by our receiving systems. But far away from the Eigenfrequency the response of the receiver is not that ideal. This is why a radio observatory provides multiple receivers to the scientific community. Each is optimized to receive the emission of a certain atom or molecule space. Today it is not feasible to construct a single receiver covering the whole accessible frequency range at the site to the telescope.

3.6 The LC Resonant Circuit

A forced oscillation describes a real receiving system and all its components amazingly well! An electromagnetic wave from a source in space acts as a periodic exciter to the antenna in the telescope. This concept carries further than just to the antenna. The incoming electromagnetic wave also causes forced oscillations in the electronic circuits of the receiving system. The forced oscillation from the electric oscillating circuits can be extended to the cables, and can be treated as a coupling of LCR oscillating circuits. Thus, it is worthwhile to look at LCR oscillatory circuits in detail.

The excursion into the world of differential equations of damped forced oscillations may be a bit unpleasant but quite instructive.

In the most simple case, the oscillating circuit consists of a capacitor C and an inductor L. The plates of the capacitor are charged and then the system is left to regulate itself. In the lossless case, the capacitor discharges, a maximum current flows through the inductor in the next phase, and then the plates of the capacitor would charge again with reversed charge sign. Then the process repeats in the opposite direction. In the lossless case, the system oscillates at its natural frequency forever.

In the real case, however, there are losses. All the cables and electronic components have resistances, which are summarized with the electrical resistance R.

At the time $t = 0$, there are no charges Q in the system, so the sum of all voltages on the components of the oscillating circuit are measured and use the Kirchhoff's second law which states that the sum of potential differences (or voltages) around any closed loop is zero

$$V_C + V_R + V_L = 0 \tag{23}$$

V_C is the voltage across the capacitor, V_R is the voltage across the resistor, and V_L is the voltage across the inductor. The values for each of these voltages is given by the equations below.

$$V_L = L \cdot \frac{dI}{dt} \tag{24}$$

$$V_R = R \cdot I \tag{25}$$

$$V_C = \frac{Q}{C} \tag{26}$$

Where I is the current, L as inductance and C as capacitance. Inserting these simple equations into Eq. 23:

$$\frac{Q}{C} + R \cdot I + L\frac{dI}{dt} = 0$$

To obtain an equation depending only on a single variable, namely Q, the charges in the system, we substitute $I = \frac{dQ}{dt}$, as current is a measure of the flow of charges.

$$\frac{Q}{C} + R \cdot \frac{dQ}{dt} + L\frac{d^2Q}{dt^2} = 0$$

and rearrangement the terms in the sense that its structure look like Eq. 14 yields:

$$\frac{Q}{L \cdot C} + \frac{R}{L} \cdot \frac{dQ}{dt} + \frac{d^2Q}{dt^2} = 0$$

This is a differential equation for a damped oscillating LCR system. Its identical in form to a dampened pendulum. The dampening corresponds to an energy loss over the resistance. In case $R = 0$ the system is not damped and oscillates with its Eigenfrequency $\omega_0 = \frac{1}{\sqrt{LC}}$. However, due to the damping the system is detuned, it oscillates with the frequency ω and not with ω_0. Eventually the damping d (Eq. 14) corresponds to $d = \frac{R}{2 \cdot L}$.

The solution of this equation is only ready known from Eq. 22.

$$\omega_{1/2} = -\frac{L}{2C} \pm \sqrt{\left(\frac{L}{2C}\right)^2 - \frac{1}{LC}} \tag{27}$$

$$Q(t) = e^{-d \cdot t} Q_0 \left(e^{\sqrt{d^2 - {\omega_0}^2} \cdot t} + e^{-\sqrt{d^2 - {\omega_0}^2} \cdot t} \right) \tag{28}$$

As long as the d is low $d^2 - {\omega_0}^2 < 0$ the system oscillates periodically, but the oscillation frequency in the damped case is always lower $\omega < \omega_0$. All results derived from above you can apply here.

Task accomplished!

Exercise 3.1

1. Derive the differential equation of an LC–circuit consisting of a resistor R, an inductor L, and a capacitor C. Use Kirchhoff's second law.

2. What is the Eigenfrequency ω_0 of this system? Is that the frequency with which the system oscillates?

3. Show that the following function is a solution of the differential equation for $R = 0$ by inserting it:

$$\Phi(t) = A\, e^{-i(\omega t + \phi)}$$

3.7 Directionality: Waveguides and Reflector Antennas TBW

3.7.1 Waveguides

3.7.2 Paraboloid Antennas

4 Radiometers IMAGES

[note [JS]: Radiometer sketch => Essentials of radio astronomy?] In radio astronomy the term receiver is used to describe a collection of devices and components that are used to turn detect, amplify and measure the incoming radio wave so that it can be used to conduct science.

The “front end” and “back end” are the most commonly discussed components of a receiver as they are what astronomers interact with the most, as they are the first and last devices in the receiver chain. The front end is the part which interacts directly with the cosmic wave, turning incoming photons into an electrical signal, while the back end is the spectrometer or the continuum flux recorder, which produces the data in a form ready to be analysed. there are many other components in between these two such as the “feed”, “bandpass” and “integrator”. These components, all acting together are also commonly called the radiometer.

4.1 Overview

A radiometer, as sketched in Fig. 4.1.1, consists of at least five components

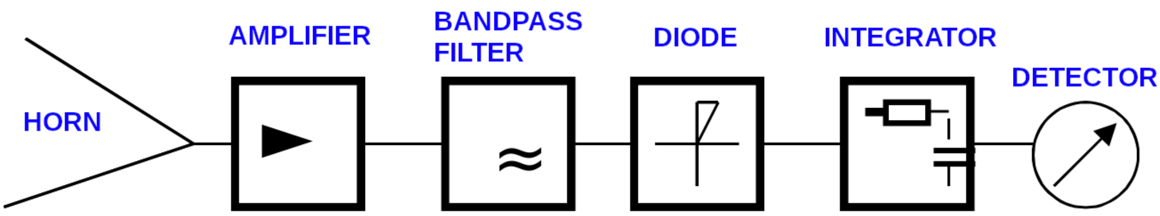

Fig. 4.1.1: Sketch of a radiometer set-up. The dipole and horn are shown as one.

- **Horn:** The horn is used to guide the radio wave towards the dipole, the receiving element. The horn’s shape needs to have an optimized geometry and be adapted to the optical parameters of the mirror system.
- **Dipole:** The dipole is the components that collects the photons from the incoming radio wave and converts it into an electrical signal. It is positioned within the horn. Dipoles are used in the MHz to GHz frequency regime to collect photons. At shorter wavelengths, such as in the optical regime more sophisticated technology is used to collect photons such as CCDs. The incoming radio wave accelerates the electrons within the dipole which is made from an excellent conductor, such as copper. The length of the dipole is chosen to optimise the reception of radio waves of certain wavelengths (keyword: resonance), beyond that

Eigen–frequency or Eigen–wavelength the other neighboring waves are detected with a lower efficiency.

- **Amplifier:** The voltages induced in the dipole by the received radio wave are very small. They must be amplified immediately so they are not lost as the signal is carried through the receiver to other components. Semiconductor elements such as high electron mobility transistors are used as amplifiers. The signal is amplified on the order of a few thousand to a few million times of the input signal.

- **Bandpass Filter:** Radio receivers detect multiple signals at the same time and therefore, it is necessary to filter out signals which are not the signal of the source of interest. Artificial signals from telecommunication or electronic devices are the most severe sources of confusion with cosmic signals. These are called radio frequency interference (RFI). They outshine the cosmic sources by a significant amount. The bandpass filter removes these unwanted signals by only allowing the signal within a certain frequency window to pass through. this frequency window can be chosen so that the unwanted are blocked and filtered out. Bandpass filters are commonly placed directly after the amplifier, however, in some cases such as when there is strong RFI, they are placed in front of the amplifier, however this is not ideal as it raises the system temperature, increasing noise.

- **Diode:** Since the voltages induced in the dipole oscillate between positive and negative values, they sum-up to zero. To solve this issue, the voltages of the signal are squared. This is achieved using a semiconductor device called a diode. Squaring warrants that the sum, I, is proportional to the voltage amplitude squared, $I \propto V^2$. At a certain operating point the characteristic curve of the I-V line is shaped like a parabola. A voltage bias move the amplified voltage variations to this operation point and squares them.

- **Integrator:** After squaring the signal to make all the voltages positive, the voltage variations must be summed up. this is done with an integrator. An integrator can be made from a combination of resistor and capacitor. The charges generated by the diode current are accumulate on the capacitor, which is charged.

- **Read out:** Finally, a device is needed to read out the amount of charges stored by the integrator, this is commonly done by a computer after passing the stored charge through an analogue to digital converter.

[note [JB]: include diagram of pipeline]

4.2 Semiconductors

Two of the devices in a basic radiometer set up are made from semiconductors. The amplifier, consisting of a transistor and the detector, consisting of a diode.

Semiconductors are devices that, as the name implies, are neither conductors nor insulators. To understand these devices and their value for radio astronomy, it is useful to recall the band model in insulators, conductors and eventually semiconductors.

The band model distinguishes between a valence band and a conduction band. The valence band is lower in energy than the conduction band. The valence band is the "host" of the charges. A moving charge in the valence band results in a current. Within the band model the energy gap between the valence and the conductor band is what defines the properties of the semiconductor.

If there is no gap, the device simply acts as a conductor. In the case where the gap is large, no current is able to flow and the device acts as an insulator. In between both extremes is the semiconductor. One needs a certain amount of energy to transfer a charge from the valence to the conductor band within a semiconductor. Without any excitation from the outside the charges are going to remain at the minimum energy level within the valence band. At the same time the conducting band is empty and no current is detectable. In the case that sufficient extra energy is transferred to the system, sufficient to bridge the gap between valence and conductor band, the charge can move freely within the conductor band.

Due to the energetic difference between both bands, charge carriers can only be transferred from the valence band to the conduction band by means of additional energy, added to the system from the outside. One way is to heat-up the semiconductor. In this case the thermal motion (Maxwell–Boltzmann velocity distribution) of the charge carriers increases. Those with the highest energies have the highest probability to cross the energy gap between valence and conductor band, then the semiconductor becomes a conductor. In those cases, when thermal heating makes a conductor out of the semiconductor the device is called a thermistor.

Semiconductor are made from elements found in the atomic group 14 of the periodic table. They all have four (valence) electrons which have a stable connection within their solid state lattice (crystal structure). Germanium and silicon are examples of elemental semiconductors. By adding elements of group 13 (with three valence electrons) and group 15 (with five valence electrons), the semiconductor properties can be changed. This "doping" of the elemental semiconductor with donors (group 15 elements, p-doping) and acceptors (group 13 elements, n-doping) cause a change in the energy gap between the valence and conduction bands. The doping can increase or decrease the gap.

The combination of n- or p-doped semiconductors allows the construction of devices which are sensitive to certain energies, e.g. the irradiation of photons (CCD cameras). Here the doping is optimized in such a way that visible light can lift electrons from the valence to the conduction band and yield an image proportional to the photon energy can be produced.

4.3 Transistor and Diode, Voltage and Current

In the case of radio astronomy, the photon energy is too low to lift an electron from the valence band to the conduction band. Semiconductors are used in transistors to amplify the radio astronomical signal or in diodes to convert the voltage fluctuations into a current. The trick is in both cases to apply the voltages in such a way to make best use of the pnp, npn and pn structure of the doped semiconductor.

First, consider the diode. In the case where the battery is connected such that the positive pole is connected with the n-doped side of the diode and the negative pole with the p-doped side, a current flows. If the polarity is reversed, the current stops. The current–voltage characteristic of a diode describes this behavior.

The polarity as well as the voltages can be changed and chosen for specific scientific tasks. Applying a bias voltage to the diode allows for a squared current function to be extracted from a varying voltage signal around a bias point. The voltage bias affects the charge structure within the pn transition zone. Changing the voltage across the diode is equivalent to modify the transition zone. The mobility of the charge carriers can be adapted appropriately.

Transistor Base Thickness

[note [JB]: Include basic diagram of transistor] Consider now the transistor. Here, we have something like two diodes connected with opposite polarity. This means that regardless of the voltage configuration, a current always flows. Consider the pnp transistor for example. Only a low positive voltage is applied on in input side of the transistor while the output side is connected to the ground. P-charges from the input side would not move further than into the n-doped base. The p-charges “fill” the “holes” in the n-doped layer. Now assume that the positive voltage is so high, that only the gaps are filled. In this case the n-doped base acts like a barrier for the p-charges. It stops their motion towards the output. This is a stable situation.

Now shrink the size of the base, the barrier function becomes less effective at stopping the motion of the p-charges. But what is gain from this procedure? A thinner base means fewer n-“holes” are available for the p-charges. Consider that all p-charges fill all n-holes and a stable situation is reached. Now imagine, that the base is connected with the dipole of a receiver from a radio telescope. The varying signal disturbs this stable configuration, the barrier becomes unstable. Positive charges move through the n-doped base, because of the unstable configuration to the ground. Positive charges from the input fill quickly the n-holes in the base. Like a valve, the weak cosmic signal received by the dipole, connected to the transistors base, modulates the transition of the p-charges to the ground. A few charges at the base modulate the strong flow between input and output. Thus, the transistor is used as an amplifier for the incoming cosmic signal.

The combination of doped semiconductors and bias voltages allows for these special devices to be used for our purposes. In all cases a bias voltages is applied at such a levels that the charge carriers gain sufficient energy to move from the valence to the conduction band. In radio astronomy the photon energies are about a fraction of a millions of that of the visible light.

Noise at Low Photon Energies

The light you see with your eyes has between 1.5 and 3.2 eV per photon. Our eyes have evolved to optimally receive light at the maximum of the Sun's thermal radiation. This thermal radiation corresponds to black–body temperatures in the range of 3000 K to 30.000 K. In radio astronomy the photon energies are much lower. The famous 21-cm of neutral atomic hydrogen corresponds to a few $E_{21-\mathrm{cm}} \sim 10^{-6}$ eV. Using Wien's displacement law the corresponding temperature of the black–body can be estimated. It is a fraction of a million times cooler than a stellar photosphere. A quick thought makes obvious that all bodies in the universe radiate thermally and confuse the signal of the 21-cm line. The physical lower temperature limit is given by the 3K CMB radiation. Nothing can be cooler in nature than 3K, except laboratory experiments. But also the 3K photons have significantly more energy that the 21-cm line photons. Next to the few photons from our source of interest a huge number of unrelated photons flow into our receiver. Radio astronomy is the science of cosmic noise.

Consequently, radiation not associated with the cosmic signal entering the receiver, also result in currents in the amplifier. These currents are much higher than that of the cosmic signal. A way to detect the faint objects signal from that of the unrelated thermal radiation is by statistics.

Cosmic Noise

Since all bodies within the universe are warmer that 3 K, they produce bright radio noise. Thermal noise, superimposed with the radio signal of source of interest, is unavoidable. Unrelated thermal noise however can be differentiated from the cosmic signal by statistical means. This is done by doing two things:

- Reducing the thermal noise to a minimum. Since most of the universe is out of reach for us, we can do a great job by simply cooling our equipment.
- Optimizing the stability of the receiving system, in the sense of long term temporal stability. This means in particular the the variance of the system noise is not a function of time. For this purpose it is necessary to check the system performance regularly by calibration measurements in between the observation.

The only way to disclose the faint cosmic signal from the unrelated but overwhelming strong signal, is using the fact that the strong signals vary statistically and sum up to zero. Because the absolute majority of these unrelated signals originate from thermal radiation, dominantly in the amplifier itself. Integrating the thermal noise signal sufficiently long leads to them being averaged out, while the cosmic signal remains. We need to integrate the signal until the noise becomes so low, that the signal sticks out.

Central Limit Theorem and Radiometer Equation

The contribution of unrelated noise within our radio astronomical measurement is caused by a large number of photons. When dealing with large samples of a population (in this case photons) the central limit theorem can be applied. The central limit theorem states that for large samples of a population, the mean of the sample will be normally distributed, even if the population is not normally distributed. Considering this for a sample of photons where N is the number of photons the we find:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} \tag{29}$$

$\sigma_{\overline{X}}$ is the mean value of the sample and σ is the variance from the mean μ and N is the number of probes. Translated into our radio astronomical measurement: σ is the systems intrinsic uncertainty, mainly due to the system temperature T_{Sys}. $\sigma_{\overline{X}}$ is the actual uncertainty of the measurement with our equipment. N corresponds to the number of photons received. Now, N is the product of two parameters, the bandwidth $\Delta\nu$ and the integration time t_{int}. The larger the bandwidth, the more photons are received. The same holds true for the integration time. So $N \propto \Delta\nu \cdot t_{\text{int}}$. The equation above can be rewriten as:

$$\Delta T = \frac{T_{\text{Sys}}}{\sqrt{\Delta\nu \cdot t_{\text{int}}}} \tag{30}$$

This equation is called the *radiometer equation*. The uncertainty ΔT is a measure for the quality and sensitivity of our radio astronomical measurement. The longer we integrate the lower the uncertainty. Obviously the same holds true if we reduce the system temperature T_{Sys} to a minimum.

As an observer, you need to identify the bandwidth $\Delta\nu$ and to determine the systems temperature. At long wavelength T_{Sys} depends on the instrumentation only, at frequencies higher than a few tens of GHz the system temperature limit is determined by the atmosphere. Here, the water vapor content is the important factor.

Exercise 4.1

Your receiver has a receiver temperature of 150 K with a standard deviation of 10 K. You measure a signal with $T = 175\,\mathrm{K}$. Assume this is just a statistical deviation. How likely is this to happen? Hint: Standard normal dsitribution on Wikipedia

How strong would the signal temperature have to be, to fall into the top 5% of the distribution, when $T_{\mathrm{rec}} = 150\,\mathrm{K}$ and $\Delta T = 30\,\mathrm{K}$?

Exercise 4.2

One way to calibrate your systems and to determine the system temperature, or more correctly the amplifier temperature, is to perform a hot–cold calibration. To do so boiling water is used as the hot-load and then ice at its melting point as the cold–load in front of the receiver. Show that:

$$T_{\mathrm{rec}} = \frac{T_{\mathrm{hot}} - Y \cdot T_{\mathrm{cold}}}{Y - 1} \tag{31}$$

Hint: $P_{\mathrm{hot/cold}} = (T_{\mathrm{hot/cold}} + T_{\mathrm{rec}}) \cdot \mathrm{const.}$ and $\frac{P_{\mathrm{hot}}}{P_{\mathrm{cold}}} = Y$. You derive $Y = 1.26$ from your measurements. What is the receiver temperature T_{rec}?

Amplifier Cooling

The amplifier in the form of a semiconductor element plays a special role in the radiometer. By the specific construction of this device, amplifications of the signal in the range of a few thousand up to a few million can be achieved. This also means that the unrelated thermal fluctuations within the transistor lift the charges from the valence to the conduction band, completely outshining the faint radio astronomical signal. Therefore it is necessary to minimize the thermal motion in the amplifier.

This is done by cooling the amplifier. For this purpose, liquid helium or nitrogen is used.

The thermal properties or the thermal noise of the components of the receiver beyond the amplifier are of secondary importance, since the amplifier boosts the signal thousands to millions of times above the noise of the following components (see Sect. 4.2).

Bandpass Cooled and Uncooled

The bandpass filter, next to the amplifier, can remain without any cooling. This is different in the case that artificial radio interferences affect our frequency range of interest seriously. Here, a bandpass filter is placed in-front of the amplifier. Here, also the bandpass filter needs to be cooled properly. The bandpass filter consists physically out of an optimized combination of RC circuits. By exchanging the position of the resistor (R) and capacitor (C) a high or a low pass filter can be constructed. The bandpass filter consists therefore out of a combination of a low and a high pass filter.

Diode

In the next step the voltage fluctuations reach the diode. The diode is a semiconductor device consisting of a combination of n- and p-doped material. Depending on the doping, the size of the energy gap between both doped semiconductor materials can be optimized. The diode has an "operating point" at a certain voltage, dependent on its doping. Voltage variations around this operating point cause a quadratic change of the diode current. This quadratic characteristic is used to create "intensities" out of "amplitudes". The charges of the generated currents are stored on a capacitor and read out by a measurement device.

Radiometer Input

The signal we are looking for is very small compared to the unrelated radiation entering our radiometer. What contributes to the signal we want to study?

Firstly there is the thermal noise of the receiver itself, especially the first amplifier $T_{\mathrm{Ampl.}}$. Then the contribution of cosmic microwave background is always present T_{CMB}. Also, the Milky Way and all extragalactic sources radiate into the telescope illuminating the antenna pattern $T_{\mathrm{gal+extragal.}}$. Then, a small fraction of the receiver usually still looks beyond the outer rim of the dish. That portion receives radiation from the ground at the telescope's site. It is called *spill over* T_{spill}. Thus we can write:

$$T_{\text{Sys}} = T_{\text{Ampl.}} + T_{\text{CMB}} + T_{\text{gal+extragal.}} + T_{\text{spill}} + (T_{\text{Atmosphere}})... \tag{32}$$

The $T_{\text{Ampl.}}$ can be reduced as much as possible by cooling the device. But what about the other components of the radiometer?

If each individual device from which the radiometer is made of is considered and attribute to each of them a specific temperature T_{i} and a gain of the signal G_{i}, the noise contribution of this chain relative to that of the first element can be quantified:

$$T = T_1 + \frac{T_2}{G_1} + \frac{T_3}{G_1 \cdot G_2} + ... \tag{33}$$

Since G_1 is in the range of 10^3 to 10^6 the thermal noise contribution of the following components is significantly reduced by its amplification. We only need to optimize and actively cool the first amplifier. The following devices can be operated at room temperature.

Temporal Stability: The Dicke receiver

In addition to these statistical fluctuations, electronic components may also exhibit systematic fluctuations that we have not considered yet. For example, the amplifier may exhibit fluctuations in gain ΔG that systematically limit the quality of our measurements. These fluctuation are not easy to differentiate from variation of the system's temperature.

$$\sigma_{\text{gain}} = T_{\text{Sys}} \cdot \frac{\Delta G}{G} \tag{34}$$

Therefore, the radiometer equation needs to be modified

$$\Delta T = T_{\text{Sys}} \sqrt{\frac{1}{\Delta\nu \cdot t_{\text{int}}} + \left(\frac{\Delta G}{G}\right)^2} \tag{35}$$

These fluctuations can be limited by using what is known as a "Dicke" receiver. A Dicke receiver works by measuring an ON and OFF position in very short time intervals, usually by means of two receivers. The time interval must be shorter than the time scale on which the amplification varies. Switching times between 10 cycles per second to 1000 are common. Its clear that a big single dish cannot be moved that fast, therefore a second feed is needed and connected it to the same radiometer chain. Feed 1 is permanently ON the source while feed 2 is OFF. A precisely clocked electronic switch is used. The difference between ON source T_1 and off source T_2 is evaluated, because this is a measure for ΔG.

$$\sigma_{\text{Gain}} \simeq (T_1 - T_2) \cdot \frac{\Delta G}{G} \ll T_{1/2} \cdot \frac{\Delta G}{G} \tag{36}$$

Since only $(T_1 - T_2) * \frac{\Delta G}{G}$ and $T_1 \simeq T_2$ are being evaluated the gain variations are down-scaled accordingly.

The problem with a Dicke receiver is that only half of the telescope time is spent observing the source, half of the time the OFF position is being observed. So, the final noise level reached with the Dicke receiver is only

$$\Delta T_{\text{Dicke}} = 2 \cdot \Delta T \tag{37}$$

Since since the source is only being observed for half the time with a Dicke receiver, the sensitivity is a factor of two less than with a simple radiometer. Therefore, very sensitive measurements can be performed but require twice the integration time. The Wilkinson Microwave Anisotropy Probe (WMAP) has by the instrumental construction used the Dicke Ansatz. They arrange two identical mirror and receivers looking in opposite directions. The whole satellite performs a rotation around its major axis within 129 seconds so that within about one minute the receivers watch the same portion of the sky. The radiometer amplifier chains are organized in such a way that two strings amplify both mirrors in parallel. By exchanging the telescope hemisphere, also the amplifier chains are swapped. Because of that construction no observing time has been lost, because the telescopes are always on source.

Exercise 4.3

1. What is a radiometer and how does it work?

2. Which components does a radiometer have? Sketch the set-up.

3. Which purpose does the diode of a radiometer serve?

4. Justify using a transistor as an amplifier.

5. Why is it necessary to cool the pre-amplifier? Why are not all components cooled?

6. In which situation could it be beneficial, to place the bandpass BEFORE the pre-amplifier? Why is this not an ideal set-up?

7. What physical process causes the skin-effect?

Nerd Knowledge: Frequency mixing

[note [JS]: frequency mixing]

5 Backends TBW

So far we have discussed how to obtain a signal with our radio telescope. The dish itself and the radiometer is what is commonly referred to as the *frontend* of a telescope (the part that points towards the source). Now we shall take a look at how this signal is measured and turned into data. The part of the telescope that performs this task is referred to as the *backend*.

Depending on the research question and the telescope there can be a number of backends to choose from, the most important being continuum backends and spectrometers.

5.1 Continuum backends

Continuum backends integrate the signal over the entire bandwidth of the receiver and give a single observed Temperature for each integration time step. They are useful for imaging, as a large bandwidth also means lower noise at lower integration times. (Remember Eq. 107!) This makes it possible to do many pointing observations in a short amount of time, which can then be mapped to create an image.

[note [JS]: Talk about squaring]

[note [JS]: Grid and OTF mapping, basket weaving?]

However a lot of information about the physical processes behind the emission is lost. The spectrum of a source tells us a lot about the circumstances under which the light we see is emitted. This is where spectrometers come in useful.

5.2 Spectrometers

A spectrometer integrates the signal over many narrow frequency ranges. To do so many modern spectrometers make use of Fast–Fourier–Transform (FFT).

Fourier transforming our signal can be understood as transitioning from the time to frequency domain. It "disentangles" the waves of different frequencies that make up our signal, and returns the spectrum: The amplitude for each frequency present.

The conceptual idea is, that every (time-dependent) input signal can be expressed (or at least approximated) as a sum of sinus waves of different frequencies. You may be familiar with this concept from music, where each note corresponds to a frequency, and the sum of all notes played becomes a song.

Our goal here then is to take a song, and find out what notes it is made of (frequencies of the sinus waves) vs. how loud they are a played (amplitude of the sinus waves). This can then be plotted in what is referred to as a spectrum (amplitude vs. frequency).

You can think of a spectrometer as an row of buckets, each bucket collecting the part of the signal of a narrow frequency range and continuously adding the recorded amplitudes until you read out the value at the end. By plotting the final content of the buckets over their central frequency you obtain the spectrum.

[note [JS]: include example spectrum]

From a mathematical standpoint the transformation between the signal amplitude vs. time to the amplitude vs. frequency plot (the so-called Fourier-Transformation) is expressed as an integral:

$$\tilde{f}(\nu) = \int_{-\infty}^{+\infty} f(t) \cdot \exp(-i\,2\pi\,\nu\,t)\mathrm{d}t$$

where $\tilde{f}(\nu)$ is our spectrum, and $f(t)$ our recorded signal, with the reverse transformation reading:

$$f(t) = \int_{-\infty}^{+\infty} \tilde{f}(\nu) \cdot \exp(i\,2\pi\,\nu\,t)\mathrm{d}\nu$$

In a spectrometer this transformation is done electronically using an array of bandpass filters, or digitally with a FFT chip set, and the resulting spectrum is binned and each bin integrated over the recording time.

In case of a small radio telescope like the one used during this course, the FFT and recording of the spectrum is usually done automatically by your computer, and you can directly save and work with the output file to analyze your data.

Exercise 5.1

Most smartphones, laptops, and PCs nowadays have audio mixers pre-installed. This software allows you to not only change the overall volume (amplitude) of your music, but also change the volume of selected frequency bands, usually in the form of sliders for bass, treble, and everything in between.

For this exercise locate the audio mixer on your device of choice, play a song, and change the settings of the sliders. What changes?

Now set all sliders to zero except the lowest ones (bass).

1. What do you hear? Can you still recognize the song?
2. What does this have to do with spectra and Fourier transforms?

Experiment 5.1

On the website spectrogram.sciencemusic.org you can make music by drawing on the spectrum on the screen. The "waterfall plot" that you see is the spectrum over time, the color corresponding to the amplitude at the respective frequency. Each line corresponds to a single spectrum.

Play around with the spectrogram and make your own music! Can you tell why this kind of plot is referred to as a "waterfall plot"?

Fun fact: Waterfall plots connect music and images. So called "spectrogram art" is the art of writing music that creates hidden images in its spectrogram. You can find a nice article with examples here: mixmag.net/feature/spectrogram-art-music-aphex-twin

Setting up a spectrometer

[note [JS]: Exercise on calculating the number of spectrometer bins required! => experiment idea: observe the spectrum towards the galactic pole. Expect single line -> What is the expected line width? How do we have to set up the spectrometer? Have them perform the observation and interpret the output]

[note [JS]: example spectrum (Stockert SRT 21-cm data?)]

[note [JS]: add background reduction already?]

5.3 Other backends

As the physical processes we are interested in are complex, many, many "speciality" backends have been developed and are still being developed to keep up with latest technology and optimize trade-offs between different properties of our systems. For example so-called "Pulsar backends" have a significantly better time resolution to observe the extremely short pulses of rapidly spinning Pulsars, but at the cost of frequency resolution.

6 Radiation Transfer

Analogous to the reception of electromagnetic waves by antennas, the radiation of cosmic sources interacts with the matter located along the line of sight to the observer, such as gas or dust clouds. The physical state of the matter can be changed. Simply put, photons do not just get emitted and travel unchanged to our telescopes. Many different processes can occur between the place they are emitted and where they are received. For example, a photon can raise an electron to a higher energy level if it is resonantly absorbed, i.e., if it matches the energy scheme of the atom. By re-configuring the electron shell, this energy is re-emitted. If the photon energy is high enough to release an electron from the Coulomb potential of the atom, this ion can become a neutral atom again by recombination with a free electron. In all these processes it is necessary that the energies of the particles/photons fit to each other, so that they can be bound by the atom or ion.

If the energy of a photon is very high, it can also release strongly bound electrons from the innermost shells of the atom. A series of non-radiative ions can take place, which bring the ion into a state with minimal energy. These non-radiative transitions are described by the Auger effect and play an important role in X-ray astronomy.

All these single atomic level processes take place where electromagnetic radiation meets (baryonic) matter. Since these processes strongly depend on the absorbed atoms and ions as well as on the photon energy, it is necessary to follow a more general approach, which does not consider the single atom or ion but an ensemble or collection of atoms or ions. This is done in the following via radiative transfer equation.

In matter-free space it can be assumed that the specific brightness I_ν of a source is constant in time:

$$\frac{dI_\nu}{dt} = 0 \tag{38}$$

If a photon hits matter, this photon can be absorbed by the atom or ion with a probability dP along a path ds. The longer the path through the interstellar gas the higher the absorption probability κ.

$$\kappa = \frac{dP}{ds} \tag{39}$$

κ is the linear absorption coefficient and has unit $[\mathrm{m}^{-1}]$. The specific brightness I_ν is estimated with probability dP on the path ds depending on the absorption coefficient κ.

$$dI_\nu = -dp \cdot I_\nu = -I_\nu \cdot \kappa \cdot ds \rightarrow \tag{40}$$

$$\frac{dI_\nu}{I_\nu} = -\kappa ds \tag{41}$$

The solution of this equation is found using an exponential function.

$$I_\nu(D) = I_\nu(0) \cdot e^{-\int_0^D -\kappa(s')ds'} + const. \tag{42}$$

This equation states that the incident brightness is decreased exponentially. The constant is in example a radiating source located in the foreground. The optical depth of matter for radiation along the line of sight D is given by:

$$\tau = -\int_0^D -\kappa(s')ds' \tag{43}$$

If $\tau \gg 1$ then light cannot pass through the matter distribution, it is **optically thick** for us. An example for such a matter distribution is e.g. the Sun. It is optically thick for electromagnetic radiation, but not for neutrinos. If $\tau \ll 1$ then light can pass through the matter distribution, it is **optically thin**. The Milky Way is optically thin on most lines of sight for the HI 21-cm line radiation of the neutral atomic hydrogen.

In addition to the electromagnetic radiation, the absorbed energy is re-emitted. This is described analogous to the above approach by an emission coefficient j_ν which gives the probability that on the distance ds the specific brightness is increased by the matter on the line of sight

$$j_\nu = \frac{dI_\nu}{ds} \tag{44}$$

The total change of the specific brightness can be written as:

$$\frac{dI_\nu}{ds} = j_\nu - \kappa \cdot I_\nu \tag{45}$$

This is the radiative transfer equation.

According to Kirchhoff, in the case of full thermodynamic equilibrium (TE, thermal equilibrium)

$$\frac{dI_\nu}{ds} = 0 \tag{46}$$

which simply has conservation of energy as its statement. If we consider as a black–body as the source of radiation, then in TE the following is valid

$$dI_\nu = B_\nu(T) \tag{47}$$

Thus we find

$$\frac{dI_\nu}{ds} = 0 = j_\nu - \kappa \cdot B_\nu(T) \rightarrow \tag{48}$$

$$.\frac{j_\nu(T)}{\kappa(T)} = B_\nu(T) \tag{49}$$

This approach is used not only in the strict case of TE, but also in the case where only local thermal equilibrium (LTE)can be assumed, for example the Earth's atmosphere. It is illuminated by an electromagnetic radiation field of a black body with a temperature of 5800 K, but its particles have velocities which can be described by a Maxwell-Boltzmann velocity distribution of 300 K. The Earth's atmosphere is only locally in thermal equilibrium with the radiation of the Sun.

6.1 Einstein Coefficients

Up to this point we have neglected to deal with the atoms directly. However, if we want to model radiative transfer in detail, it is necessary to describe the radiative processes on an atomic level quantitatively.

For this purpose the Einstein coefficients are used. They describe the probability of a reconfiguration of the electron shell of an atom or ion. In simple terms, how likely it is for an atom to emit or absorb a photon. The coefficient A_{UL} describes the spontaneous emission of a photon by the transition from an upper (U) to a lower (L) energy level. B_{LU} is the Einstein coefficient for the absorption and B_{UL} that for the stimulated emission of a transition by a photon which is not absorbed by the atom or ion.

In the case of TE, the occupation of the energy levels of an atom or ion is given by the Boltzmann distribution

$$\frac{n_{\mathrm{U}}}{n_{\mathrm{L}}} = \frac{g_{\mathrm{U}}}{g_{\mathrm{L}}} \cdot e^{-\left(\frac{E_{\mathrm{U}}-E_{\mathrm{L}}}{\mathrm{k}T}\right)} \tag{50}$$

Here, n_{U} describes the number of atoms/ions in the excited level, n_{L} in the lower energy level, g_{U} and g_{L} the statistical weights of the energy levels. The statistical weights are a measure of the number of possible quantum states an electron can assume at an energy level. So it depends on the orbital angular momentum, the magnetic quantum number, etc. In the case of the hydrogen atom $g_{\mathrm{n}} = 2n^2$ with n as the principal quantum number.

Thus, the radiative transfer equation is given by:

$$\frac{dI_\nu}{ds} = j_\nu - \kappa I_\nu \tag{51}$$

$$= \left(\frac{\mathrm{h}\nu_0}{4\pi}\right) n_U A_{\mathrm{UL}} \Phi(\nu) - \left(\frac{\mathrm{h}\nu_0}{\mathrm{c}}\right) \left(n_{\mathrm{L}} B_{\mathrm{LU}} - n_{\mathrm{U}} B_{\mathrm{UL}}\right) \Phi(\nu) I_\nu \tag{52}$$

where $\mathrm{h}\nu_0$ is the energy difference between the two energy levels and thus the photon energy. The $\Phi(\nu)$ describes the shape of the line that is emitted or absorbed. To calculate the radiative transfer it is necessary to know the abundances of the different atoms and ions in a gas as well as the position of the energy levels and their occupation as a function of temperature, which is quite complex and time consuming.

6.1.1 Line width

[note [JB]: include basic figure of spectral peak with line width marked]
When the spectral emission is measured, it can be seen that the emission does not take place at a single frequency but instead across a short range of frequencies centered on the characteristic frequency of the process. How much the emission is spread around this rest frequency gives insight to the many physical and chemical conditions of the environment in which the emission is taking place.

If we initially focus only on the atomic level and neglect everything concerning the environment of the atom, we find the so-called natural line width in addition to the principal quantum number n which is used as a measure of the energetic spacing of the levels in the atom. An atom at rest in the laboratory relative to the spectrometer shows this width of the emission line.

It results from Heisenberg's uncertainty Principle:

$$\Delta E \cdot \Delta t \geq \frac{\mathrm{h}}{2\pi} \tag{53}$$

where E is the energy of the photon released and t is the time. ΔE and Δt are the uncertainty in these two measurements. Since the energy of the photon is related to the wavelength or frequency of the photon, ΔE describes the line width of the emission. The uncertainty relation is an estimate for minimum line width.

$$\Delta\nu = \frac{1}{\Delta t} \geq \frac{\Delta E\, 2\pi}{\mathrm{h}} \tag{54}$$

This natural or intrinsic line width is much, much smaller than what is actually observed from an ensemble of atoms. There are multiple reasons for this. One of these is thermal broadening. The atoms or molecules which make up the ensemble move according to the temperature of the ensemble. The thermal motion of the particles in the ensemble is described by the Maxwell velocity distribution. This means that each particle in the ensemble moves with both the large radial velocity of the entire ensemble and the smaller individual motion caused by the thermal energy in the ensemble. The radial motion of the ensemble determines where the center of the spectral peak is, while the thermal motion determines how broad the peak is. Think of a coffee to go on the bus. The warm coffee makes small thermal motions inside the cup, therefore the cup is warm. The bus gives this system a high velocity, only the radial component we can measure with a spectrometer of it. This large spatial motion is superimposed on the smaller ones in the cup. With an infrared spectrometer we will observe the cup with a group velocity equal to the radial velocity component of the bus. Around this systemic velocity a Gaussian velocity distribution of the warm coffee can be seen.

The motion of an atom with mass M in a gas with temperature T is described by the Maxwell velocity distribution. Here it is only important for us to know which parameters influence the line width.

$$\Delta\nu = \sqrt{\frac{8\ln(2)\mathrm{k}}{\mathrm{c}^2}}\sqrt{\frac{T}{M}} \tag{55}$$

$$= const.\sqrt{\frac{T}{M}} \tag{56}$$

According to this equation, the line width $\Delta\nu$ is proportional to the root of the temperature and inversely proportional to the root of the mass of the atom or ion. The higher the gas temperature the higher the average speed to the moving constituents. An increase of the particles mass however reduces the line width, because the heavier the particle the more energy is required to accelerate it.

Exercise 6.1

- Estimate the natural (Lorentz) line width of the excited state that leads to the emission of the HI 21-cm line. Use the Einstein coefficient $A_{10} = 2.85\times10^{-15}\,\mathrm{s}^{-1}$. Hint: Use the Heisenberg uncertainty principle for energy and time.

- Estimate the life-time of the excited state that leads to the emission of the 21-cm line, based on the Einstein coefficient A_{10}.

Exercise 6.2

- The absorption coefficient κ_ν for the 21-cm line is:

$$\kappa_\nu = \frac{c}{8\pi\nu^2}\frac{g_1}{g_0}n_0 A_{10}\left(1 - \mathrm{e}^{-\frac{h\nu}{\mathrm{k_B}T}}\right)\cdot\phi(\nu) \tag{57}$$

Where g_i are the statistical weights of the energy levels ($g_1 = 3$, $g_0 = 1$), n_0 the number of atoms, A_{10} the Einstein coefficient of the $1 \rightarrow 0$ transition, ν the frequency, and $\phi(\nu)$ the line shape (described by a Gaussian).

Show that for the 21-cm line ($\nu \approx 1402.4 \times 10^6\,\mathrm{s}^{-1}$) and a spin-temperature $T_S = 150\,\mathrm{K}$, eq. 57 can be simplified using series expansion.

Use $n_H = n_0 + n_1 \approx 4n_0$ to rewrite the equation, so that all constants are separated from the variables.

What does this mean for measurements of the HI line?

7 Influences of the Atmosphere TBW

7.1 Atmospheric Noise

The gas that makes up the atmosphere can have significant effects on the radio waves that pass through it, and therefore can seriously effect the quality of the observations that take place. The effects which influence radio waves occur in a regions of the atmosphere known as the troposphere and ionosphere.

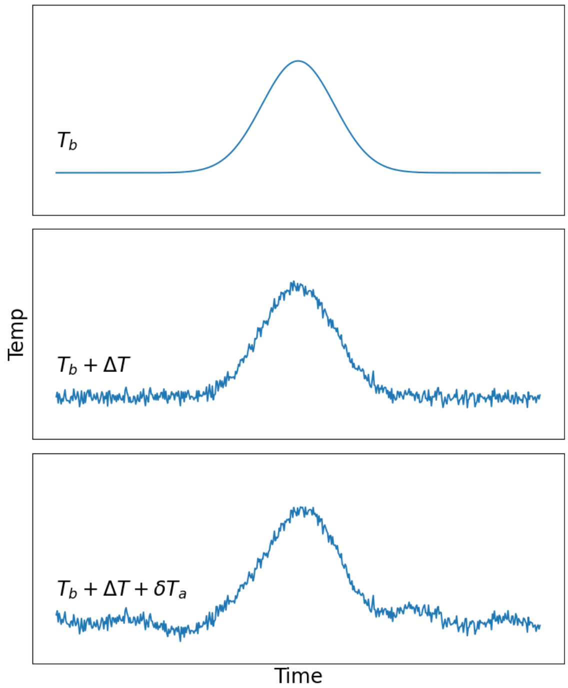

Fig. 7.1.1: Fluctuation superimposed over an astronomical signal.

Figure 7.1.1 illustrates how the atmosphere impacts the measurement of an astronomical signal. The Radiometer Equation (30) introduces noise, ΔT, to the signal. It determines the sensitivity of the measurements that can take place. ΔT depends on the system temperature T_{sys}. A number of different factors contribute to T_{sys}, some of which depend on the telescope's instruments such as receiver temperature, or the source such as the brightness temperature, but also the atmosphere temperature, T_a contributes to this noise.

The atmospheric temperature depends on the total column density of water vapour that the astronomical signal has to travel through to reach the receiver. However the distribution of water vapour in the atmosphere is not constant or homogeneous, due to inhomogenous temperature distribution and atmospheric turbulence. This means that there is a time-varying atmospheric signal superimposed on top of the astronomical signal, $\delta T_a(t)$. This is shown in figure 7.1.1 (bottom). This further limits the quality of the astronomical measurements that can take place.

7.2 Troposphere

The troposphere is the lowest layer of the atmosphere, and reaches approximately 18 km in height in the tropics and decreases in size closer to the poles. The troposphere contains approximately 75% of the mass of the total atmosphere, and almost all the water vapour. It is the layer in which most weather phenomena occur. This water vapour is very important as it is what causes most of the atmospheric effects present in radio observations. Some of the other atmospheric components also cause absorption and emission, but they are minor compared to the water vapor. The list below presents the most common molecules in the atmosphere and their effect on radio signals.

- N_2: No effect on signals due to lack of permanent dipole moment.
- O_2: Has a magnetic dipole with transitions at 60 and 118GHz.
- O_3: Many possible transition at mm/submm wavelengths. Has a low optical depth and therefore does not significantly influence most radio observations.
- H_2O: Multiple possible transitions in the radio regime and has a strongly varying optical depth due to weather. Strongly influences radio observations.

[note [JB]: Find absorption line diagram for atmosphere similar to fig 5.2 U.Klien script]

7.2.1 Emission and Absorption

The same equations of radiative transport, discussed in Section 6, are also used to describe the absorption and emission which occurs in the troposphere. As radiation passes through the atmosphere, there are two processes that take place. Some of the radiation gets attenuated as it passes through the gas. The atmosphere also emits radiation itself as a black–body. Thus, the flux measured by the telescope when pointing at the source is a combination of flux from the source, minus what has been absorbed, and the black–body spectrum of the atmosphere.

Absorption can be described as the infinitesimal reduction of intensity $I\nu$, by an amount dI_ν as it passes through a section of the atmosphere with path length ds. This is shown in equation 58, where κ is the absorption coefficient with units of m^{-1}.

$$\begin{aligned} dI_\nu &= I_\nu(s+ds) - I_\nu(s) \\ &= -\kappa I_\nu ds \end{aligned} \tag{58}$$

The absorption which occurs as radiation passes through a medium is also described by the optical depth. It is defined as the fraction of intensity that remains after passing through a gas column. Mathematically it is given by the integral of the absorption coefficient of the medium over a path length from the lower bound s_0 to an upper path length s.

$$\tau_\nu = -\int_{s_0}^{s} \kappa ds \tag{59}$$

Along the path length of ds, radiation is not only absorbed, it as also emitted by the molecules in the atmosphere, increasing the intensity of the radiation. This is described using the emission coefficient j_ν.

$$dI_\nu = j_\nu ds \tag{60}$$

Equation 61 combines both the absorption and emission into a single equation. This is shown illustratively in Figure 7.2.1.

$$dI_\nu = -\kappa I_\nu ds + j_\nu ds \tag{61}$$

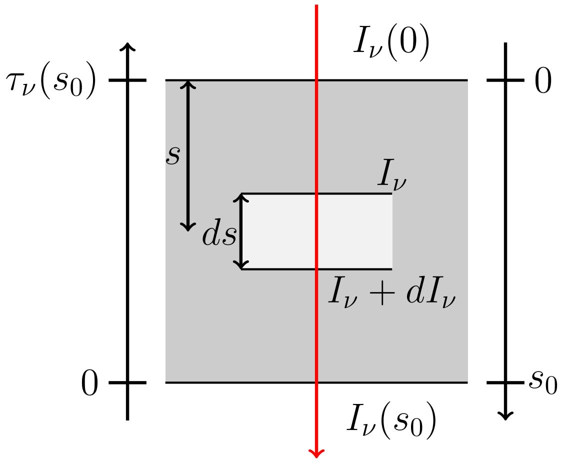

Fig. 7.2.1: Effect of atmosphere on the intensity I of radiation as it passes through a distance s of atmosphere.

The source function, S_ν is defined as the ratio of the absorption and emission coefficients. When the medium is in local thermodynamic equilibrium (LTE), the source function is the same as the Planck intensity from a black–body. This is referred to as Kirchhoff's Law. This means that the emission properties only depend on temperature and not the properties of the material the radiation is passing through. This is only strictly true in special situations such as stellar interiors. However, in the radio domain it remains a good approximation due to the Rayleigh-Jeans approximation discussed in Section **??**.

$$S_\nu = \frac{j_\nu}{\kappa} = B_\nu(T) \tag{62}$$

Using the relations derived above, it is possible to rewrite equation 58 in terms of optical depth and the black–body spectrum.

$$I_\nu(s_0) = I_\nu(0)e^{-\tau_\nu(0)} + B_\nu[1 - e^{-\tau_\nu(0)}] \tag{63}$$

Using the Rayleigh-Jeans approximation this equation can be rewritten in terms of the brightness temperature T_b, where the the black–body temperature is replaced by the atmospheric temperature T_{atm} and the brightness temperature of the source.

$$T_b(s_0) = T_s(0) \cdot e^{-\tau_\nu(0)} + T_{\text{atm}} \cdot [1 - e^{-\tau_\nu(0)}] \tag{64}$$

The properties of the atmosphere, including its frequency dependence, are encoded in the optical depth τ_ν, which is determined through experiments. This, and the atmospheric temperature, T_{atm}, are the components that are highly time variable due to weather conditions.

7.3 Ionosphere

The ionosphere is part of the Earth's upper atmosphere, covering the region between approximately 50 to 1000 km above sea level. High energy radiation from the Sun ionises this layer of the atmosphere resulting in free electrons which interact with incoming radiation. The number density of free electrons is not constant and is dependant on the amount of solar radiation it receives. This means that it is dependant on time of day, season, and level of solar activity.

7.3.1 Limiting Frequency

The number density of free electrons, n_e, determines the refractive index n of the ionosphere, which determines the propagation of radiation through it. However, the refractive index is also dependant on the frequency of the radiation passing through it.

$$n = \sqrt{1 - \frac{4\pi n_e e^2}{m_0 \nu^2}} \tag{65}$$

Equation 65 is the equation for the refractive index of the ionosphere as a function of the frequency of the incoming radiation ν, and electron number density n_e. It is also dependant on the charge of the electron e, and the mass of the free electron m_0. The derivation of this equation is beyond the scope of this course. As can be seen from the form of the equation, for every value of n_e, there is a minimum frequency ν, below which the refractive index becomes imaginary. electromagnetic waves with frequencies below this value can not propagate through the ionosphere. this is called the plasma frequency and corresponds to the frequency of mechanical oscillations in a plasma. During a period of low solar activity during autumn at night, the limiting frequency has a value of $\nu_{\min} = 3\text{MHz}$ and during a period of high solar activity during the daytime during spring it has value of $\nu_{\min} = 24\text{MHz}$.

Unlike when the troposphere, when radio signals can not pass through the ionosphere they are not absorbed by the atmosphere. Instead, they are reflected. As discussed in Section 1.1.1, this effect is has been used for a long time to broadcast radio signals across the globe. This can cause two issues for radio astronomy , as signals are not only effected by the ionosphere but there is also interference caused by signals coming from Earth.

7.4 Refraction

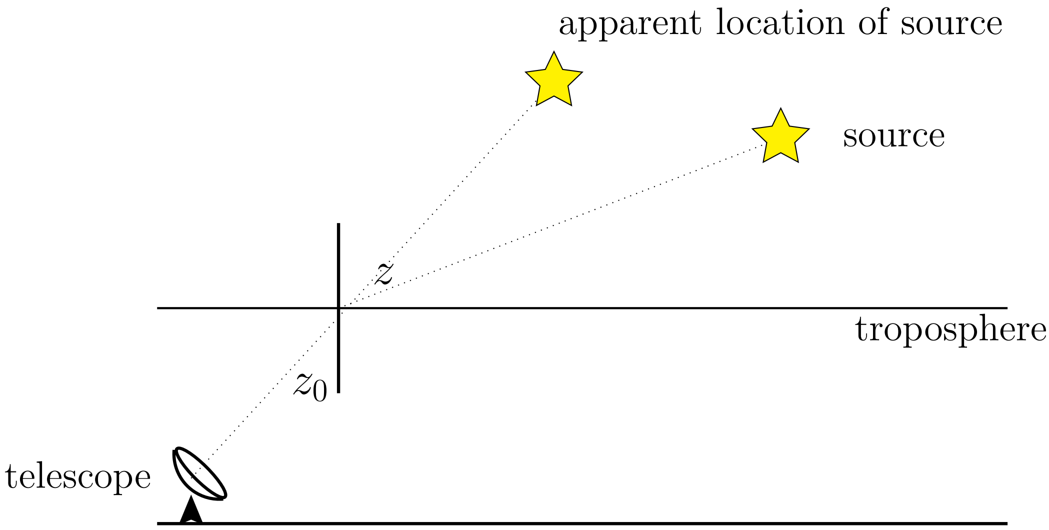

Fig. 7.4.1: Refraction of beam by troposphere modelled as a uniform slab with constant refractive index.

As a light propagates from one medium to another of different density at an angle, the ray of light gets deflected. This is called refraction. This occurs as radio waves enter the atmosphere from outer space. This primarily occurs at the troposphere mainly due to the water vapour present.

The refractive index of the atmosphere depends on the total air pressure, the amount of water vapour in the atmosphere and the temperature of the atmosphere. These are all weather dependent and hence can change greatly from place to place, day to day and also even throughout an observation.

The effect of refraction is that it causes sources to appear higher in the sky than they truly are. For example, when the sun touches the horizon during a sunset, it is in fact already set below the horizon and the atmosphere deflects the ray of light and causes it to still be visible. This atmospheric refraction is illustrated in Figure 7.4.1. In this figure, the troposphere is modelled as a uniform, plane-parallel slab with constant refractive index. As can be seen, the source appears higher in the sky than it truly is. The refraction of the radiation from the source is governed by Snell's Law

$$n_0 \sin z_0 = \sin z \tag{66}$$

where z is the zenith angle at the top of the troposphere, where the refractive index is assumed to be $n = 1$, and z_0 is the zenith angle as viewed by the telescope.

This situation is clearly an oversimplification, as in reality the troposphere cannot be modelled as a uniform, plane-parallel slab with constant refractive index. It has a decreasing refractive index gradient and does not have a distinct border with the higher layers of the atmosphere. This means ray tracing beams through the atmosphere is more complicated. However, the above example is useful to describe the basic principle of the effect. Luckily, this atmospheric effect is often taken into account by the steering software of many telescopes, as otherwise the telescope would quickly lose track of its target as it changes elevation.

7.5 Scintillation

Scintillation is caused by the time-variable refraction in the atmosphere. In the ionosphere, the distribution of free electrons is homogeneous and causes phase shifts between rays of light passing through different paths. The rays of light coming from the source interfere with each other both constructively and destructively. This results in changes of intensity of the observed source. Scintillation also causes the angular broadening of sources.

7.6 Site Selection

When choosing the location for a radio telescope there are two aspects to consider. Firstly there is the Radio Frequency Interference which must be minimised. RFI can occur naturally or the result of human activity. Natural sources of RFI include electrical storms, solar radiation and cosmic radiation. RFI which originates from human activity can be from radio transmitters, power lines, computers. There are a number of ways to minimise RFI. The Green Bank Radio Telescope in the USA does this by implementing a National Radio Quiet Zone around the observatory. Within this zone, radio transmissions are restricted by law. In some areas, even WiFi, microwaves and digital cameras are prohibited in order to minimise unintentional radio emission.

Another method to reduce RFI is to locate the radio telescope in a valley. The surrounding mountails help to shield the telescope of from unwanted radiation. This was the approach taken by the Effelsberg 100-m Radio Telescope in Germany which is located in a valley in the Eifel mountain range.

[note [JB]: include pic of effelsburg]

Another aspect to consider when choosing the location of a radio observatory is the climate. As discussed previously in Section 7.2, the amount of water vapour present in the atmosphere significantly impacts the quality of radio observations. Therefore it is beneficial to place radio telescopes in areas which have a high altitude and are dry. This high altitude means that the radio signal has less atmosphere to travel through. The high altitude also reduces the amount of humidity present. The Atacama Large Millimeter/submillimeter Array (ALMA) was constructed on the Chajnantor Plateau in the Atacama Desert in northern Chile, at an elevation of 5,000-m.

[note [jb]: pic of ALMA in desert]

8 Radiation Processes IMAGES

[note [JS]: Exercise ideas: emission processes in the Sun -> what kind of spectrum would we expect to observe overall? What do we expect to observe with the oven pipe antenna? Sun transit; Similar to what is done in the Radio master lab course]

The emission of light at optical wavelengths is what most of us are familiar with. The emission of optical light occurs when an electron changes distance from the nucleus of its host atom. A change in distance results in a change in energy due to the Coulomb potential:

$$\Delta E = const. \cdot q_1 \cdot q_2 \left(\frac{1}{r_1} - \frac{1}{r_2} \right) \tag{67}$$

$$= \mathrm{h} \cdot \nu \tag{68}$$

Quantum mechanical systems such as atoms and molecules are most stable in the state which minimises their energy, this is called the ground state. If a system has more energy, it is less stable, and it will attempt to release this energy and return to its ground state. Spontaneous emission is the process which occurs when an atom or molecule goes from a higher energy state to a more stable lower energy state by changing the distance of an electron to the nucleus.

In the case of visible light the change of energy between to energy states, ΔE, amounts to a few eV and corresponds to changing the electron configuration, changing the main quantum number N, which is equivalent to changing the atomic shell. These energies are about a million times greater than the amount of energy of radio photons. So the radio line emission relies on very low energy transitions.

In general there are two line emission processes for atoms that are important for radio astronomy. Either hyperfine structure transitions or the re-configurations of electrons within the outermost atomic shells. N is here on the order of few tens. At high N's the energy difference ΔE between adjacent level is sufficiently low to emit radio photons.

The emission of electromagnetic radiation is not only caused by transitions between states of different energies. It is also caused by the acceleration of charges in electric and magnetic fields. This type of emission results in continuous radiation spectra, while changes in energy state results in line emissiion in spectra.

Both of these emission types are observed in radio astronomy, like in the optical light domain from stellar photospheres, continuous and line emission. So let us start to explore the field of radiation processes.

8.1 *Basics:* Emission from Charged Particles.

This section does not wish to substitute for a comprehensive course on and therefore, will only briefly show the basic principles of why accelerated electric charges emit electromagnetic waves. In chapter **??** we looked at simple cases for an undamped oscillation.

Maxwell's equations

Consider Maxwell's four equations in a space free of electric charges:

$$\boldsymbol{\nabla} \cdot \mathbf{E} = 0 \tag{69}$$

$$\boldsymbol{\nabla} \times \mathbf{E} = -\frac{d\mathbf{B}}{dt} \tag{70}$$

$$\boldsymbol{\nabla} \cdot \mathbf{B} = 0 \tag{71}$$

$$\boldsymbol{\nabla} \times \mathbf{B} = \mu_0 \, \epsilon_0 \, \frac{d\mathbf{E}}{dt} \tag{72}$$

B represents the vector of the magnetic field, **E** the vector of the electric field. Due to the charge free (neutral) space, $div\mathbf{E} = 0$.

Calculating the curl of the two non-zero equations, using the identities:

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times E) = \nabla(\boldsymbol{\nabla} \cdot E) - \nabla^2\mathbf{E} = -\nabla^2\mathbf{E} \tag{73}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times B) = \nabla(\boldsymbol{\nabla} \cdot B) - \nabla^2\mathbf{B} = -\nabla^2\mathbf{B} \tag{74}$$

and

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times E) = \boldsymbol{\nabla} \times \left(-\frac{d\mathbf{B}}{dt}\right) = -\frac{d}{dt}(\boldsymbol{\nabla} \times B) = -\mu_0 \, \epsilon_0 \, \frac{d^2\mathbf{E}}{dt^2} \tag{75}$$

$$\boldsymbol{\nabla} \times (\boldsymbol{\nabla} \times B) = \boldsymbol{\nabla} \times \left(-\frac{d\mathbf{E}}{dt}\right) = -\frac{d}{dt}(\boldsymbol{\nabla} \times E) = -\mu_0 \, \epsilon_0 \, \frac{d^2\mathbf{B}}{dt^2} \tag{76}$$

$$\tag{77}$$

Yields:

$$0 = \nabla^2\mathbf{E} - \mu_0 \, \epsilon_0 \, \frac{d^2\mathbf{E}}{dt^2} \tag{78}$$

$$0 = \nabla^2\mathbf{B} - \mu_0 \, \epsilon_0 \, \frac{d^2\mathbf{B}}{dt^2} \tag{79}$$

$$\tag{80}$$

Beginning here with Ampere's induction law. It results after some calculation:

$$\Delta\vec{E} = \mu_0 \cdot \epsilon_0 \frac{d\vec{E}^2}{dt^2} \tag{81}$$

This is an equation of oscillation as we have already treated. You also already know the solution, a periodic oscillation. Therefore, Maxwell's equations lay the foundation for understanding electromagnetic waves.

8.2 The Lamor Equation

We follow here the path of explanation also followed in the Essential Radio Astronomy. It is based on J. J. Thomson's understanding of the motion of a charged particle and the finite nature of the speed of light.

Thomson considers the intensity and angular distribution of the radiation of a charged particle that experiences a small acceleration $\Delta v/\Delta t$. Due to the acceleration of the charge, the electric field, which is directed radially outward, is disturbed. These disturbances are transmitted with the speed of light. After a time Δt, the disturbance reaches a distance of $r = \mathrm{c} \cdot \Delta t$. What happens to the electric field lines? For this purpose we look more closely at the following figure.

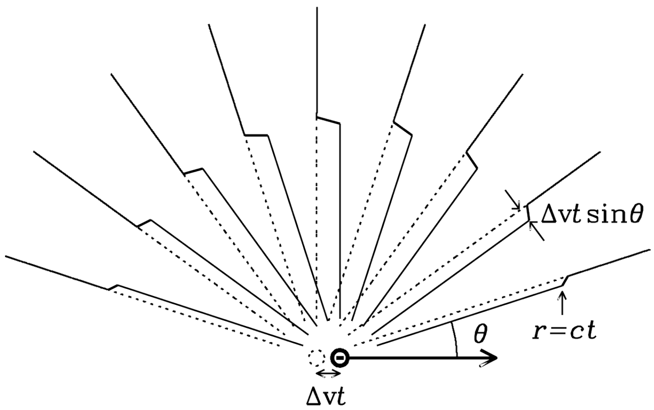

Fig. 8.2.1: The charged particle, while at rest, is at the center of the figure. Then it is accelerated for a short time $\Delta v/\Delta t$ and after a time t it reaches the velocity $\Delta v{\cdot}t$. This velocity is maintained. The field lines, however, show a perturbation. This disturbance propagates in space with the speed of light c and reaches a distance $r = \mathrm{c} \cdot t$ after t. In the direction of acceleration nothing happens, but perpendicular to it a maximum amount happens. This angular dependence is described via $\sin(\theta)$. Thus, the perturbation is maximally observed perpendicular to the orientation of the acceleration vector. **image source:**Condon *Essential Radio Astronomy* Fig. 2.22

The perturbation of the field is maximal perpendicular to the x-axis, the direction of motion of the particle. Along the direction of motion nothing can be seen. Therefore, the radiation is perpendicular to the orientation of the acceleration vector.

This can be expressed in an equation which compares the electric field perpendicular to the direction of motion $E_\perp$ and along the direction of motion $E_|$.

$$\frac{E_\perp}{E_{||}} = \frac{\delta v \cdot t \cdot \sin(\theta)}{\mathrm{c} \cdot t} \tag{82}$$

Inserting Coulomb's law for $E_|$ it then follows:

$$E_\perp = E_{||} \cdot \frac{\delta v \cdot t \cdot \sin(\theta)}{\mathrm{c} \cdot t} \tag{83}$$

$$= \frac{q}{r^2} \cdot \frac{\delta v \cdot t \cdot \sin(\theta)}{\mathrm{c} \cdot t} \tag{84}$$

$$= \frac{q}{r^2} \cdot \left(\frac{\delta v}{\delta t}\right) \frac{r \cdot \sin(\theta)}{\mathrm{c}^2} \rightarrow \qquad = \frac{q}{r} \cdot \frac{\sin(\theta)}{\mathrm{c}^2} \cdot \dot{v} \tag{85}$$

Perpendicular to the direction of acceleration, an accelerated charge radiates maximally.

It is now possible to calculate the radiated power as a function of angle θ. The following distribution is obtained.

The total power radiated is equal to the acceleration received by the charged particle. It is described by the Lamor equation:

$$P = \frac{2}{3} \cdot \frac{q^2}{\mathrm{c}^3} \cdot \dot{v}^2 \tag{86}$$

Lamor's equation states that any charged particle radiates electromagnetic waves when accelerated. The total power radiated is proportional to the acceleration squared. The acceleration is a function of the mass of the particle. Therefore, the radiation from the EM wave is proportional to the charge/mass ratio of the particle. Therefore, electrons/positions radiate much more powerfully than heavier protons:

$$\left(\frac{m_\mathrm{e}}{m_\mathrm{P}}\right)^2 \simeq 4 \cdot 10^6 \tag{87}$$

Consider again fig. **??**. It can be seen that the spectral progressions of three radiation processes: The synchrotron radiation, the black body, and the recombination radiation. But why are they emit in frequency in this order? Can the radiation processes

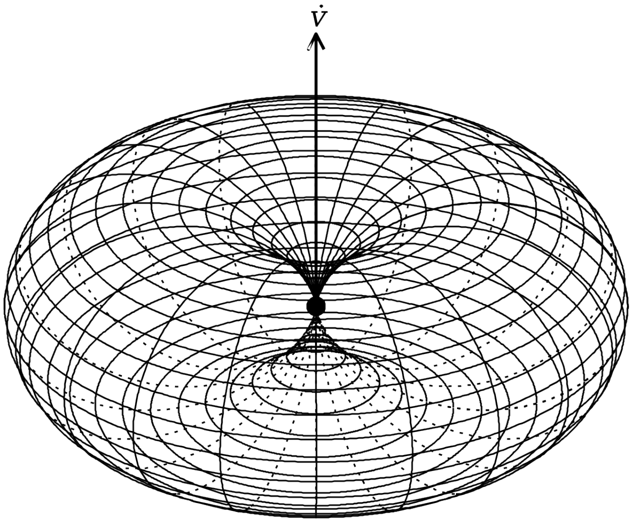

Fig. 8.2.2: The power distribution of an accelerated charged particle. **image source:**Condon *Essential Radio Astronomy* Fig. 2.23

be observed alone or are they only observable together? These may be the questions that are going through your mind and the figure does not help. Therefore, we will now take a closer look at these radiation processes and see under which physical conditions they are observed.

8.3 Radiation Types

Any accelerated charged particle emits electromagnetic radiation. Electric fields can accelerate charged particles. Likewise, magnetic fields can accelerate charges by drawing the charges into a circular path around magnetic field lines.

The lightest charged particles like electrons and positrons are accelerated more than the more massive protons or heavier ions. Therefore, the emitted radiation is mostly caused by the acceleration of electrons.

The kinetic energy, or more precisely the velocities with which the electrons move relative to the field decides which form of radiation we observe.

- $E_{\text{kin}} \ll E = m\,\mathrm{c}^2$ then the radiation spectrum is called gyromagnetic radiation.
- $E_{\text{kin}} \simeq E = m\,\mathrm{c}^2$, where the particle is non- relativistic, is called cyclotron radiation.
- $E_{\text{kin}} \gg E = m\,\mathrm{c}^2$, where the partiocle is relativistic, is called synchrotron radiation.

8.3.1 Gyromagnetic Radiation

Assume the Lorentz force

$$\vec{F} = q \cdot \frac{\vec{v}x\vec{B}}{\mathrm{c}} \tag{88}$$

The charged particle follows the right-hand rule, according to which the vectors of the charged particle's motion and the magnetic field are oriented perpendicular to each other, i.e. the magnetic field cannot change the velocity of the charged particle by the action of a force. The force action is always perpendicular to the orientation of the magnetic field and the direction of motion of the particle. Thus, the Lorentz force forces the charged particle to follow a circular path around the magnetic field line.

We can take advantage of this by a) allowing us to assume $\frac{dv}{dt} = 0$ and b) to set $v = \omega_{\mathrm{G}} \cdot r$. ω_{G} stands for the angular frequency and is in its unit s^{-1}. Rewrite Eq. 88:

$$\vec{F} = m \cdot \vec{a} = m \cdot \omega_{\mathrm{G}}^2 \cdot \vec{r} \tag{89}$$

$$= \frac{q}{\mathrm{c}} \cdot \vec{v}x\vec{B} \rightarrow \ |F| \qquad = \frac{q}{\mathrm{c}} \cdot \omega_{\mathrm{G}} \cdot r \cdot B \tag{90}$$

This is called the gyro frequency.

$$\omega_{\mathrm{G}} = \frac{q \cdot B}{m \cdot \mathrm{c}} \tag{91}$$

It is independent of the velocity of the particle, when calculated non-relativistically.

It follows

$$\frac{\nu}{[\mathrm{MHz}]} = 2.8 \cdot 10^{-4}\,[\mathrm{T}] \tag{92}$$

The magnetic field of the Sun at the location of the Earth is of the magnitudes of a few nT, the magnetic field of the Earth is in the magnitudes of a few tens of μT and for a neutron star it can reach a few kT.

If we extend our consideration now to the cyclotron and synchrotron range, we must apply only a small correction:

$$\omega_{\mathrm{Syn}} = \frac{\omega_{\mathrm{G}}}{\sqrt{1 - \left(\frac{v}{\mathrm{c}}\right)^2}} \tag{93}$$

$$= \frac{\omega_{\mathrm{G}}}{\gamma} \tag{94}$$

This correction seems small, but it has significant consequences. If the charged particle is moving at a speed close to the speed of light, $1/\sqrt{\gamma}$ becomes very large. The angular frequency becomes smaller and smaller the closer the charged particle gets to the speed of light. Substituting typical values for the magnetic field strength in the interstellar medium 10^{-8} T and a $\gamma \simeq 10^5$ into Eq. 92, we find:

$$\frac{1}{t_{\text{cycle}}} = \frac{\omega_{\text{G}}}{2\pi \cdot \gamma} = 30 \cdot 10^{-5}\,\text{s}^{-1} \tag{95}$$

$t_{\text{cycle}} \simeq 3300$ seconds or roughly one hour. This is how long the circulation of an electron/position around such a magnetic field line takes. However, it covers a considerable distance in the process. We can estimate the gyration radius simply by $2\pi \cdot r = \text{c} \cdot t_{\text{cycle}}$ and find

$$r = 157.6 \cdot 10^9\,\text{m} \simeq 1\,\text{AU} \tag{96}$$

The electron/positron travels a tremendous distance in that 1 hour. If the magnetic field is even much stronger, t_{cycle} becomes proportionally smaller. The synchrotron radiation around a neutron star is therefore placed correspondingly close to its magnetosphere. In other words high frequency synchrotron radiation originates close to the source itself, while low frequency synchrotron radiation, such as can be observed with LOFAR radio interferometers, traverses the spatial scales of the size of the halo of a galaxy.

8.3.2 Synchrotron Radiation

Synchrotron radiation is the most common case in astronomy. It is caused by relativistic, charged particles being accelerated or deflected in a magnetic field. It accounts for most of the radio emission from active galactic nuclei (AGNs). It dominates the radio continuum emission from galaxies such as the Milky Way at frequencies below $\nu \simeq 30$ GHz. Jupiter's magnetosphere is a very strong synchrotron radio source. Thunderstorms produce synchrotron radiation just as AGN radio jets have a perfect synchrotron spectrum.

8.4 The radiation of an electron

Our single charged particle circling around a magnetic field line shines in our direction once per hour. Because of its rapid motion, however, we see an extremely brief flash of light for a duration of about 10^{-10} s. Every hour, such a flash of light from a single electron/positron is seen. The shortness of the flash is due to the fact that the relativistic velocity completely changes the radiation pattern (see Fig. 8.4.2).

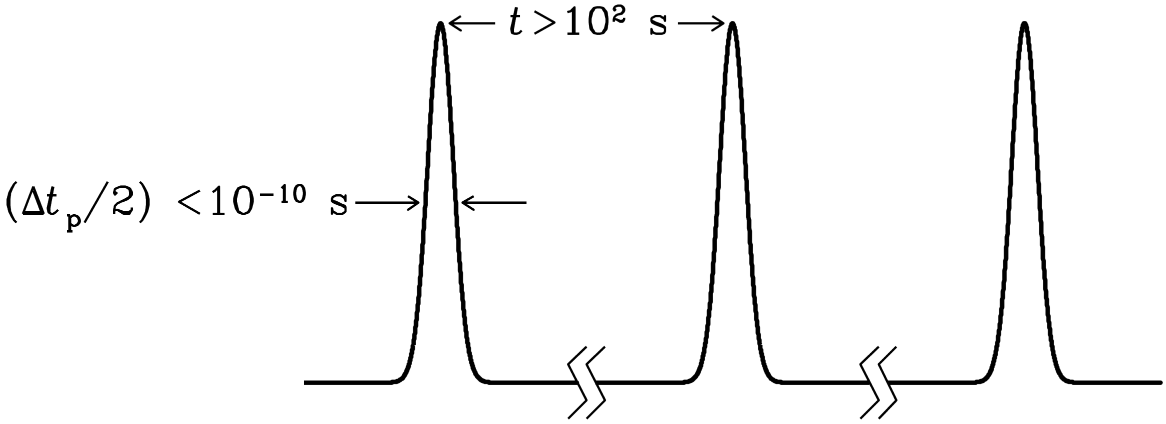

Fig. 8.4.1

However, we as radio astronomers do not observe a single electron, but a multitude of electrons. All of them have a slightly different energy, a slightly different orientation to our line of sight, etc. Therefore, we do not see one flash of light but a huge number.

From the observations we can derive a power law for the energy of the electrons

$$n(E) \cdot dE = E^{-\delta} \cdot dE \tag{97}$$

The δ indicates how much the energy decays. That is, it is easy to accelerate an electron through an electric field to velocities suitable for cyclotron radiation. But the larger the velocity of the electron/positron becomes, the larger the radiation losses. The δ is therefore also found in the radiation spectrum we can observe:

$$j_\nu \propto B^{\frac{\delta+1}{2}} \cdot \nu^{\frac{1-\delta}{2}} \tag{98}$$

The acceleration mechanism of the charged particles decides which spectrum is observed from a source. Directly, the magnetic field strength can be measured via the Zeeman effect, but only the part perpendicular to the magnetic field line. However, we unfortunately cannot infer the orientation. The same problem is also associated with the use of Faraday rotation. Unfortunately, we therefore cannot easily determine the strength of the magnetic field and must rely on estimates and plausible assumptions.

Note that the Larmor equation only describes the non-relativistic case. It is only applicable in this form when $v \ll$ c. To describe particles moving at near light speed relative to the observer, we first calculate the radiation according to Larmor's equation in its rest frame and transform the result relativistically correct into the observer frame.

Fig. 8.4.2: Relativistic radiation chattering is based on the transformation of the Larmor radiation pattern from the electron rest system (dashed line) to the observer rest system. In the observer's rest system, it determines the velocity of the electron which corresponds to $\gamma = 5$. **image source:**Condon *Essential Radio Astronomy* Fig. 5.3

8.5 Free–Free Radiation

Free-free radiation describes a continuous radiation process which is observed very often in radio astronomy.

The term "free" describes a charged particle (most commonly electron) which is not part of an atom or molecule. It is released from the Coulomb potential of its nucleus by a collision or by the absorption of a photon. The collision or photon imparts a certain amount of energy to the electron. Part of this energy goes to releasing the electron from its Coulomb potential. Any energy in excess of this, is given to the electron as kinetic energy. Because of the high difference in mass between electron and nucleus the part of the kinetic energy of the nucleus can be neglected.

If a free electron passes close to a an ion and is influenced by its Coulomb potential but however does not become bound to the ion, it releases an electron. This emission of electromagnetic radiation is called free-free radiation.

We can calculate the acceleration the electron experiences via:

$$|\dot{v}| = \frac{F_{\text{Coulomb}}}{m_{\text{e}}} = \frac{Z \cdot e^2}{m_{\text{e}} \cdot r^2} \tag{99}$$

Also in the case of free-free radiation, due to its smaller mass, the electron radiates more strongly than a proton or even a heavy ion. The acceleration and thus the power radiated according to the Lamor characteristic is proportional to the charge of the particle squared.

If the free electron comes too close to the ion, then it can be captured by it and become bound. However, this danger is small in the radio wave range because the photon energy in the radio range is a few 10^{-5} eV and thus the velocity vector $\vec{v}$ of the electron is only minimally changed. If the distance is so large that the electron does not interact with the ion, then no emission takes place. But the distance cannot become arbitrarily large. It is not as if this process takes place in empty space. It

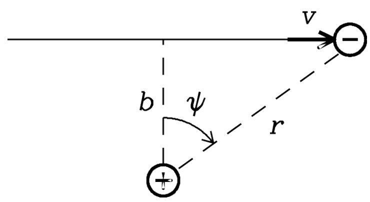

Fig. 8.5.1: The passage of an electron past an ion is shown. The photon energy in the radio range is so small, think of $E = \mathrm{h} \cdot \nu$ that there is very little change in the direction of motion of the electron given by $\vec{v}$. The smallest distance between the electron and the ion b is called the collision parameter and the derivable time interval $\tau = b/v$ is called the collision time. **image source:**Condon *Essential Radio Astronomy* Fig. 4.2

takes place near a source of ionizing radiation, e.g. near an O-star. The volume density of the particles then determines the average distance. If this is small, thus the density is high, then the impact parameter is accordingly small and the acceleration is high.

By the consideration that an electron can be captured if $E_{\mathrm{kin}} \simeq E_{\mathrm{coulomb}}$ follows:

$$b_{\mathrm{min}} \simeq \frac{< missing > \cdot e^2}{m_{\mathrm{e}} \cdot v^2} \qquad (100)$$

In the radioastronomical range, this corresponds to some 10^{-10} m.

According to Kirchhoff, like any thermal radiation, there exists an equilibrium between emission and absorption $B_\nu = j_\nu / \kappa_\nu$. Therefore, an ionized gas can also absorb radiation not just emit it. Therefore, we are again dealing with the two cases,

- The optically thin case, where all photons reach the observer.
- The optically thick case, where some or even nearly all photons are absorbed along their path through the medium.

In the optically thick case, only photons emitted at the surface of the ionized gas cloud are visible to the observer. In the optically thin case, the observer looks through the gas cloud, and the measured radiant power is a measure of the emission processes in the gas cloud.

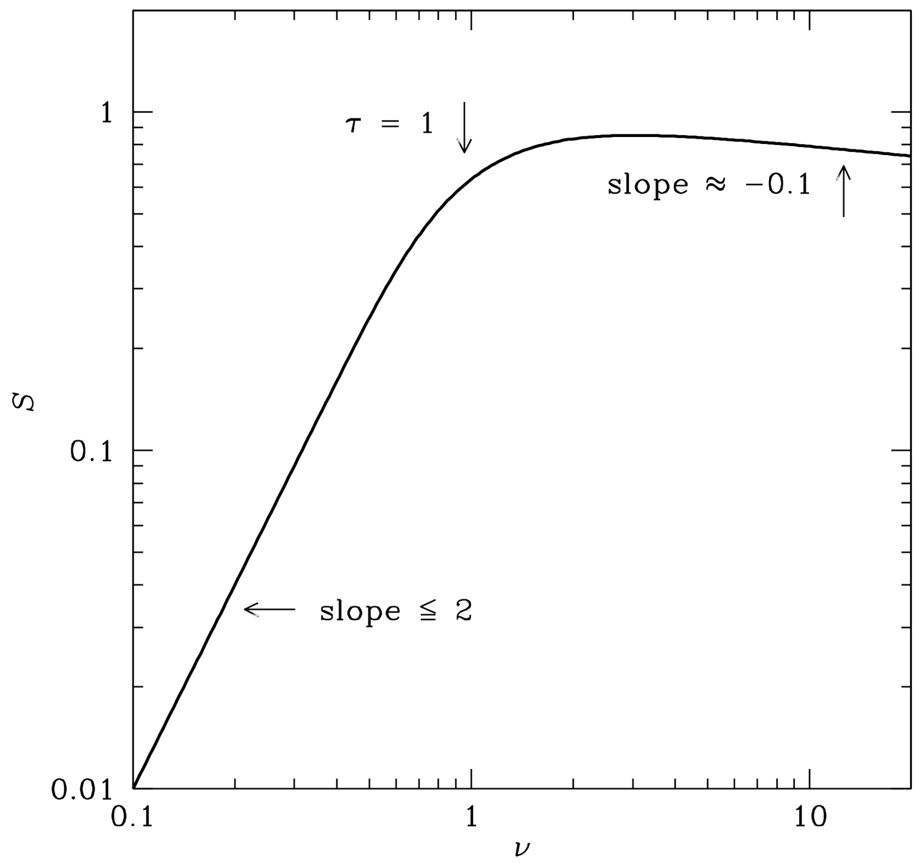

Fig. 8.5.2: The radio spectrum of the free-free radiation. At low frequencies we observe the radiation of a black body, because we can only detect the surface. Therefore $B_\nu \propto \nu^2$. At a certain frequency ν, which is determined by the density of the ionized gas cloud, the transition to the optical thin case occurs. From then on $\nu \simeq -0.1$, the observer looks into the plasma at high frequencies. At low frequencies we observe the radiation of a blackbody, its radiation temperature is therefore proportional to the electron temperature. At high frequencies, the radiation power is proportional to the emission processes and thus to the density of the ionized gas. **image source:**Condon *Essential Radio Astronomy* Fig. 4.8

9 The Milky Way: HI 21-cm line observations

When we look up at the sky on a dark and clear night, we can make out a diffuse band of light that extends across the entire sky. Using binoculars or a small telescope, you may verify the observations made by Galilei in 1609, when he discovered that this band is made of stars. This is the Milky Way: our own galaxy, as you can see it from Earth. It consists of billions of stars and our sun is merely one of them. In the entire universe there are few billion other galaxies and our Milky Way is merely one of them, too. Mostly we refer to the Milky Way as just "our galaxy".

Astronomers took a long time to find out what our galaxy really looks like. The most convincing method would probably be to board a spaceship and take a look at our galaxy from the outside. Sadly traveling in and around the galaxy is impossible due to the large distances involved (and will probably stay impossible for a while longer). We are damned to observe our galaxy from within, limited to the view we have from where our solar system is. This is what you will do in this experiment.

The sun is located in the outer regions of the Milky Way. It takes about 25 000 years for the light from the galactic center reaches our telescope, and an additional 20 000 years for it to leave our galaxy. Observations of other close-by galaxies in addition to what we can observe about ours, have led to us unraveling the structure of our own galaxy. Today, astronomers are convinced that they know how the gas and stars are distributed in the Milky Way.

Our galaxy is a very thin disk that consists of stars, dust, and gas. This disk is, relative to its size, much thinner than a vinyl record. It is only 300 lyr thick - with a diameter of 100 000 lyr! (Compare this ratio with that of a vinyl record.)

The stars in this thin disk are not homogeneously distributed, but take the form of spiral arms. This was found not long ago: Measuring the 21-cm line with radio telescopes verified the spiral arm structure as recently as the 1950s! You will repeat this experiment in this lab course, and obtain the spiral arm structure of the Milky Way (see Fig. 9.0.1).

But you will explore even more: Since the 1970s, astrophysicists have been puzzled by a new mystery: that of so-called dark matter. Most of the mass of our Galaxy and other galaxies seems to consist of some form of invisible matter. Dark matter acts solely through gravitational attraction. Despite numerous experiments, it has not yet been possible to detect electromagnetic radiation from this type of matter. Likewise no interaction with baryonic matter could be observed. The prevailing opinion considers dark matter as non-baryonic matter. Despite many years of experiments at CERN, in Switzerland, no elementary particle has yet been found that exhibits the physical properties of dark matter. Therefore, its existence is only indicated by it effect on visible matter. All attempts to investigate dark matter are based solely on

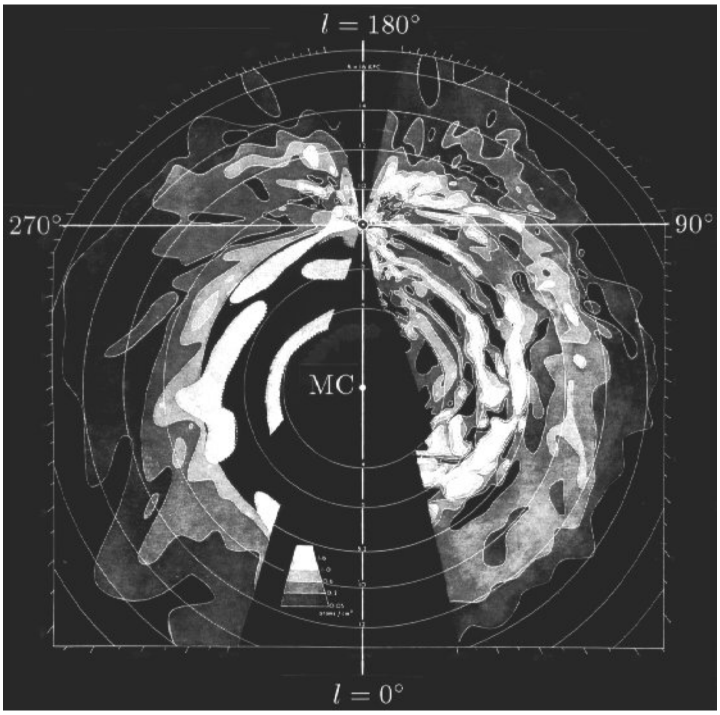

Fig. 9.0.1: The first map of neutral atomic hydrogen published by Oort, Kerr and Westerhout in 1958 [5]. You will repeat the experiment of these three scientists in the following.

its gravitational effect on gas and stars within its range. This effect we will also find with this experiment.

Dark matter, or invisible matter, was introduced as a hypothesis by Fritz Zwicky in 1933. He observed that the galaxies of the Coma galaxy cluster are moving so fast, that the visible matter of the cluster is not sufficient to bind the cluster with its self-gravity. Zwicky deduced, that if galaxy clusters are stable and closed objects, i.e. have existed for a long time already, then they must contain more invisible than visible matter. Zwicky's insight also applies to galaxies, such as the Milky Way. The stars and the gas in our galaxy rotate too fast around the galactic center to be kept on a closed orbit by the mass of visible matter alone. The mass of some hundreds of billions of stars is not sufficient to describe the rotational motion of the galaxy as a stable object. And it is not a small amount that is missing. The visible mass is too small by almost an order of magnitude compared to what would be needed to compensate for the centrifugal force experienced by the stars. A popular solution to this obvious problem is the hypothesis of dark matter, whose extra mass and gravitational effect holds the stars in our galaxy together and prevents them from flying away. The main argument for the existence of dark matter, on the size scales of galaxies, is the measured velocities at large distances from their center. For our galaxy, you will take these measurements yourself in this experiment. On smaller scales, e.g. on the size of planetary systems or star clusters, the effect of dark matter is not detectable.

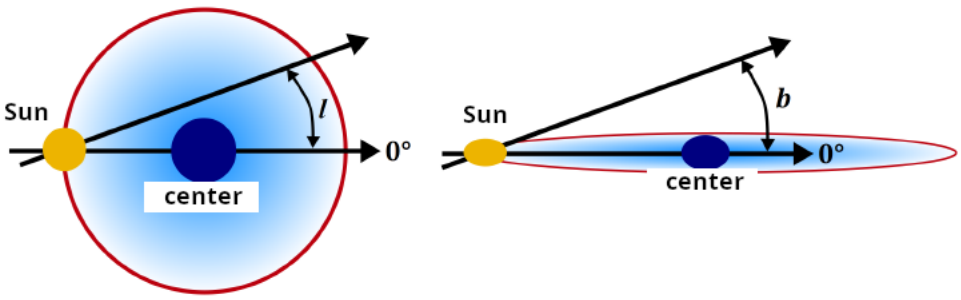

Fig. 9.0.2: How the galactic coordinate system is defined. Note that the origin lies at the position of the sun, and ($l = 0°$,$b = 0°$) corresponds to the position of the galactic center. l ("galactic longitude") and b ("galactic latitude") are measured in degrees. l increases counterclockwise and the Galactic latitude b takes positive values on the northern and negative values on the southern sky. The figure is taken from the bachelor thesis of H. Waimann from 2022, and has been translated to English.

9.1 Galactic coordinates: The position of the Sun in the Milky Way

Our solar system is located in the outer parts of the galaxy, at a distance of about 8.5 kpc (about 25 000 lyr) from the galactic center. Most of the stars and gas lie in a thin disk and rotate in the same sense of rotation around the galactic center. The Sun has a circular orbital velocity of about 220 km/s, and completes one full orbit around the center of the galaxy in about 240 million years.

To describe the position of a star or a gas cloud in the galaxy, it is useful to introduce galactic coordinates (l, b). Here l denotes the galactic longitude and b the galactic latitude (see Fig.9.0.2). The origin of the Galactic coordinate system is centered on the location of the Sun. $b = 0°$ corresponds to the Galactic plane, the center of the thin disk. The direction $b = +90°$ is called the Galactic North Pole, $b = -90°$ the Galactic South Pole. The longitude l is measured counterclockwise relative to the line connecting the Sun and the Galactic center. Thus, the Galactic center has the coordinates ($l = 0°$, $b = 0°$). There, in the galactic center, is indeed something very special: a very large mass concentration in the form of a supermassive black hole (SMBH). Its mass is about four million times the mass of the Sun. Although so much heavier than our Sun, the linear size of this supermassive black hole is about the dimension of our planetary system. Although we cannot directly observe the black hole itself, like dark matter, its gravitational effect on the stars and gas around it is essential for studying it. In the radio wavelength range, the Galactic Center is blindingly bright, and the source of this electromagnetic radiation is called Sagittarius A*. The strong emission of electromagnetic radiation in the radio is caused by electrons moving in the strong magnetic fields around the black hole. It is not the black hole itself that glows, but the baryonic matter in its vicinity does.

The galaxy is divided in four quadrants, that are numbered with roman numerals:

Quadrant I: $0° \leq l \leq 90°$

Quadrant II: $90° \leq l \leq 180°$

Quadrant III: $180° \leq l \leq 270°$

Quadrant IV: $270° \leq l \leq 360°$

In quadrants I and IV we observe mainly the inner part of our galaxy. Quadrants II and III contain material located at galactocentric radii larger than the distance of the Sun from the galactic center (the radius of the Sun's orbit around the galactic center). This finding will help us greatly in our study of the spiral structure of the Milky Way. Seen from the galactic north pole, the Milky Way rotates clockwise, i.e. the gas masses in the first quadrant move away from us, those in the fourth quadrant move towards us.

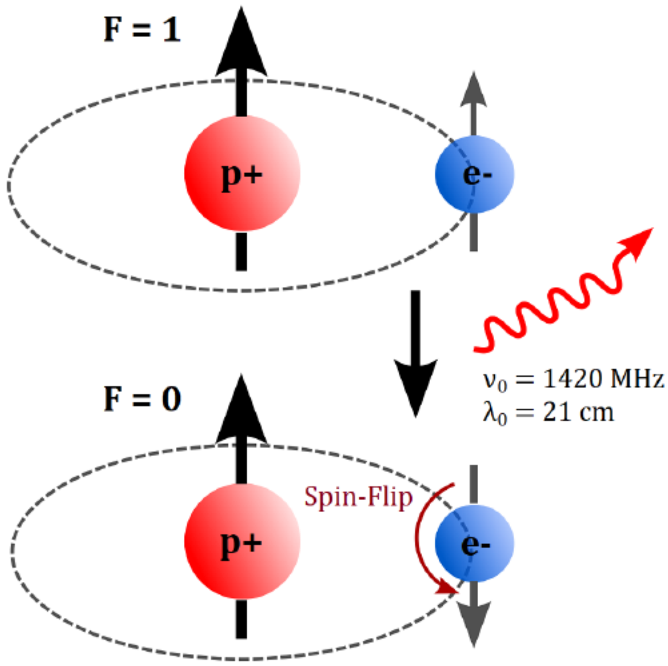

Fig. 9.1.1: The 21-cm line is formed by the re-orientation of the proton and electron spins relative to each other. The excited state corresponds to the parallel alignment of the two spins. The ground state, and thus the lowest energy state, is associated with the anti-parallel alignment of the spins. Such a transition is called a hyperfine structure transition. The energy released in this process is less than 1 millionth of the energy we can perceive with our eyes. In elastic collisions with atoms, ions and electrons lead to the parallel alignment of the spins. The system returns to the ground state only after an average lifetime of about 11 million years. The figure is taken from the bachelor thesis of H. Waimann from 2022

9.2 Hydrogen

Most of the gas in the galaxy is neutral atomic hydrogen (HI). HI is the simplest, but most common[1] atom in the entire universe: It consists of only one proton and one electron. Therefore it is electrically neutral. Close to very hot stars (spectral type O or B) hydrogen also exists in ionized form. This is called H^+ or HII. Far from hot

[1] by far (12 hydrogen atoms to one helium atom, both together add up to 99% of all baryonic matter in space)

stars, e.g. even in the photosphere of the Sun, hydrogen is neutral. Electron and proton form a stable unit. Neutral atomic hydrogen emits a radio emission line at a wavelength $\lambda = 21\,\mathrm{cm}$, or in terms of frequency $f = \mathrm{c}/\lambda = 1420\,\mathrm{MHz}$, (c denotes the speed of light). This is the signal we want to detect.

The 21-cm hydrogen signal is emitted by a neutral hydrogen atom when the spin of the electron changes from a parallel to an anti-parallel state relative to the proton spin (Fig. 9.1.1). Here the principal quantum number remains unchanged - in contrast to the visible lines of the hydrogen atom. This change is also called a hyperfine structure transition. The anti-parallel state of the two spins requires less energy. As an analogy, you could imagine two bar magnets that you put next to each other. If the magnets have the same polarity (alignment) you have to apply force to bring them closer together. This corresponds to the excited, the higher energetic, state. If the magnetic poles point in different directions, they stick together. This state requires no external force and corresponds to the lowest energy state a hydrogen atom can be in. The external force that leads to the parallel alignment of the spins is applied to the atom by inelastic collisions with atoms, ions, and electrons in the vicinity of the atom.

The hydrogen atom remains in the excited state for a very long time. For the transition from the excited to the ground state, the hydrogen atom needs about eleven million years. This very long average lifetime of the hydrogen atom means that the 21-cm emission line has an extremely narrow width. For the study of physical conditions in the interstellar gas (the gas between stars), this is an incredible advantage. Analysis of the observed line-width provides direct information about the density and temperature of the gas in which the atom is located. The observed line-width is therefor a direct indicator of the maximum kinetic temperature of the atoms and ions that collided with the neutral hydrogen atom. Although eleven million years seems very long, the HI 21-cm emission line is very bright. The reason for this is the enormous amount of hydrogen present in the Milky Way and in all galaxies, so the 21-cm line can be observed very clearly, and nowadays quite easily, with radio telescopes.

9.3 Doppler-effect

We can study the motion of hydrogen gas clouds in our galaxy using the so-called "Doppler–effect", which links the observed frequency of our neutral atomic hydrogen emission signal to the velocity of the emitting gas. The Doppler–effect, named after the austrian physicist Christian Andreas Doppler (1803-1853), can also be observed in everyday life: i.e. when you are standing in a street and an ambulance approaches. The siren of the ambulance appears to increase not only in volume (amplitude), but also in pitch (frequency). Similarly, the volume and pitch decrease once the ambulance has passed and is now moving away from you. Because sound waves propagate through a medium (the air), the waves are compressed by the forward motion of the ambulance as it approaches: the time-difference between the maxima of the pressure waves decreases. Shorter wavelengths correspond to a higher frequency, and therefor a higher pitch of the siren. This change in frequency due to the relative velocity of the source (ambulance) and observer (you) is the Doppler–effect.

To compute the motion of gas clouds we need a formula that describes this dependency of the frequency on the relative velocity. We have decided to skip the lengthy derivation of this formula and present you with the end result:

$$\frac{\Delta f}{f_0} = \frac{v}{\mathrm{c}} \tag{101}$$

where $\Delta f = f - f_0$ is the frequency shift, f the observed frequency, f_0 the (emitted) rest frequency of the 21-cm line, and v the relative velocity. We use the following sign convention: When the source is moving away from us $v > 0$. For a source moving towards us $v < 0$.

Origin of the interstellar velocity scale

To derive the motion of the emitting gas clouds in the galaxy, we set the receiver of our radio telescope to measure frequencies close to the rest frequency of the 21-cm line, meaning the frequency we would observe if there was no relative motion between us and the cloud. The reference frame for velocities is called the *local standard of rest* (LSR). It is not relative to the movement of the sun around the galactic center. The LSR is defined to be the center of the average motion of 30 stars close to the sun. Relative to the LSR even the sun is moving.

Clouds with different velocites, relative to the LSR, are observed at different frequencies. The frequency is shifted up (blue) when the gas cloud moves towards us, and down (red) when it is moving away.

Rotation curve and spiral structure of the Milky Way

In this chapter we describe how you can use a small radio telescope to measure the structure and kinematics of the Milky Way. First, we explain how to measure the rotation curve, that is, the rotational velocity of the gas at various distances from the galactic center. Then we use our results to create a map of the spiral arms.

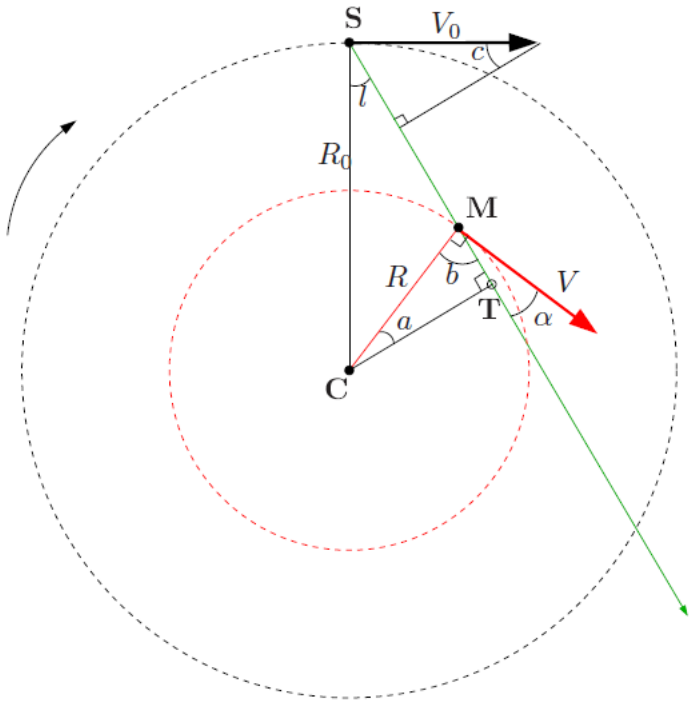

Fig. 9.3.1: Geometry of the galaxy. C denotes the position of the galactic center, S that of the sun and M that of a gas cloud we want to observe. The line between S and M is the line of sight (LOS). The radio telescope receives the HI 21-cm line radiation of interstellar matter along this line of sight. The arrows on the line segments indicate the velocity of the sun (V_0) and the gas cloud (V). Please note the direction of rotation of the Milky Way, represented in this figure by the curved arrow in the upper left. Viewed from the galactic north pole, the Milky Way rotates clockwise. In the first quadrant, the maximum velocity is found at positive v_{LSR} in the fourth quadrant at negative v_{LSR}. This figure was taken from the SALSA script.

9.4 Rotation curve

A rotation curve is a function that describes the rotational velocity of the interstellar gas of a galaxy at different distances from its center. The rotation velocity is usually denoted by $V(R)$. To derive a rotation curve, we must first understand how the velocity we measure through the Doppler shift is related to the motion of the gas clouds in the Milky Way. Let us imagine that we are pointing our radio telescope at a gas cloud in the Galaxy, i.e., we are observing along the green line in Fig. 9.3.1. For further information, we list some variables that are used in this figure

V_0	the velocity of the sun around the galactic center, roughly 220 km/s
R_0	the distance of the sun to the galactic center, 8.5 kpc
l	galactic longitude at which the gas cloud is located
V	velocity of the observed gas cloud
R	the distance of the gas cloud to the galactic center

There may be many gas clouds along this line of sight, but for the purpose of this derivation we consider only a single cloud at position M in Fig. 9.3.1. Since both the sun and the cloud are in motion, we do not directly measure the absolute velocity of the cloud in space. Instead, we measure only the relative velocity, V_r, of us and the cloud. Using the Doppler effect me measure only the projection of the velocity vector on the line of sight, hence the subscript r for "radial". Using the angles in Fig. 9.3.1 we can write:

$$V_r = V \cdot \cos(\alpha) - V_0 \cdot \sin(c)$$

For this expression to be useful, we need to relate the angles to the galactic coordinates we discussed in Section 9.1. We know that the sum of the angles in the upper right-angled triangle (the one containing point S) must be 180°, which means that we can relate the angle c to the galactic longitude l:

$$(90 - l) + 90 + c = 180 \quad \Rightarrow \quad c = l$$

We now also want to relate the angle α to the length l. We find that the angle between V and R is 90° and can be written as the sum of a and α, i.e. $90° = b + \alpha$. From the triangle CMT we can also find that $b = 90° - a$. Putting all this together we obtain:

$$90 = 90 - a + \alpha \quad \Rightarrow \quad a = \alpha$$

Looking at the triangles CST and CMT, we find that the distance between the galactic center (C) and the tangential point (T) can be expressed in two different ways:

$$CT = R_0 \cdot \sin(l) = R \cdot \cos(a) = R \cdot \cos(\alpha) \quad \Rightarrow \quad \cos(\alpha) = \frac{R_0 \cdot \sin(l)}{R}$$

This is the relation between l and α we have been looking for. Using it we find:

$$V_r = V \cdot \frac{R_0}{R} \cdot \sin(l) - V_0 \cdot \sin(l) \tag{102}$$

This equation is valid for all longitudes l. However, measuring V_r for any l alone is not sufficient to solve this equation and derive both V and R. The solution is to restrict the range of possible l to the first quadrant and use only the determined maximum velocity for our calculations. Let's see how this simplifies the problem!

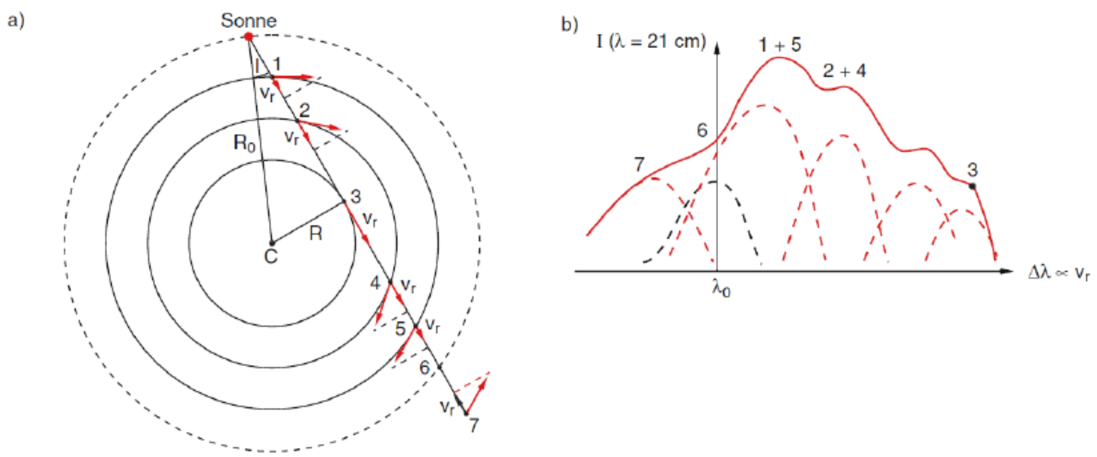

Fig. 9.4.1: This figure shows how we measure the radial velocity component (left) of the HI 21-cm line emission from seven gas clouds along the line of sight in the spectrometer (right). Where the full velocity vector exactly coincides with the radial component, we measure $V_{\rm r,max}$, this is position 3. Positions 2&4 and 1&5, cannot be uniquely determined by the spectral measurement. If we move the radio telescope slightly north or south from the plane ($b = 0°$), the spectral emission of 1 and 2 is expected to be preserved. The more distant gas clouds 4 and 5 are too far away to be observed that far above the Galactic plane. Figure from [2].

In any given direction, we can observe emission from multiple clouds simultaneously (see Fig. 9.4.1). Since the clouds move at different relative velocities, we measure multiple components in the spectrum. The largest velocity component, $V_{\rm r,max}$, is from the cloud at the tangential point (3, Fig. 9.4.1 or T in Fig. 9.3.1). This assignment is geometrically unambiguous, since the maximum possible projected velocity is observed when the projection angle is $\alpha = 0°$. For a cloud at the tangent point we see in Fig. 9.4.1 that the position of the cloud is given by

$$R = R_0 \cdot \sin(l) \tag{103}$$

This simplifies Eq. 102, and for the tangential point we find:

$$V = V_{\mathrm{r,max}} + V_0 \cdot \sin(l) \tag{104}$$

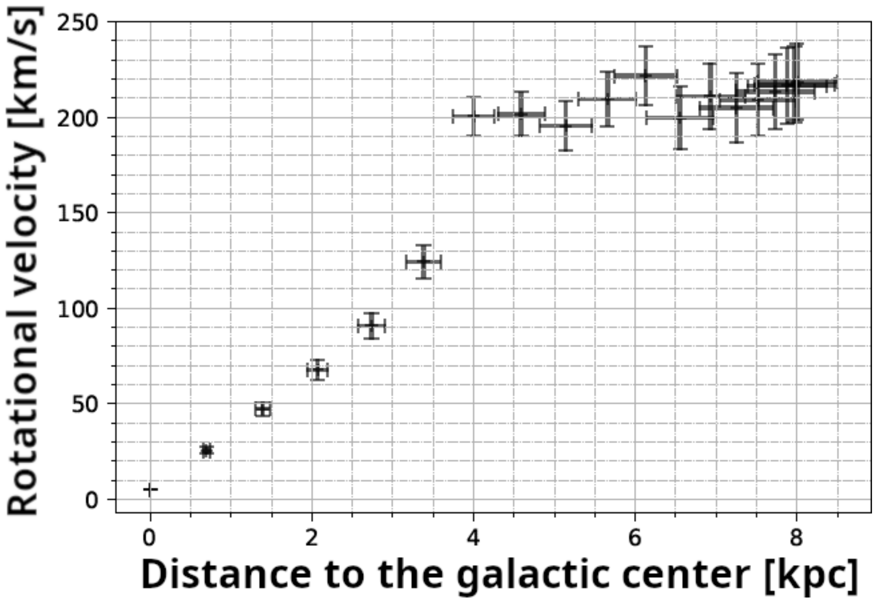

Fig. 9.4.2: Rotation curve of the Milky Way derived from observations of the HI 21-cm line emission with a 2.3 meter telescope. Taken from the bachelor thesis of R. Hürten 2022 and translated to English.

We can now use the radio telescope to measure $V_{\mathrm{r,max}}$ at different l in the first quadrant. Using equations 103 and 104 we can then derive the rotation curve $V(R)$. Assuming that the Milky Way is a symmetric spiral galaxy, it is reasonable to assume that the measured rotation curve is also valid in the other quadrants.

After reading the previous sections, you should have an idea of how to measure the rotation curve of the Milky Way with a small radio telescope. Instructions on how to operate the telescope can be found in another document, the AIfA SRT User's Guide. Once you have obtained spectra at a few selected longitudes, extract the maximum velocity of each spectrum and plot the rotation curve. Immediately after recording the spectrum, you can use the cursor, directly in the control program, to determine the maximum velocity. However, it makes more sense if to examine the data saved as text or FITS file at your own pace. The goal is to obtain a result that is reproducible. In the appendix you will find a short Python script which allows you to read in the spectra as a FITS or CSV file, subtract the baseline and perform a Gaussian decomposition. Essential here is the subtraction of the baseline and the determination of the thermal noise, because you will have to decide which emission is statistically significant and which disappears in the noise.

For the final plot of the rotation curve, you can use your favorite plotting program. A simple solution is to use the Excel-like program `LibreOffice Calc`, which you use to enter the calculated values for V and R into a spreadsheet and then plot the data using the xy-chart function. Your final graph should look similar to Fig. 9.4.2. The origin of the coordinate system is at the Galactic Center. The Sun is located at about 8.5 kpc in this plot.

You can see two components in Fig. 9.4.2: Inside, the rotation of a rigid body with $v_\mathrm{r} \propto R$ and, at distances greater than 4 kpc, a constant rotational velocity. The latter is clearly different from the Keplerian rotation e.g. in the solar system. The orbital velocity of the planets around the Sun behaves like $v \propto \frac{1}{\sqrt{r}}$ with r the distance of the planet from the Sun. Physically, this Keplerian velocity gradient means that the dominant mass fraction is completely enclosed by each of the planetary orbits. However, the constant rotational velocity of the Milky Way tells you that none of your measurement points is so far from the center of mass that the entire mass is completely within the orbit of the gas cloud. Therefore astronomers speak of the "dark matter halo" which described a wide-spread distribution of the dark matter around the stellar and gas disk of a galaxy. For large spiral galaxies a Keplerian decay of the radial velocity could not be measured in any way so far.

This is essential for the following part of our experiment. Assume, that the rotation curve is flat for $R \geq 4$ kpc meaning that there $V(R) = V_0$.

9.5 The spiral structure of the Milky Way

Now we want to find out where the HI gas we discovered is located. Therefore we return to equation 102. When we determined the rotation curve, we only determined the maximum velocity component in the spectrum and assumed that it came from gas at the tangent point. This assumption allowed us to solve Eq. 102 in the first quadrant. Now we will use **all** velocity components that we see in our spectra, and we want to be able to map them in all observable directions, not just the first or fourth quadrant directions. We can now directly use our knowledge of the rotation curve obtained in section 9.4. Motivated by the shape of our measured rotation curve, we may assume that the gas in our Milky Way obeys differential rotation, i.e., the rotation velocity is constant with $V(R) = V_0$. Differential means here, since V is constant, that apparently the gas clouds lying further out ($R > R_0$) have a lower angular velocity, since these need longer for an orbit around the Galactic center.

$$\omega(R) = \frac{V}{R} = \frac{V_0}{R}$$

$\omega(R)$ is the angular velocity of the cloud at distance R, as observed from the Galactic Center. Since $V = V_0$, ω becomes smaller as R increases. The gas clouds at large R have a longer orbital period ($t = \frac{2\pi \cdot R}{V}$) around the Galactic Center. Therefore, from the point of view of the Galactic Center, their angular velocity is lower. For this reason we speak of a differential rotation. This property of the rotation of the Milky Way around its Galactic Center had been derived by Jan Hendrik Oort already in 1927.

Assuming this, Eq. 102 simplifies to:

$$V_r = V_0 \cdot \sin(l) \left(\frac{R_0}{R} - 1 \right)$$

This equation can be rewritten as an expression for the cloud distance R in terms of the known (or rather observable in case of V_r) quantities as:

$$R = \frac{R_0 \cdot V_0 \cdot \sin(l)}{V_0 \cdot \sin(l) + V_r} \tag{105}$$

We now want to map the Milky Way and determine the position of the cloud we have detected. From our measurement of the radial velocity V_r we have just calculated the distance of the cloud to the galactic center R and we also know in which direction we have pointed our telescope (the galactic longitude l). Please note that the position of the emitting gas clouds can only be determined unambiguously if you observe in quadrants II or III. However, if you observe in quadrants I or IV, there may be two possible positions corresponding to the given values of l and R: closer to us than the tangential point T (the actual point M in the figure), or farther away, at the intersection of the ST line and the inner circle (see Fig. 9.4.1).

This ambiguity of distance can also be represented mathematically. Let r denote the distance from the sun to the cloud, i.e., the distance between points S and M in Fig. 9.3.1. Using the law of cosines on the triangle CSM we obtain the following relation:

$$R^2 = R_0^2 + r^2 - 2 \cdot R_0 \cdot r \cdot \cos(l)$$

This is a second order equation in r that has two possible solutions. Using the pq formula we can write the solutions as:

$$r_\pm = \pm \sqrt{R^2 - R_0^2 \cdot \sin^2(l)} + R_0 \cdot \cos(l) \tag{106}$$

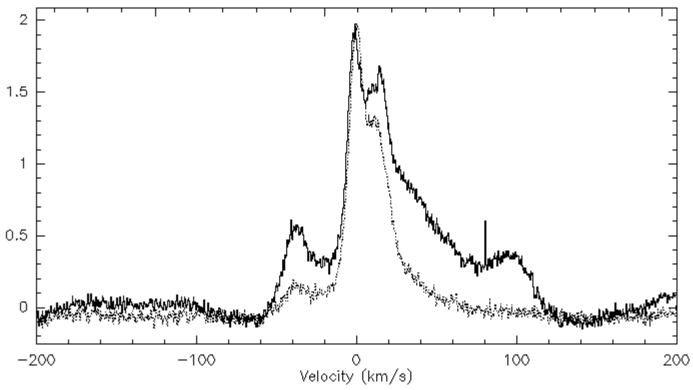

Fig. 9.5.1: This figure shows two spectra in the direction $l = 25°$, at $b = 0°$ (solid line) and $b = +5°$ (dotted line). The component at $v_{\rm LSR} \simeq 100\,{\rm km\,s^{-1}}$ is not present in both spectra. That means, this gas cloud is at too great a distance to belong to the spiral arm under study. The other components, however, are present in both spectra.

There are a few things to note about the above equation. For some data points, one finds a positive and a negative solution for r. The negative solution does not make sense (it would point to a position on the opposite side of the sun) and should be discarded. In quadrants II or III, R is always greater than R_0 and $\cos(l) < 0$, which means that there is only one positive solution, r_+. In quadrants I or IV, there can be two positive, and therefore possible, solutions. In cases with two possible solutions, it is not possible to determine which solution is the correct one without additional observations. To resolve this ambiguity, one can again observe in the direction of the same Galactic longitude, but in the direction of a slightly non-zero (few degrees) Galactic latitude. If the ambiguous cloud is far away, it should not be seen. If the cloud is close, it should still appear, even if observed slightly out of the Galactic plane. Some experimentation may be necessary here to find the appropriate Galactic latitude.

Cartesian representation of polar coordinates

For graphical representation, it may be impractical to use the coordinates r (distance of the cloud from the Sun) and l (galactic longitude) to describe the cloud positions. Instead, it is usually more convenient to convert these coordinates to a Cartesian system of perpendicular coordinates x and y. To convert to x - y coordinates, we need to relate them to r and l.

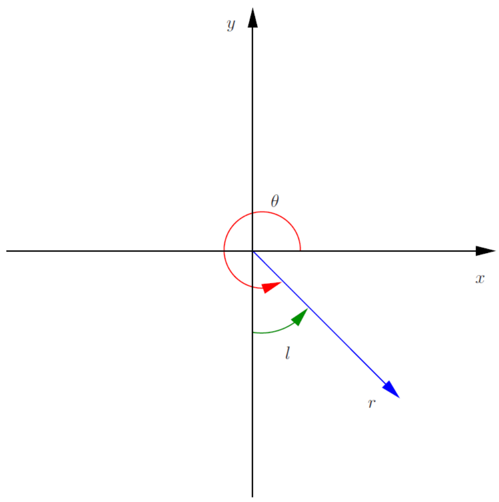

Fig. 9.5.2: Connection between the Cartesian (x,y) and polar coordinate system (r,θ) and the measured coordinates (r,l). The figure was taken from the SALSA script

You are probably familiar with polar coordinates, which are usually defined as follows:

$$x = r \cdot \cos(\theta)$$

$$y = r \cdot \sin(\theta)$$

where r is the distance from the origion, and θ the angle. x and y are the Cartesian coordinates. When we plot the two together in Fig. 9.5.2 it becomes clear, that polar coordinates are very similar to our r-l-system. We find that $\theta = 270° + l$, or $\theta = l - 90°$. This means that we can convert our positions given as r, l to Cartesian x-y coordinates:

$$x = r \cdot \cos(l - 90°)$$

$$y = r \cdot \sin(l - 90°)$$

This format is usually best for plotting our maps of the Milky Way. Note that this is the map with the Sun at the origin, at position (0,0). If instead you want the galactic center to be at position (0,0), add R_0 to the y coordinate of each point.

Now you should have an idea of how to make a map from the measured velocities. The observations are performed in the same way as in section 9.4, but now you don't have to and shouldn't measure exclusively in the first quadrant. Also, you should extract all velocity components in your spectra, not just the maximum component. Again, instructions for operating the telescope can be found in another document, the AIfA SRT User's Guide.

The easiest way to obtain velocities from a spectrum is to use the cursor to examine the spectrum directly in the control program. Reproducible results can be obtained by analyzing the FITS or CSV files. The provided Python script allows you to fit up to four Gaussian components to the spectral data in addition to determining the *baseline*, and retrieve the center velocities.

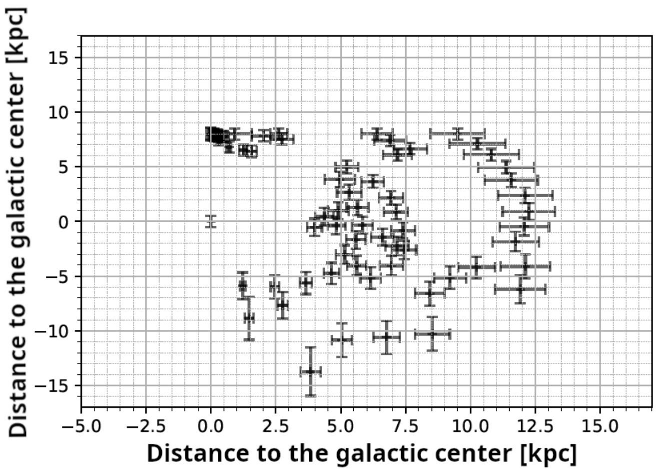

Fig. 9.5.3: Reconstruction of the spiral structure of the Milky Way from measurements taken with a 2.3 m telescope. Figure taken from the bachelor thesis of R. Hürten 2022, translated to English.

Once you have a list of the measured velocities at several galactic longitudes, you can use equation 105 to derive the position of the clouds. Note that more measurements are required to clarify ambiguities, when observing clouds in quadrant I or IV. In this case use Eq. 106, as discussed in the previous section. Use a plotting software of your choice to create the final map. Here, too, Excel or `LibreOffice Calc` can help you out: Enter the calculated positions (x, y) into a worksheet and create an xy-diagram. Your final map should look somewhat similar to the map in Fig. 9.5.3, although it depends on the direction of your observation how much of the galactic disk you can see.

9.6 The radio telescope

For this experiment you use a radio telescope. The goal of this section is to give you an idea of how this radio telescope works. If you understand the operating principle of a mirror telescope, you are already half-way there!

The number of photons your receiver detects, depends on the size of your main (or "primary") mirror (also called a "dish" in radio astronomy). The larger the collecting area, the more photons are bundled in the focus of the dish. Here, too, the known formula for the angular resolution of a circular aperture holds:

$$\theta = 1.22 \cdot \frac{\lambda}{D}$$

where θ is the angular resolution of your telescope in units of radians, λ the wavelength of the observed radiation (0.21 m in our case), and D the diameter of the aperture.

You will need to determine the angular resolution in the context of the project in order to meaningfully search for the far components in the spectrum in the first and possibly fourth quadrants (see Fig. 9.5.1).

The receiver of our telescope is placed in the primary focus. We also refer to the receiver as the *front end*. On a very basic level, it is a system capable of electromagnetic oscillation. There is a copper rod, a dipole, mounted inside. Within the dipole incoming electromagnetic waves accelerate the electrons. The response of the system is maximal in the resonant case:

Imagine two tuning forks: one is the emitter, the other the receiver. The vibrations of the emitter are transmitted to the receiver through sound waves. Even if the latter does not vibrate at first, the emitter will make it vibrate at the same, resonant, frequency.

The dipole in our receiver is tuned to the resonant frequency of the radiation we want to detect.

Naturally, our receiver system comes with losses, which smears our resonant frequency out into a frequency band, that best matches the incoming radiation. Outside of this frequency band the losses are so large, that we do not detect radiation with these frequencies. We call this behaviour a bandpass filter. When you take a look at the incoming signal during the experiment, you can nicely see this effect. Outside of our bandpass frequency range, the antenna temperature T_{A} drops to zero very quickly.

But what is the antenna temperature? Let's go back to the Rayleigh–Jeans law: Our approximation of the Planck law for the blackbody spectrum at the large wavelengths we work with in radio astronomy. The power of a blackbody emitter then is approximately:

$$B_\nu(T) \propto T \cdot \nu^2$$

with T the effective temperature of the blackbody, and $\nu = \frac{c}{\lambda}$ our observing frequency. The radiation flux that we measure is called the antenna temperature T_{A}. The value you measure depends on your telescope and the external conditions during your observation. Different telescopes give different antenna temperatures for the same source. It is an *uncalibrated* observable.

Sadly, or rather luckily, the photon energy of the 21-cm line is merely 1/1,000,000th of that of visible light. Therefore almost everything, even non-astronomical sources, emits close to the energy of the 21-cm line. Even you emit photons of this wavelength. Try placing your hands in front of the receiver. The measured power level will rise. In Fig. 9.6.1 you can see this background noise. It covers the entire spectrum. This is because we receive much more radiation from all kinds of directions than from the source itself.

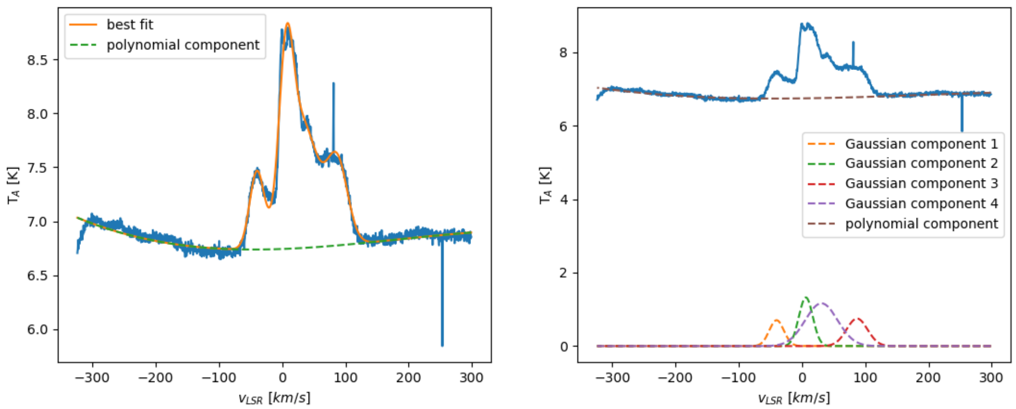

Fig. 9.6.1: Left: Observed spectrum with a 2.3-m radio telescope. The line emission from the Milky Way is superimposed on the continuous radiation that the telescope receives from its surroundings and generates itself. The *polynomical component* models this radiation component, which is composed of a large number of radiation components. It is also referred to as *baseline* in the script. Right: Same spectrum, but the contributions of the continuous components are plotted separately from those of the HI 21-cm line. The amplitude difference between the two signals is clearly visible. The approximation of the line emission by four Gaussian components is sufficient for a good approximation of the spectrum. This graph was created with the ipython notebook, which will be provided.

This means that our astronomical line emission is embedded is a large number of photons that have nothing to do with what we are trying to observe. The largest source of thermal noise is the receiver itself. Its un-cooled amplifier, a *high electron mobility transistor* (HEMT for short) has the highest noise temperature of the entire frontend. We will determine its temperature later on. By cooling the receiver we could therefore increase the quality of our observation. However our telescope receiver does not have a cooling system.

Let us introduce the system temperature T_{sys} at this point, which is sufficient to describe the state of our telescope. It is dominated by thermal noise of the HEMT–amplifier. But the weather, the brightness of the Milky Way, and the cosmic microwave background also contribute to the system temperature.

Now, how do we differentiate between our source and the dominating background noise?

The idea is simple: Instead of looking at the total received power, we use a spectrometer as our *backend*. It counts how often a given temperature is measured at a certain frequency, after the amplification.

White noise has a uniform probability distribution, meaning that the probability to detect a photon at each frequency is always the same. But our astronomical line emits only at one frequency! Meaning that photons from our line only appear within a small velocity and thus frequency range. Their probability distribution differs significantly from the uniform white noise. The HI 21-cm lines we are interested in are the Gaussian components on the right in Fig. 9.6.1.

Your task now is to differentiate between the noise and astronomical signals. To do that, you have to derive the noise width σ. Based on the discussion above, the noise increases with the system temperature: $\sigma \propto T_{\text{Sys}}$. You can improve the (statistical) accuracy of your measurement, by including as many photons as possible in your analysis, even those that originate from background noise.

So integrating over a long time Δt ($\sigma \propto 1/\Delta t$) and detecting high energies $E = \text{h} \cdot \nu$. Therefore $\sigma \propto 1/\Delta\nu$, where $\Delta\nu$ is the frequency width of the bandpass filter. From this you can get a rough understanding of the radiometer equation:

$$\sigma = \frac{T_{\text{Sys}}}{\sqrt{\Delta t \cdot \Delta\nu}} \tag{107}$$

where Δt is the integration time, and $\Delta\nu$ the bandwidth.

With this equation you can now compute the width of the noise. Default values are: $\Delta t = 20\,\text{s}$ and $\Delta\nu = 1464\,\text{Hz}$. Inserting these and assuming an upper limit of room temperature (305 K) for the system temperature we obtain:

$$\sigma = \frac{305\,\text{K}}{\sqrt{20\,\text{s} \cdot 1464\,\text{Hz}}} = 1.78\,\text{K}$$

To detect a signal with 99% certainty, it has to have an amplitude of atleast 3σ, so in our case 2.36 K.

Determining the system temperature

Finally, you are to determine the absolute value of T_{Sys}. This is quite simple. We consider the position $l = 25°$ and $b = 0°$. You determine the noise σ_{T_A} in your spectra on your T_{A}-scale right next to the hydrogen emission lines, but not on these lines themselves. This is where the iphyton script helps you. With it, you can determine the

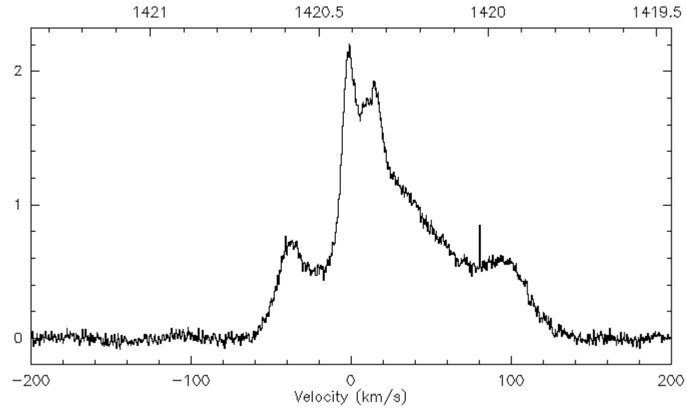

Fig. 9.6.2: Calibration spectrum: this is an observation of the HI 21-cm line emission at position $l = 25°$ and $b = 0°$ made with the 25-m telescope at Dwingeloo (Kalberla et al. 2005). It was reduced to the angular resolving power of an SRT. The spectrum is absolutely calibrated in its flux, so you can consider the intensity maximum to be absolute and compare it to the one you measured. The quotient of the two values is the calibration factor.

left side of Eq. 107. However, the amplitude of the noise is still an uncalibrated value since it is based on T_A. Since our system is linear, we can quantitatively compare the measured spectrum with a calibrated spectrum. We read the maximum value of the calibrated spectrum and compare it to the maximum value in our measured spectrum. The quotient of both values is your calibration factor. With this you get σ absolutely calibrated by to multiplying with σ_{T_A}. With this you are almost done, because you know the integration time $\Delta t = 20\,\mathrm{s}$ and the frequency resolution $\Delta\nu = 1454\,\mathrm{Hz}$ of the spectrometer. Use this to determine T_Sys. It is lower than 305 K, I promise.

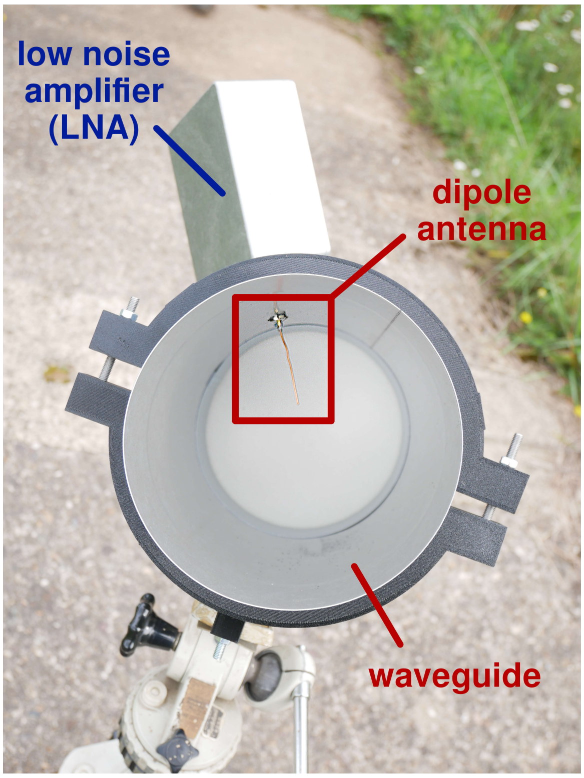

Fig. 9.7.1: Frontend: Here, we look straight forwardly into the waveguide. Labeled is the dipol, the antenna of our radio telescope, a straight copper wire. The waveguide itself is a cylinter with a metal plate at one end. The box is host to the HEMT and the bandpass unit.

9.7 Setup of the radio meter

The radiometer is the heart of a radio telescope. It comprises the so-called frontend, which receives the astronomical signal from space, and the backend, which is eventually used for data analysis and displaying purposes.

Let us focus on the frontend first.The first contact of the astronomical signal is with the dipole. At long wavelength a simple dipole receives the electromagnetic wave directly. Recall cars and the dipole attached to the car body for listening music and broadcast news. At the 21-cm wavelength of neutral atomic hydrogen however its is better to put the dipole antenna, this is a copper wire about 5cm long, inside the waveguide. The waveguide is a one end closed cylinder of about 15 cm diameter, with a short circuit which optimally reflects the space wave onto the dipole with a separation between both of about $\lambda/4$. This construction accounts for the fact, that at the short circuit a standing wave zero point is localized and the $\lambda/4$ distance to the copper wire warrants a maximum of the amplitude.

Directly attached to the dipole is the low noise amplifier (LNA). In practice, this electronic part consists out of two components: a high electron mobility transistor, which amplifies the signal by orders of magnitude, and a band pass filter. The band pass filter blocks all unrelated emission at neighboring frequencies.

The amplifier needs a power supply. It is realized by the cable which also transmits the amplified signal received. For this aim a bias tee. A bias tee has an input and an output channel plus a connection to a power supply unit which sets the DC bias point. By construction the bias tee can be placed close to the spectrometer. A location within the immediate vicinity of the amplifier is not necessary.

Now the signal reaches the backend of the radiometer. The celestial signal is connected via the bias tee to the spectrometer. We use for the dispersion of the radio astronomical signal a software define radio (SDR). It samples digitally the frequency portion received an amplified by the frontend. An analog to digital converter (ADC) samples the signal. The complex values of the ADC is transmitted via a USB-C cable to the Raspberry-PI mini computer. Here a fast-fourier transformation is calculated an the power-spectrum can be displayed and analysed.

This principle setup of a radio meter is a basic concept and demands only minor changes at cm wavelength range. The length of the dipol antenna has to be adopted as well as the HEMT transistor and bandpass. The backend remains however unchanged.

9.8 Raspberry-PI

The Raspberry-PI is mini or single board computer. We are using this mini computer with a Ubuntu Linux distribution. For this operating system a wealth of hardware modules are available. The operation system is stored on a micro-SD chip at the backside of the Raspberry PI. At this SD card stored is also a python distribution as well as the programs necessary to operate the SDR.

Keyboard and mouse can be connected via the USB-2 connectors. Next to them is a pair of USB-3 connector located. A monitor needs to be connected via the micro HDMI to HDMI adapter.

To the left of the two micro-HDMI connectors the power supply of the Raspberry-PI is visible. When connecting the Raspberry-PI to the power connector it starts immediately booting. The "user" is simply "pi" without a password. The boot procedure last up to one minute. Then the Raspberry-PI is ready for radio astronomical observations.

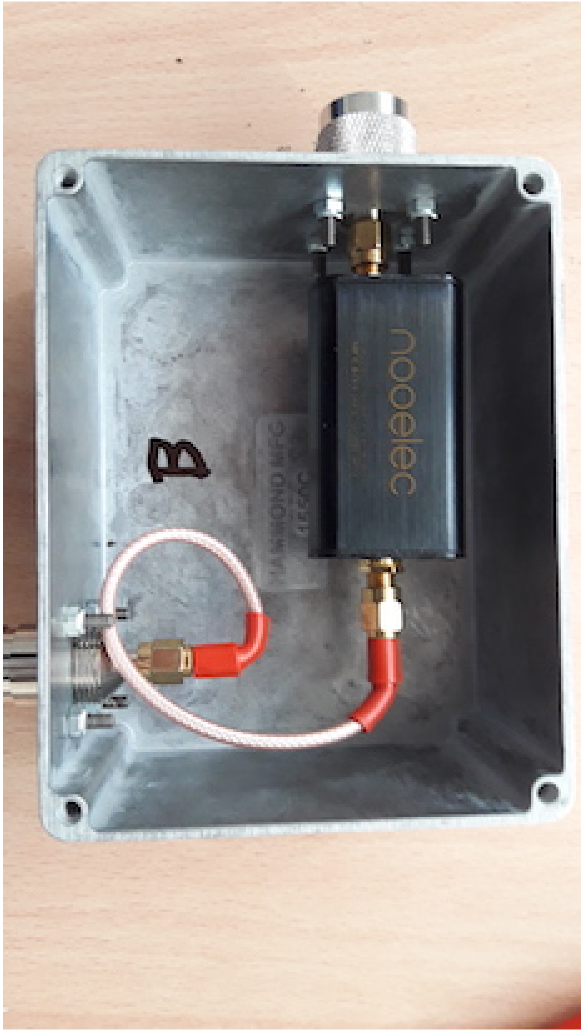

Fig. 9.7.2: View inside the aluminium box hosting the HEMT and band pass. This device amplifies the signal and limits it to the HI 21-cm line at 1420.4 MHz. Because of mechanical stability the aluminium box is equipped with N-type connectors. Within the box the SMA type connectors are used. This device need a DC voltage. It is provided by the bias-tee element and the connected high frequency cable.

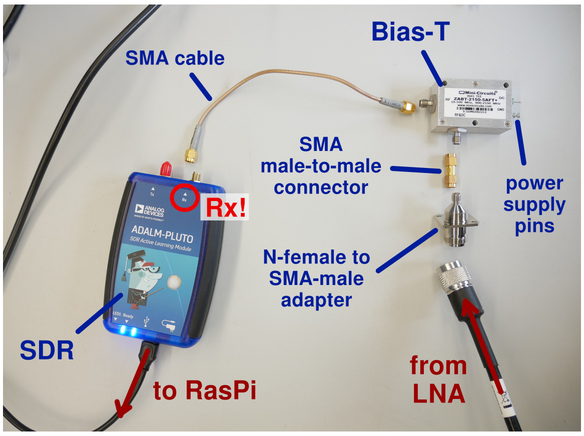

Fig. 9.7.3: Backend: The HF-cable with the attached N-connector comes from the frontend . We need to adapt this to the bias-tee. We use a N-female to SMA-male connector and an SMA-male-male connector here. The bias-tee itself is connected to a DC 5 V power supply. This provides the power for the amplifier. The bias-tee output is connected via a SMA-cable to the SDR. Which itself is connected with a USB cable to the Raspberry-PI mini computer.

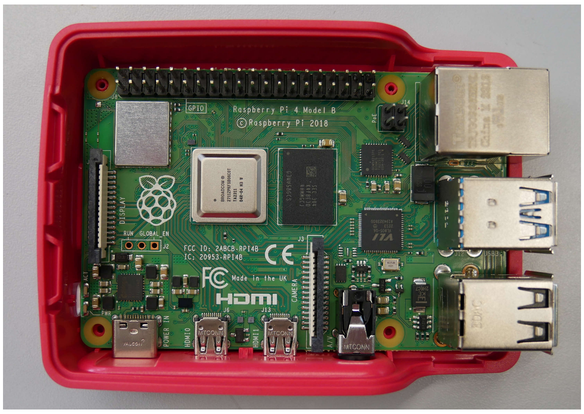

Fig. 9.8.1: Top view onto the Raspberry-PI version 4.0 mini or single-board computer. On the right hand side the network connector as well as the USB ports are visisble. Towards the bottom the micro-HDMI connectors and the power supply plug is visible.

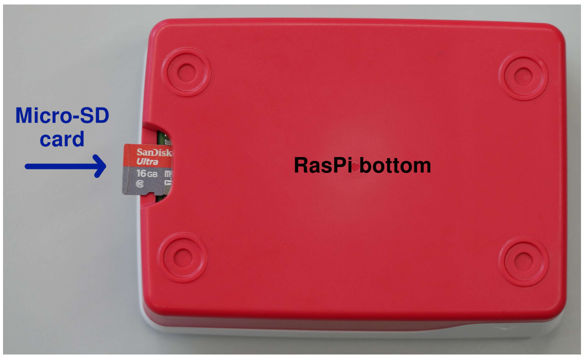

Fig. 9.8.2: The whole operating system, here a Ubuntu Linux distribution, as well as Python and the radio astronomical software is stored at the micro-SD card. If it is not inserted correctly the Raspberry-PI might not boot.

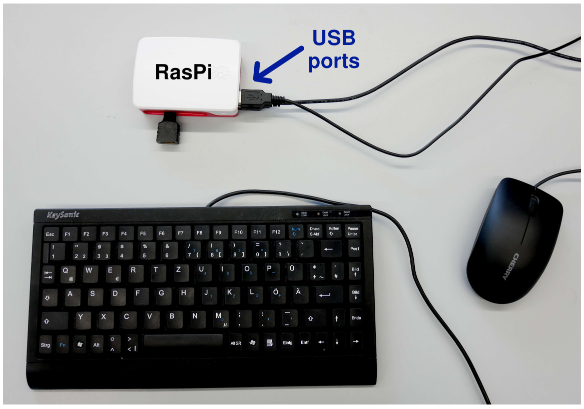

Fig. 9.8.3: Keyboard and mouse are connected to the Raspberry-PI via the standard USB-2 plugs on the right hand side of the board.

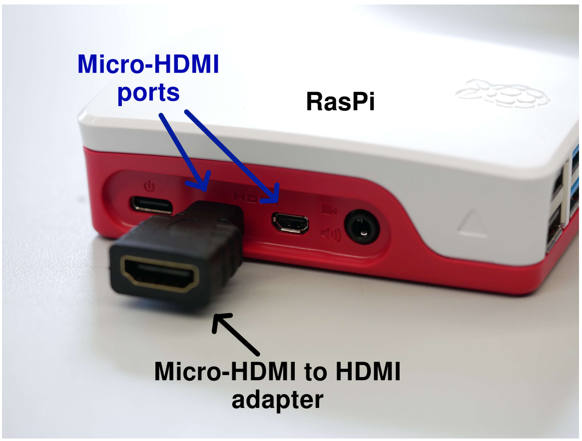

Fig. 9.8.4: Because of the tiny size of the single board computer only micro HDMI plugs are available. One needs an adapter to connect a standard HDMI connector. On the left to the micro-HDMI plugs the power supply to the Raspberry-PI is located. When connecting the Raspberry-PI to the power, the mini computer will boot immediately.

9.9 Radiometer Equation

The noise of a radio astronomical receiver follows the radiometer equation. The aim of this measurement is to test, whether your equipment is limited by systematic issues or only by the noise of the amplifier. In the latter case, the data points would distribute according to the radiometer equation. The test is a critical one for electronic devices, revealing their quality.

To do so, one needs to perform measurements of the same source with constant conditions. In cm-wavelength radio astronomy a wall can serve as black body radiator.

We measure the irradiation of our radiometer with constant integration times. 20 seconds in example per measurement. We repeat the measurements 10, 20 or a 100 times, up to the maximum integration time we are aiming to accumulate.

Now we use a program which calculates the sum of the measurements:

$$a_0 = \textit{1st spectrum}$$

$$a_1 = \textit{1st spectrum} \quad + \quad \textit{2nd spectrum}$$

$$a_2 = \textit{1st spectrum} \ + \ \textit{2nd spectrum} \ + \ \textit{3rd spectrum}$$

$$... = ...$$

We plot the a values on a logarithmic y-axis and time axis and obtain the Allan variance plot. In the case the our radiometer is only limited by statistical uncertainties, we would observe a lowering to the noise proportional to $\frac{1}{\sqrt{t}}$, in the double logarithmic plot a straight declining line with a slope of -0.5. If our device is limited by some systematic effects at a certain integration time the noise value remains unchanged and constant. Implying that it is not meaningful to spend more time in observation because we do not gain information on fainter structures.

9.10 Hot–Cold Calibration

References

[1] Suzanne Débarbat, James Lequeux, and Wayne Orchiston. "Highlighting the history of French radio astronomy. 1: Nordmann's attempt to observe solar radio emission in 1901". In: *Journal of Astronomical History and Heritage* 10.1 (Mar. 2007), pp. 1–10.

[2] Wolfgang Demtröder. *Experimentalphysik 4*. Jan. 2017. ISBN: 978-3-662-52883-9. DOI: 10.1007/978-3-662-52884-6.

[3] C. A. Muller and J. H. Oort. "Observation of a Line in the Galactic Radio Spectrum: The Interstellar Hydrogen Line at 1,420 Mc./sec., and an Estimate of Galactic Rotation". In: 168.4270 (Sept. 1951), pp. 357–358. DOI: 10.1038/168357a0.

[4] NRAO/AUI. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. Image courtesy of NRAO/AUI. 2017. URL: https://commons.wikimedia.org/wiki/File:JanskyatAntenna_hi.tif.

[5] J. H. Oort, F. J. Kerr, and G. Westerhout. "The galactic system as a spiral nebula (Council Note)". In: *MNRAS* 118 (Jan. 1958), p. 379. DOI: 10.1093/mnras/118.4.379.

[6] Grote Reber. 1937. URL: https://commons.wikimedia.org/wiki/File:Grote_Antenna_Wheaton.gif.

[7] Grote Reber. "Cosmic Static." In: 100 (Nov. 1944), p. 279. DOI: 10.1086/144668.

[8] Robert A. Rohde. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. 2019. URL: https://commons.wikimedia.org/wiki/File:Solar_spectrum_en.svg.

[9] Jarek Tuszyński. This file is licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license. 2009. URL: https://commons.wikimedia.org/wiki/File:Green_Banks_-_Grote_Reber_Radio_Telescope.jpg.

[10] J. Wilsing and J. Scheiner. "1896, On an Attempt to Detect Electrodynamic Solar Radiation and on the Change in Contact Resistance when Illuminating Two Conductors by Electric Radiation". In: *Classics in Radio Astronomy*. Vol. 10. 1982, p. 147. DOI: 10.1007/978-94-009-7752-5_13.