Individual Research Study Report

Umut Cansu

May 2025

Abstract

In this study, I used a 0.9-meter diameter offset satellite dish to observe the 21 cm hydrogen line emitted by the arms of the Milky Way. The project encompasses a variety of topics within electrical engineering, making it a comprehensive and valuable undertaking. These topics include, but are not limited to, electromagnetic and antenna theory for constructing the antenna, as well as digital signal processing techniques to acquire data from a software-defined radio (SDR) and perform spectral line analysis and data visualization.

Contents

1 Introduction

The observation of the 21 cm hydrogen line holds a special place in the fields of radio astronomy and cosmology. This spectral line, emitted by neutral hydrogen atoms due to a hyperfine transition in their ground state, enables scientists to study regions of space that are otherwise obscured by interstellar dust—such as the center of the Milky Way. Since neutral hydrogen is the most abundant element in the universe, tracking this line provides critical insight into the structure, dynamics, and evolution of galaxies. The 21 cm line was first detected in 1951 by Harold Ewen and Edward Purcell at Harvard University, marking a turning point in astronomical observation. Today, it remains a foundational tool for mapping the large-scale structure of the universe, measuring galactic rotation curves, and exploring the early stages of cosmic formation. In this study, I focus on detecting the 21 cm hydrogen line using a ground-based, low-cost radio telescope setup centered at 1420.4058 MHz. The aim is to capture and analyze radio signals from neutral hydrogen regions in the sky, with an emphasis on practical signal processing methods such as FFT-based spectral analysis and background subtraction and smoothing with Savitzky-Golay filter.

2 Antenna Specifications

2.1 Half-wave Dipole Antenna

The half-wave dipole antenna is one of the simplest and most practical choices for a front-feed antenna in radio telescope setups. It is especially suitable for small-scale or amateur observations. For larger dish antennas and more advanced radio astronomy applications, helical antennas or horn feeds are often preferred due to their higher gain and better directional characteristics. At the hydrogen line frequency of approximately 1420.4058 MHz, the corresponding wavelength is $\lambda = 21.1061$ cm. As the name suggests, the total length of a half-wave dipole antenna is half the wavelength, which comes out to about 10.55 cm. This means each arm of the dipole measures approximately 5.275 cm.[2]

2.2 Dish Antenna

The dish used in this project is a commercial satellite dish antenna designed to operate in the 10.7–12.75 GHz range. Therefore, it is not optimized for hydrogen line observations at 1.42 GHz and introduces some signal loss in terms of dB performance. However, with proper digital signal processing, it is still sufficient to detect the 21 cm hydrogen line. The dish has a theoretical offset angle of 20 degrees, but measurements based on the elevation required to reflect sunlight onto the feed suggest that the actual offset is approximately 22 degrees. The angular resolution of the telescope setup is around 14 degrees, which is relatively coarse and not ideal for high-resolution hydrogen line mapping. Nevertheless, as shown in previous studies [4], even a 60 cm dish can successfully detect the 21 cm hydrogen line.

2.3 Coaxial Cable

I used an LMR-240 coaxial cable to transmit the signal from the dipole antenna to the low-noise amplifier (LNA). This cable is suitable for frequencies ranging from 0.2 GHz to 6 GHz, making it adequate for hydrogen line observations at 1.42 GHz. At 20°C, its maximum attenuation is approximately 33.4 dB per 100 meters at 1.5 GHz.[11] However, since only 3 meters of cable was used in this setup, the actual signal loss is minimal. LMR-240 was chosen for its balance of flexibility, relatively low attenuation, and practicality for short-range amateur radio applications.

2.4 LNA

I used a 1–2000 MHz, 64 dB LNA RF wideband module to amplify the weak signals received by the antenna.[9] This component is essential for boosting the hydrogen line signal before it reaches the SDR for digitization. Its working specifications are listed in Table 1.

Aspect	Specification
Working Voltage	12 VDC
Current	70 mA
Gain	64 dB
Input Signal	≤ -54 dB
Bandwidth	1–2000 MHz
Impedance	50 Ω
Noise Figure	2 dB @ 1.5 GHz
Dimensions	7.1×5 cm
Weight	0.1 kg

Table 1: LNA Specifications

2.5 Receiver

As the receiver, I used the HackRF SDR. It is capable of both receiving and transmitting signals over a wide frequency range, from 1 MHz to 6 GHz. The device operates with 8-bit precision and supports quadrature sampling rates from 2 MSPS to 20 MSPS. It uses direct conversation method, that helps the price but causes local oscillator leakage in the signal. I selected the HackRF for this project due to its broad feature set, open-source support, and excellent compatibility with Python-based tools such as GNU Radio.[8]

3 Signal Processing

3.1 GNU Radio

GNU Radio is a digital signal processing (DSP) framework with a graphical user interface, based on Python and designed to work with SDRs. It includes a wide range of built-in signal processing blocks, such as low-pass filters, FFT blocks, and moving average filters. In my setup, I process a signal that is offset by 500 kHz from the center frequency in order to avoid local oscillator leakage, which is common with the direct conversion method used by many SDRs. I then apply a *Frequency Xlating FIR Filter* block with a cutoff frequency of 1500 kHz to isolate the frequency chunk centered at -500 kHz. After this, I use a low-pass filter with a 500 kHz cutoff to further clean the signal. Next, I compute a 4096-point FFT and take the square of the absolute value to obtain the power spectrum. Typically, I save the data at this point, but I also apply a moving average filter and plot the result to provide real-time visual feedback during the observation process. Since my system isn't calibrated, the results don't correspond to any physical units. Instead, the values are in arbitrary units and will be shown as cts (counts).

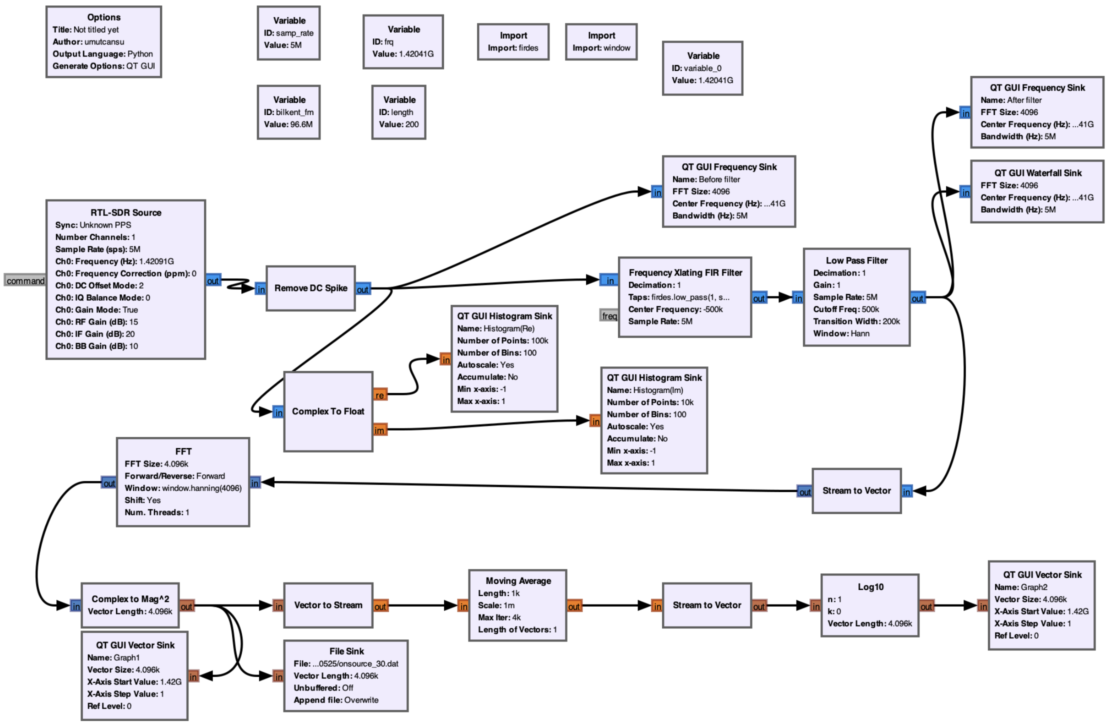

Figure 1: GNU Radio Block Diagram

Here are the graphs from the observation.

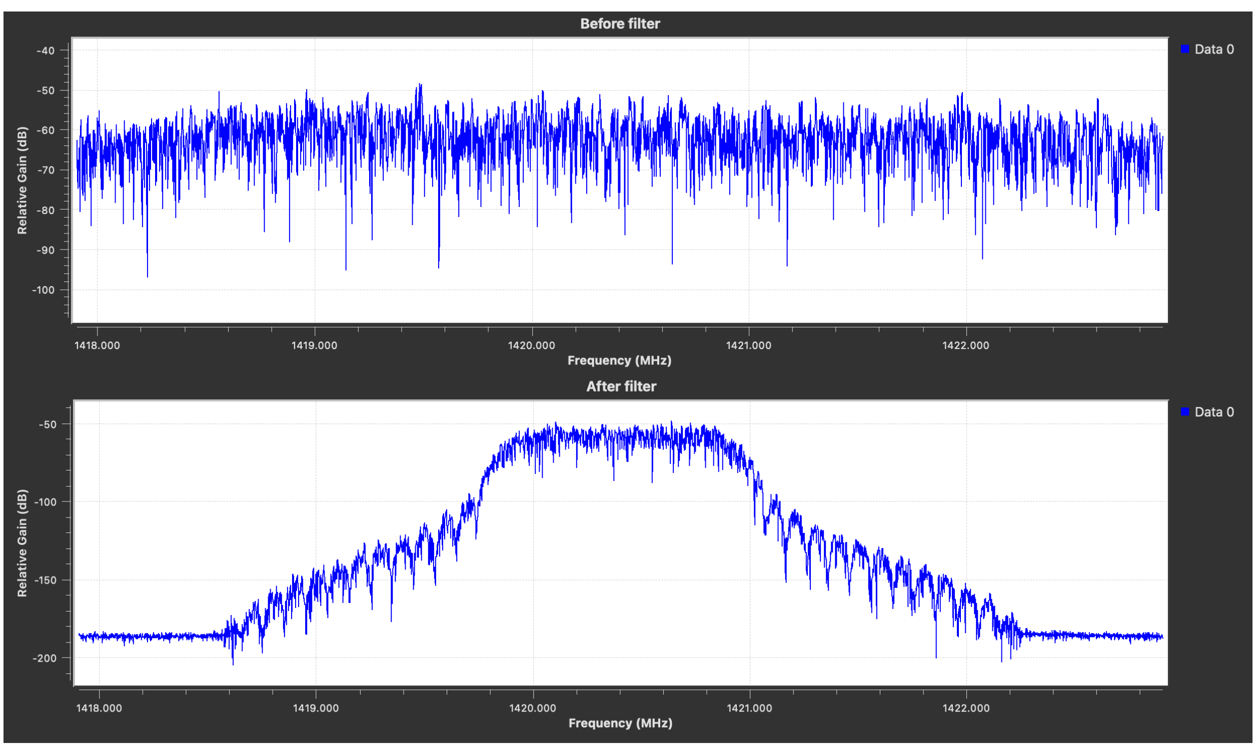

Figure 2: Signal Before and After the Filter

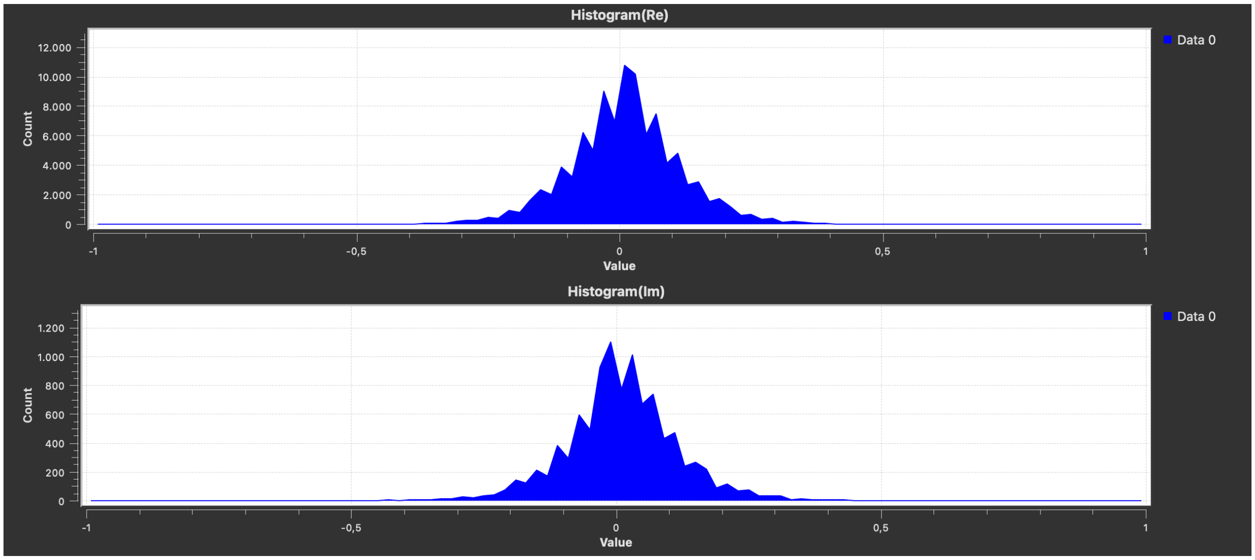

Figure 3: Histogram of the Input Signal

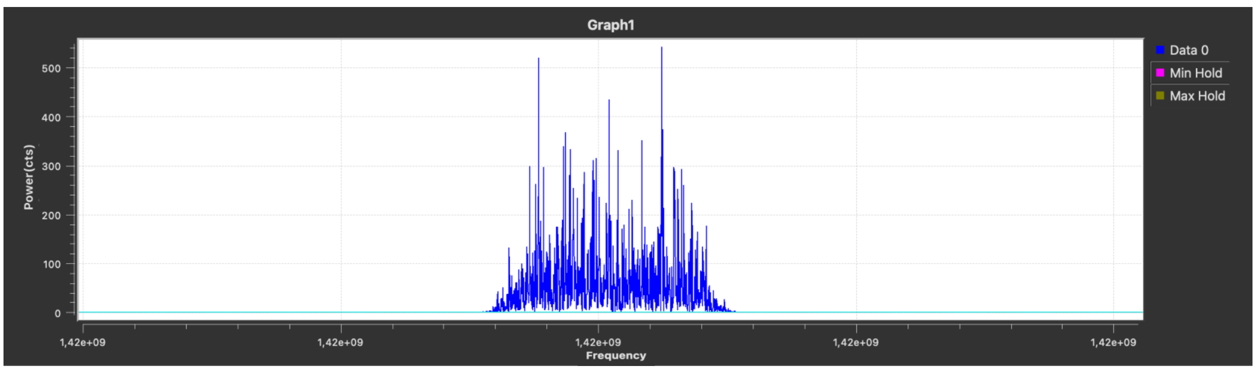

Figure 4: $|FFT|^2$ Result

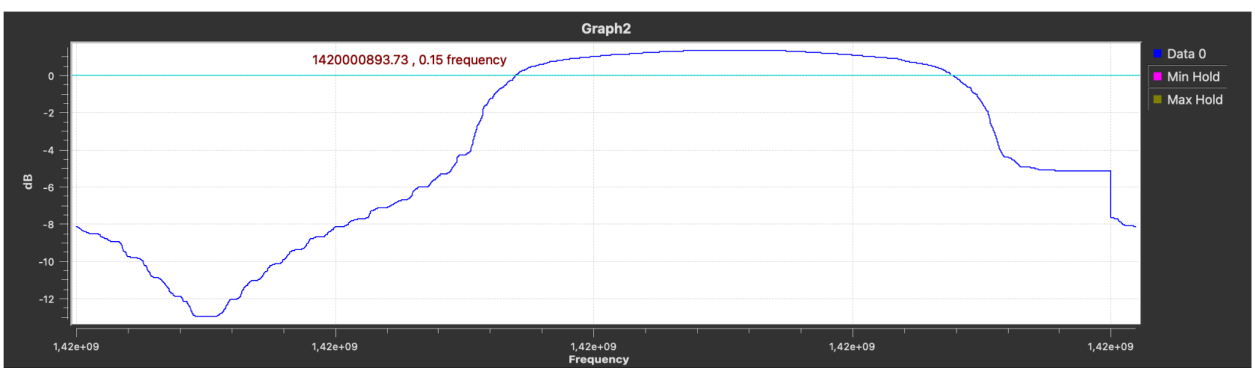

Figure 5: $10log_{10}|FFT|^2$ Result

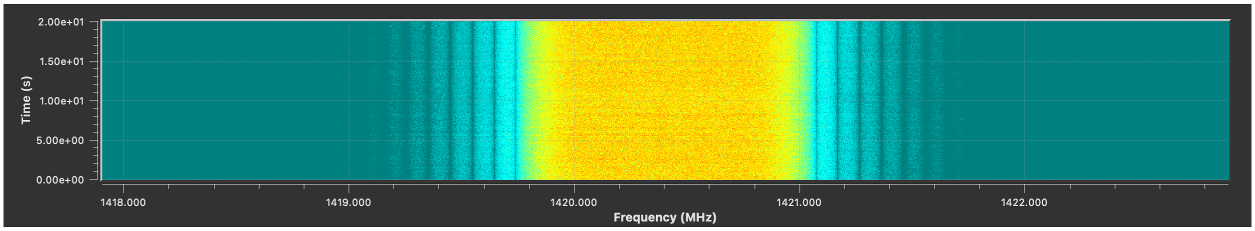

Figure 6: Waterfall Graph

3.2 Python

I used Python to further process the observed data. First, I reshaped the data into $m \times 4096$ matrices, where each matrix represents five minutes of observation. I then computed the non-coherent average for both the on-source and off-source datasets. Next, I applied two background subtraction methods. The first was direct subtraction; however, due to differences in observation times, the power levels of the on-source and off-source data did not match. To compensate for this, I projected the off-source observation vector onto the on-source vector and subtracted the projection component:

$$\hat{S}_a = S_a - \frac{\langle S_b, S_a \rangle}{\|S_b\|_2^2} S_b \tag{1}$$

After background subtraction, I smoothed the resulting signal using a Savitzky–Golay filter with a window length of 31 and a polynomial degree of 3. Finally, I converted the power spectrum to decibels using the formula $10 \log_{10}(\cdot)$. Further implementation details can be found in the appendix.

4 Observation Method and Results

4.1 Observation Method

I used the Stellarium application, which provides comprehensive sky maps and positional data for amateur astronomers. This tool was essential for determining the correct azimuth and elevation angles required for my observations.

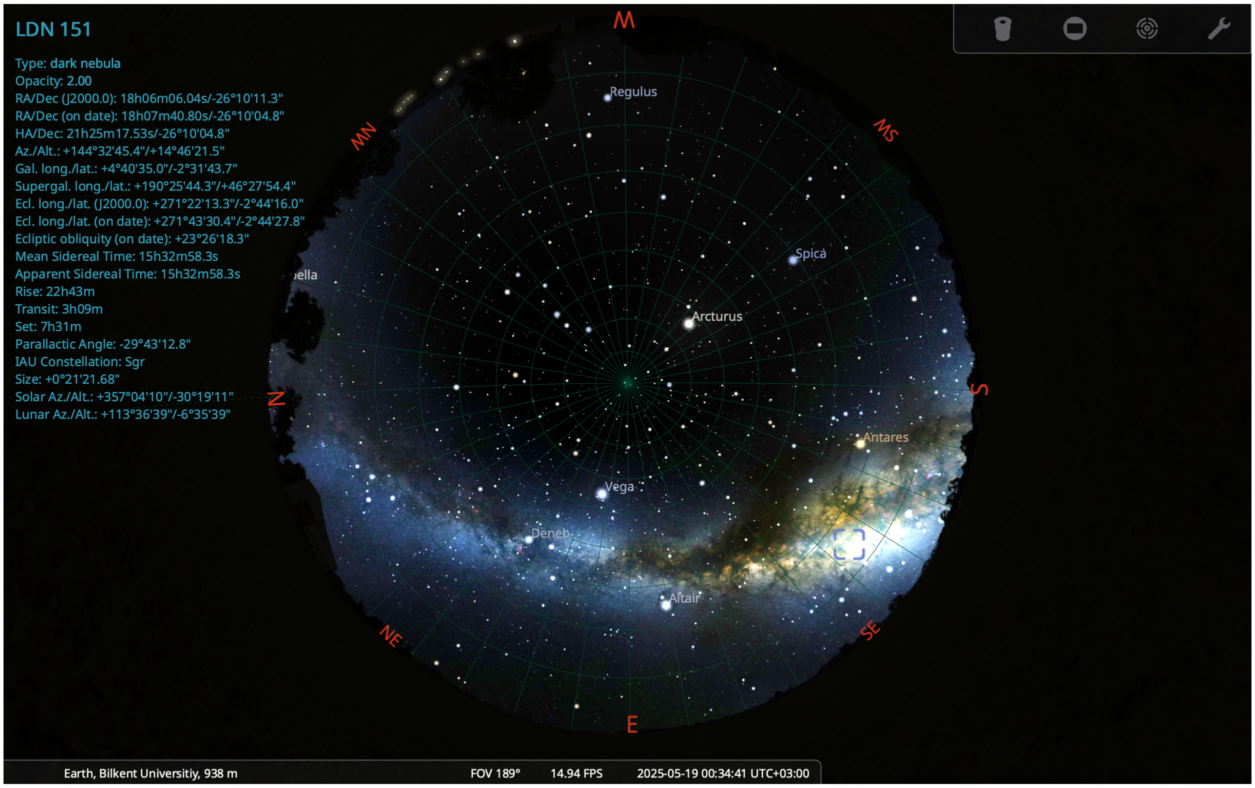

Figure 7: Stellarium App

After identifying one of the Milky Way's arms using Stellarium, I conducted a series of 5-minute observations over the course of one hour. During this period, I manually adjusted the azimuth and elevation of the telescope every five minutes to remain aligned with the target region. Following the on-source observations, I redirected the telescope to a nearby region of blank sky—at the same elevation but a different azimuth—to collect background data for subtraction. Here is my observation time table:

Interval	Start Time	Observation Type
1	06:06	On-source
2	06:11	On-source
3	06:19	On-source
4	06:26	On-source
5	06:32	On-source
6	06:38	On-source
7	06:50	On-source
8	06:56	On-source
9	07:02	On-source
10	07:09	On-source
11	07:17	On-source
12	07:29	On-source
13	07:37	Off-source (background)

Table 2: Observation Schedule: On-source and Background

4.2 Results

Prior to the main one-hour observation, I conducted six shorter test observations—each time identifying and correcting a single issue in the setup (frequency calibration, baseline subtraction, band-pass filtering, etc.). By the seventh session, these incremental improvements yielded a clear detection of the H I 21 cm line.

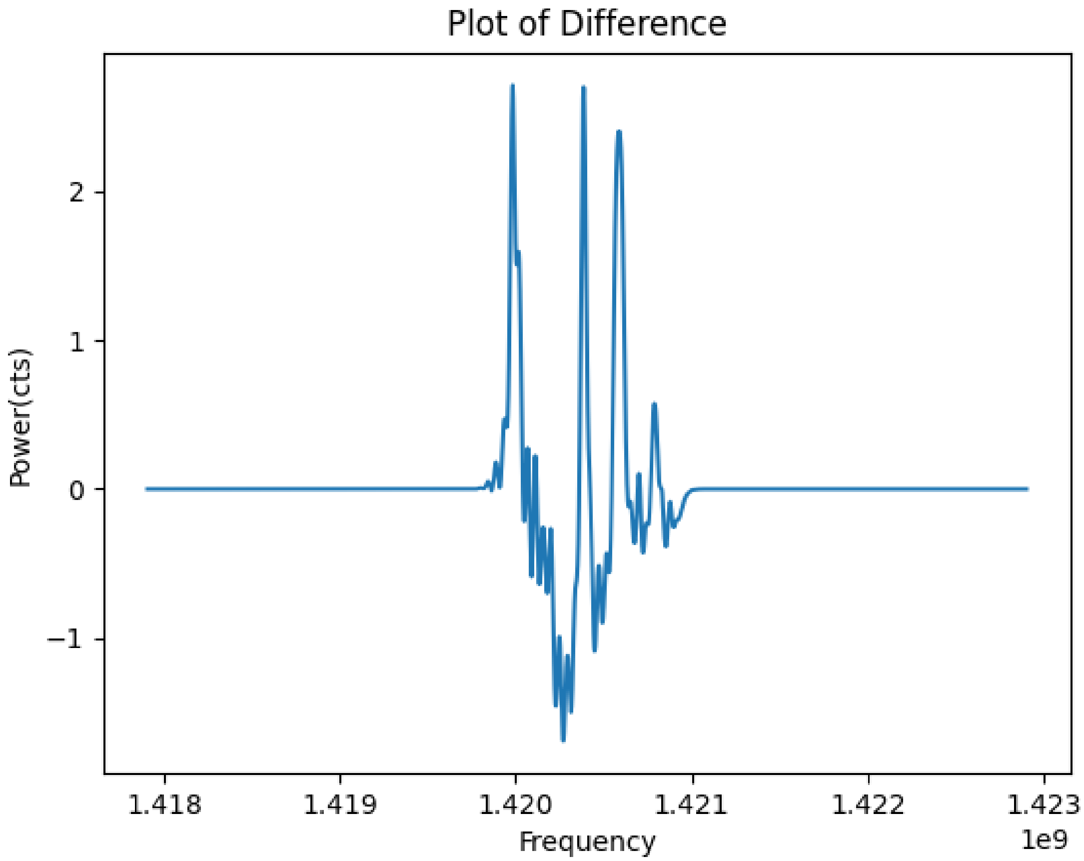

Figure 8: On–off difference spectrum of the H I 21 cm line at Galactic coordinates $(\ell, b) = (50.54°, 6.47°)$. Three positive features appear: the leftmost is a subtraction artifact, while the two rightmost correspond to genuine Galactic emission from the Local and Perseus spiral arms.

The three positive "hills" in Figure 8 arise because on–off subtraction leaves any signal present in the on–source beam but absent (or weaker) in the off–source beam. The leftmost peak is an artifact (a small gain mismatch producing a derivative-shaped residual). The two genuine H I components occur at

$$f_{\text{peak},1} = 1420.3884 \text{ MHz}, \quad f_{\text{peak},2} = 1420.5858 \text{ MHz},$$

which we associate with the Local (Orion) arm and the Perseus arm, respectively.

From each measured f_{peak} we compute the *topocentric* Doppler shift via

$$V_{\text{topo}} = c\,\frac{f_0 - f_{\text{peak}}}{f_0}\,, \qquad f_0 = 1420.4058 \text{ MHz},$$

yielding

$$V_{\text{topo},1} \approx +3.7 \text{ km/s}, \quad V_{\text{topo},2} \approx -38.0 \text{ km/s}.$$

These values still include Earth's spin and orbital motion. To compare with published Galactic surveys, we convert to the Local Standard of Rest (LSR) by applying the observer-to-LSR correction

$$V_{\text{obs}\to\text{LSR}} = -36.637 \text{ km/s},$$

obtained from the NRAO "Radial Velocity Calculator" for the Green Bank Telescope at our observation date, UTC time, and pointing [10]. The resulting LSR velocities are

$$V_{\text{LSR},1} = V_{\text{topo},1} + V_{\text{obs}\to\text{LSR}} \approx -32.9 \text{ km/s}, \quad V_{\text{LSR},2} \approx -74.6 \text{ km/s}.$$

Figure 9 overlays these LSR velocities on the brightness-temperature profile extracted from the Leiden/Argentine/Bonn (LAB) 21 cm survey at the same (ℓ, b) with a 14° effective beam [3]. Both markers fall on the known emission ridges of the Local and Perseus arms, confirming that the two detected features are genuine H I lines.

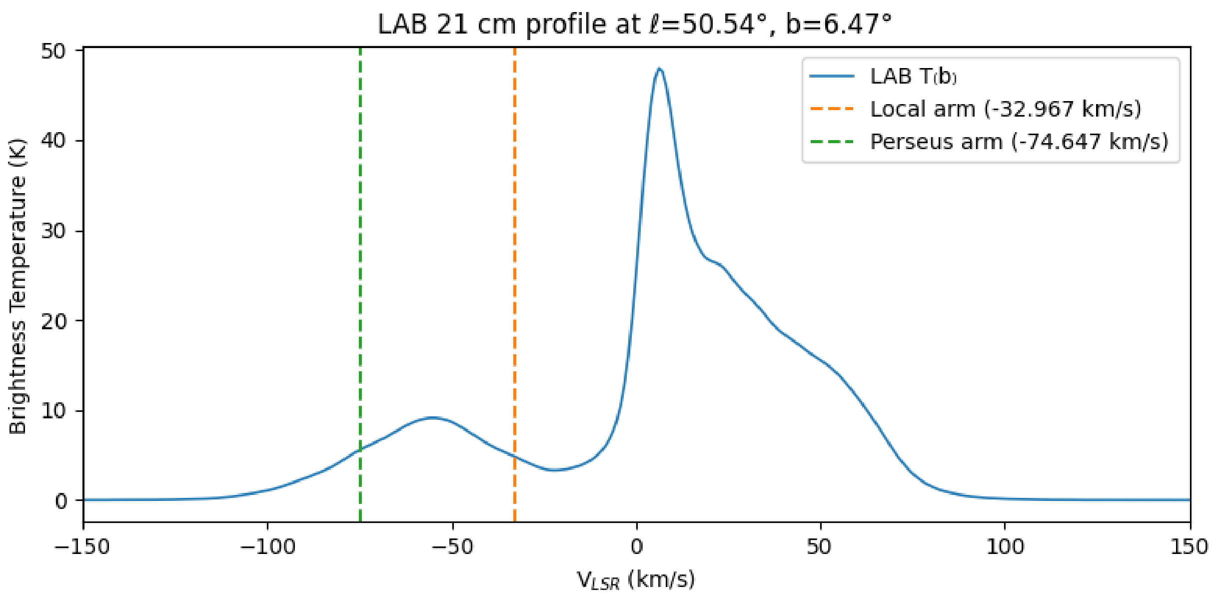

Figure 9: LAB 21 cm brightness-temperature profile at $(\ell, b) = (50.54°, 6.47°)$ with overlaid LSR velocities $V_{\text{LSR},1} = -32.9\,\text{km/s}$ (Local arm) and $V_{\text{LSR},2} = -74.6\,\text{km/s}$ (Perseus arm)

5 Possible Upgrades

There are several aspects of the setup that can be improved. First, although it is possible to detect the hydrogen line with even smaller dish antennas [4, 5, 6, 7], increasing the dish size would improve the signal-to-noise ratio (SNR) and angular resolution. Second, the current feed—a basic dipole antenna—could be replaced with a front-feed horn antenna to enhance signal reception and overall system performance. Another potential issue is the absence of a band-pass filter before the receiver. Without this filtering, radio frequency interference (RFI) may have masked the already weak hydrogen line signals, burying them in noise and making detection more difficult. Additionally, the low-noise amplifier used in this setup has a noise figure of 2 dB at 1.5 GHz, while hydrogen line observations typically require a noise figure below 0.5 dB [1, 5]. This discrepancy may have further reduced system sensitivity. Beyond observational sensitivity, several practical upgrades could improve usability and precision. Incorporating motorized systems—such as azimuthal rotators and linear actuators for dish tilting—would enable more accurate targeting and tracking of celestial sources. Implementing such a system would require feedback mechanisms and possibly custom-designed components to support and dynamically adjust the dish structure.

6 Conclusion

This study aimed to observe the 21 cm hydrogen line from the arms of the Milky Way and was ultimately successful in detecting it. The project provided valuable experience in both theoretical and practical aspects of radio astronomy. It deepened my understanding of signal processing, instrumentation, and the challenges involved in detecting weak astronomical signals using amateur equipment. In addition to technical skills, this project strengthened my ability to design, troubleshoot, and antenna construction and alignment to data acquisition, processing, and interpretation. The successful detection demonstrates that meaningful astronomical observations are possible even with modest resources when combined with careful planning and appropriate digital techniques. This work lays the foundation for future improvements in sensitivity, resolution, and automation.

7 Citations

References

[1] Venkat Ramana Aitha and Mohammad Kawsar Imam. Low noise amplifier for radio telescope at 1.42 ghz. Master's thesis, Halmstad University, School of Information Science, Computer and Electrical Engineering, Halmstad, Sweden. Technical Report IDE0747; `https://www.diva-portal.org/smash/get/diva2:238215/FULLTEXT01.pdf`; accessed: 2025-05-24.

[2] Antenna-Theory.com. Half-wave dipole antenna. `https://www.antenna-theory.com/antennas/halfwave.php`. Accessed: 2025-05-23.

[3] Argelander-Institut für Astronomie. Lab survey hi profile search tool. `https://www.astro.uni-bonn.de/hisurvey/euhou/LABprofile/index.php`. Accessed: 2025-05-25.

[4] AstroPeiler. Hydrogen i. `https://astropeiler.de/wp-content/uploads/2020/01/Hydrogen_1.pdf`. PDF; accessed: 2025-05-24.

[5] AstroPeiler. Hydrogen ii. `https://astropeiler.de/wp-content/uploads/2020/01/Hydrogen_2.pdf`. PDF; accessed: 2025-05-24.

[6] AstroPeiler. Hydrogen iii. `https://astropeiler.de/wp-content/uploads/2020/01/Hydrogen_3.pdf`. PDF; accessed: 2025-05-24.

[7] AstroPeiler. Hydrogen iv. `https://astropeiler.de/wp-content/uploads/2020/01/Hydrogen_4.pdf`. PDF; accessed: 2025-05-24.

[8] Great Scott Gadgets. Hackrf documentation (latest). `https://hackrf.readthedocs.io/en/latest/`. Accessed: 2025-05-24.

[9] Motorobit. 1–2000 mhz 64 db lna rf geniş bant amplifikatör modülü. `https://www.motorobit.com/1-2000mhz-64db-lna-rf-genis-bant-amplifikator-modulu`. Accessed: 2025-05-23.

[10] National Radio Astronomy Observatory. Gbt radial velocity calculator. `https://www.gb.nrao.edu/cgi-bin/radvelcalc.py`. Accessed: 2025-05-25.

[11] RFmarket. Lmr-240 coaxial cable. `https://www.rfmarket.com.tr/urun/lmr240`. Accessed: 2025-05-23.

8 Appendix

8.1 Antenna Photos

Figure 10: Antenna Structure

Figure 11: One of Early Observations

Figure 12: Last Observation 1

Figure 13: Last Observation 2

8.2 Python Code

```
#!/usr/bin/env python3
# -*- coding: utf-8 -*-
"""
Created on Tue May 20 14:15:10 2025

@author: umutcansu
"""

import numpy as np
import scipy.signal as scp
import matplotlib.pyplot as plt
import pandas
from scipy.ndimage import gaussian_filter1d
from astropy.coordinates import SkyCoord
import astropy.units as u
#%%

f_cen = 1420.4058e6
f_samp = 5e6
N = 4096



dir_data = "/Users/umutcansu/Desktop/Radyo_teleskop/observation_data/d200525"
min_length = float("inf")
#%%
for i in range(1,13):
    dir_name = dir_data + f"/onsource_{i*5}.dat"
    data = np.fromfile(dir_name, dtype=np.float32)
    size = len(data)
    if min_length > size:
        min_length = size



#%%
dir_offsource = "/Users/umutcansu/Desktop/Radyo_teleskop/observation_data/d200525/offsource
back_g = np.fromfile(dir_offsource, dtype=np.float32)
size_b = len(back_g) // 4096
back_g = back_g[0:4096*size_b]
back_g = np.reshape(back_g,[-1,4096])
mean_g = np.mean(back_g,axis = 0)


X = np.zeros([12,4096],dtype = np.float32)
for i in range(1,13):
    print(f"Looking into observation {i}/12")
    dir_name = dir_data + f"/onsource_{i*5}.dat"
    data = np.memmap(dir_name, dtype=np.float32,mode ='r')
    data = data[0:4096*size_b]
```

```
    data = np.reshape(data, [-1, 4096])
    mean = np.mean(data,axis=0,dtype=np.float32)
    X[i-1,:] = mean





#%%
data_mean = X[2:6,:]
data_mean = np.mean(data_mean, axis = 0)

smooth_justdata = scp.savgol_filter(data_mean,window_length=31,polyorder=3,mode = 'mirror')
smooth_noise = scp.savgol_filter(mean_g,window_length=31,polyorder=3,mode = 'mirror')
#data_mean = np.mean(X, axis = 0,dtype=np.float32)
smooth_dif = smooth_justdata -smooth_noise
diff = data_mean  - mean_g
freqs = np.linspace(f_cen - f_samp/2, f_cen+f_samp/2, N, endpoint=False)
toggle_mask = (freqs > f_cen-500e3) & (freqs < f_cen+500e3)
mask = (freqs < f_cen-500e3) | (freqs >f_cen+500e3)
coeffs = np.polyfit(freqs[mask], diff[mask],3)
baseline = np.polyval(coeffs, freqs)
coarse_d = diff - baseline
smoothed =scp.savgol_filter(coarse_d,window_length=31,polyorder=3,mode = 'mirror')
smoothed2 =scp.savgol_filter(diff[toggle_mask],window_length=31,polyorder=3,mode = 'mirror
mag = 10*np.log10(smoothed.clip(min = 1e-12))
mag_data = 10*np.log10(smooth_justdata)
mag_backg = 10*np.log10(mean_g)


#%%
#Here i ll substract the min from both mean data and mean back ground noise

s_a = smooth_justdata
s_b = smooth_noise
transpose = np.matmul(s_a , np.transpose(s_b))
mult = transpose*s_b
s_ahat = (mult/(np.linalg.vector_norm(s_b)*np.linalg.vector_norm(s_b)))
s_be = s_a - s_ahat
s_be = scp.savgol_filter(s_be,window_length=31,polyorder=3,mode = 'mirror')


plt.figure(7).clf()
plt.subplot(4,1,1)
plt.plot(freqs,s_be)
plt.title("S_be")
plt.subplot(4,1,2)
plt.plot(freqs,s_a)
plt.title("S_a")
plt.subplot(4,1,3)
plt.plot(freqs,s_ahat)
plt.title("S_ahat")
plt.subplot(4,1,4)
plt.plot(freqs,s_b)
```

```
plt.title("S_b")

plt.figure(8).clf()
plt.plot(freqs,s_be)
plt.title("Plot of Difference")
plt.xlabel("Frequency")

#%%



data_slice = data_mean[toggle_mask]
data_slice = scp.savgol_filter(data_slice,window_length=31,polyorder=3,mode = 'mirror')
back_slice = mean_g[toggle_mask]
diff_slice = data_slice - back_slice

plt.figure(1).clf()
plt.subplot(3,1,1)
plt.plot(freqs,smoothed)
plt.title("Baseline subs, smoothing after diff")
plt.subplot(3,1,2)
plt.plot(freqs[toggle_mask],smoothed2)
plt.title("NO Baseline subs, smoothing after diff")
plt.subplot(3,1,3)
plt.plot(freqs,smooth_dif)
plt.title("NO Baseline subs, smoothing before diff")



plt.figure(2).clf()
plt.plot(freqs, mag_data)
plt.axvline(f_cen)
plt.axvline(f_cen - 500e3)
plt.axvline(f_cen + 500e3)
plt.title("onsource")

plt.figure(3).clf()
plt.plot(freqs, mean_g)
plt.title("back ground")

plt.figure(4).clf()
plt.subplot(311)
plt.plot(freqs[toggle_mask],10*np.log10(data_slice))
plt.axvline(f_cen)
plt.axvline(f_cen - 100e3)
plt.axvline(f_cen + 100e3)
plt.subplot(312)
plt.plot(freqs[toggle_mask],back_slice)
plt.axvline(f_cen)
plt.subplot(313)
plt.plot(freqs[toggle_mask],diff_slice)
plt.axvline(f_cen)

plt.figure(5).clf()
```

```
plt.plot(freqs, baseline)

#%%


# freq_hz     : 1D array of frequency bins [Hz]
# spec_diff   :  on off  difference **linear** power spectrum

f0 = 1_420_405_800.0    # rest freq of H I [Hz]

# 1) local arm  component: within  100  kHz of f0
mask1 = np.abs(freqs - f0) < 100e3
f_peak1 = freqs[mask1][np.argmax(s_be[mask1])]

# 2) distant arm  component: between +100 kHz and +300 kHz
mask2 = (freqs - f0 > 100e3) & (freqs - f0 < 300e3)
f_peak2 = freqs[mask2][np.argmax(s_be[mask2])]

print(f"Local arm peak:   {f_peak1/1e6:.6f} MHz")
print(f"Distant arm peak: {f_peak2/1e6:.6f} MHz")

c     = 299_792.458    # km/s
v1    =  c*(f0 - f_peak1)/f0
v2    =  c*(f0 - f_peak2)/f0
print(f"V_local     {v1:.1f} km/s,  V_perseus     {v2:.1f} km/s")


#%%



sky = SkyCoord("19h00m20.02s", "18d36m03.6s", frame="icrs")

# convert
l = sky.galactic.l.deg
b = sky.galactic.b.deg
print(f"Galactic (l, b) = ({l:.2f}, {b:.2f})")



V0 = 220.0                          # km/s
l_rad = np.deg2rad(l)
b_rad = np.deg2rad(b)

v_pred = V0 * np.sin(l_rad) * np.cos(b_rad)
print(f"Expected V_LSR     {v_pred:.1f} km/s")

#%%


import numpy as np
import matplotlib.pyplot as plt
```

```python
vel, Tb = np.loadtxt('spectrum.txt', skiprows=4, usecols=(0,1), unpack=True)



v_local   = -32.967 # km/s
v_perseus = -74.647 # km/s

# Plot
plt.figure(figsize=(8,4))
plt.plot(vel, Tb, lw=1.2, label='LAB T b ')
plt.axvline(v_local,   ls='--', color='C1', label=f'Local arm ({v_local} km/s)')
plt.axvline(v_perseus, ls='--', color='C2', label=f'Perseus arm ({v_perseus} km/s)')

plt.xlim(-150, 150)
plt.xlabel('V$_{LSR}$ (km/s)')
plt.ylabel('Brightness Temperature (K)')
plt.title('LAB 21 cm profile at l=50.54, b=6.47')
plt.legend()
plt.tight_layout()
plt.show()
```