The Utility of the 21-cm Hydrogen Emission Line

Cleo Lepart

PHYS 43 (Spring 2021)

1 Abstract

We are all familiar with optical telescopes but the truth is that most of the stunning astronomy photos we see of nebulae and galaxies were created with the help of radio telescopes. They may be composited with optical photos, but processed radio telescope "images" are how we can see gases and other elements in space. The most abundant of these gases is neutral hydrogen. It makes up vast clouds of the galaxy and its mapping led to a revolution in astronomy, as it was possible to visualize our spiral galaxy and interstellar gas for the first time [1]. All this was achieved with a horn telescope that an amateur can basically build themselves [2]. The 21-cm hydrogen emission line that we detect is due to the hyperfine structure of atomic energy levels [3]. As this semester's course deals with atomic and quantum phenomena, it was doubly relevant to take on the project of building a pyramidal horn that could receive signals from the hydrogen line of our galaxy, calculate its speed, and gain hands-on competence with electronic hardware and signal processing software.

2 Cause

The hyperfine structure splitting that produces 21-cm electro-magnetic radiation was first observed in 1881 by Albert A. Michelson [4]. However, it could only be explained in terms of quantum mechanics in 1924 when Wolfgang Pauli proposed the existence of a small nuclear magnetic moment [4].

Hyperfine structure is defined as small shifts in degenerate (non-ground states) energy levels and the resulting splits of those energy levels of atoms (or molecules or ions) as the nucleus and electron clouds interact [5]. In its lowest energy configuration, the state of neutral hydrogen consists of an electron bound to a proton. Both the electron and the proton have intrinsic magnetic dipole moments which we describe as their spin. The nucleus-electron cloud interaction results in a slight increase in energy when the spins are parallel, and a decrease when anti-parallel. These states are only ever parallel or anti-parallel, as a result of the quantum mechanical discretization of the total angular momentum of the system. When the spins are parallel, the magnetic dipole moments are anti-parallel (this is because the electron and proton have opposite charge); this means that this particular configuration has higher energy [3].

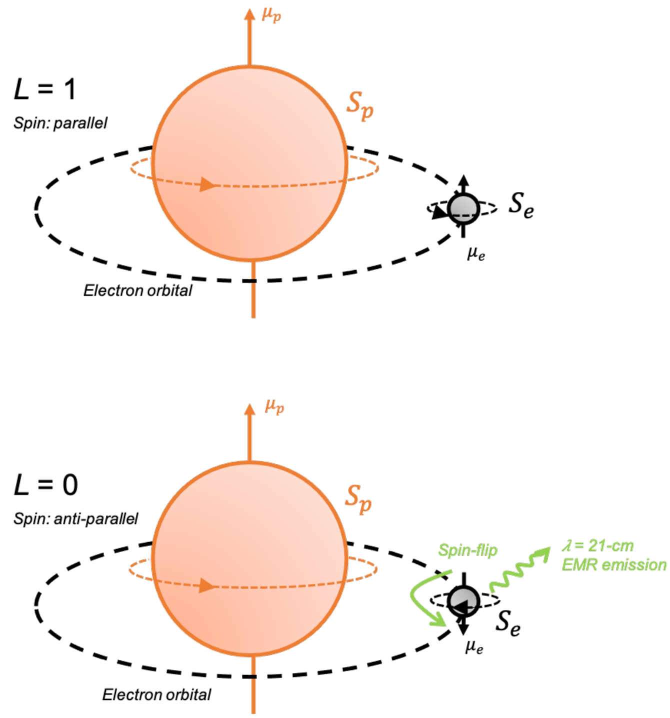

Figure 1. Magnetic dipole moment of the nucleus and electron in neutral hydrogen. (Figure created by student.)

Because of the proton and the electron's magnetic moment, every energy level of the hydrogen atom actually consists of two levels, very close together. Each magnetic moment creates a magnetic field, which means that the energy of the H-atom is slightly different, depending on the relative orientation of the electron and proton moments [6]. For simplicity, imagine a point on the axis of μ, the proton's magnetic moment, on the magnetic field. Then the magnetic field at a distance of r from μ, is

$$B = \frac{\mu_0 \mu}{2\pi r^3} \tag{1}$$

where μ_0 is the permeability of free space.
From here we can estimate the magnetic field at a distance R_p from the proton, given that μ is roughly equal to the nuclear magneton [6]:

$$\mu_N = \frac{e\hbar}{2m_N} \tag{2}$$

Thus,

$$B_p = \frac{\mu_0 g_p \mu_N}{2\pi R_p^3} \tag{3}$$

Note the presence of the dimensionless proton g-factor, g_p, which is 5.585694713 [7].

Now we can calculate the energy difference, this is the difference in energy potential depending on whether the electron-proton spins are anti-parallel or parallel.

$$U = -P \langle \mu_e \cdot B_p \rangle \tag{4}$$

The potential energy equation, without considering the spin-coupling term, includes the probability function of finding the electron within the radius of the proton, $P = \frac{4R_p^3}{3a_0^3}$ where $a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}$ is the Bohr radius of the electron [8].

$$U = \frac{4}{3}\frac{\mu_0 g_p \mu_N g_e \mu_B}{2\pi a_0^3} \tag{5}$$

Where we have an electron g-factor of -2.002319 and μ_B is the Bohr magneton, identical to the nuclear magneton except that the mass is that of an electron and not a proton, $\mu_e = \frac{e\hbar}{2m_e}$. In order to complete our description of the potential energy difference, we need to multiply it by the coupled (dotted) spins of the electron, S_e, and the proton, S_p, over reduced Planck's constant squared,

$$= \frac{4}{3}\frac{\mu_0 g_p \mu_N g_e \mu_B}{2\pi a_0^3}\frac{S_e \cdot S_p}{\hbar^2} \tag{6}$$

The way to determine this spin-spin coupling term is to take the dot product of the total angular momentum, L, with itself and rearrange,

$$S_e \cdot S_p = \frac{1}{2}(L \cdot L - S_e \cdot S_e - S_p \cdot S_p) \tag{7}$$

Now we need to apply spin operators to the wave function, to obtain the eigenvalues.

$$S_e \cdot S_p|\psi\rangle = \frac{\hbar^2}{2}(L(L+1) - S_e(S_e+1) - S_p(S_p+1)|\psi\rangle \tag{8}$$

As we see in Figures 1 and 2, total angular momentum can take on values of either 0 or 1, corresponding to parallel or anti-parallel states. While electron spin and proton spin both equal $\frac{1}{2}$. For the L = 0 anti-parallel state, we have

$$S_e \cdot S_p|\psi\rangle = -\frac{3\hbar^2}{4}|\psi\rangle \tag{9}$$

And for the L = 1 parallel state, we have

$$S_e \cdot S_p|\psi\rangle = -\frac{1\hbar^2}{4}|\psi\rangle \tag{10}$$

We can now take the difference of (9) and (10) and substitute it into (6), yielding

$$= \frac{4}{3}\frac{\mu_0 g_p \mu_N g_e \mu_B}{2\pi a_0^3}\left(\frac{1}{4} - \left(-\frac{3}{4}\right)\right) \tag{11}$$

Using the known values of our constants, we can determine the potential energy difference

$$\Delta U = 9.43 \times 10^{-27} J = 5.88 \times 10^{-6} eV. \tag{12}$$

This corresponds to a frequency of

$$f = \frac{\Delta U}{h} = 1.42 GHz, \tag{13}$$

Which when divided into c, the speed of light, gives us a wavelength of

$$\lambda = \frac{c}{f} = 0.21 m = 21 cm. \tag{14}$$

This has shown that the 1s state of hydrogen is really two closely spaced levels, as described in the figure below.

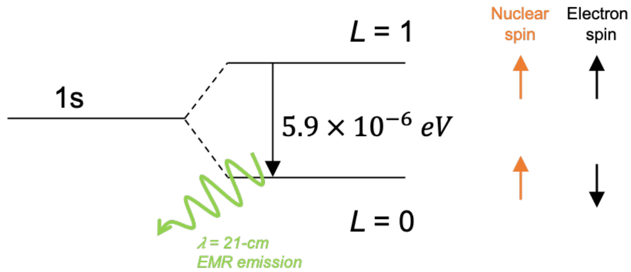

Figure 2. Hyperfine structure of hydrogen. (Figure created by student.)

This idea of spin angular momentum was actually first proposed to explain hyperfine splitting in atomic spectra in a 1925 publication by George Uhlenbeck and Samuel Goudsmit [4]. Though this explanation was provided with a rigorous theoretical foundation by Paul Dirac in 1928 in the Dirac equation for the wavefunction of the electron [4].
This transition is normally termed "highly forbidden", in that it violates the orbital angular momentum selection rule in which $\Delta l = \pm 1$. In this hyperfine splitting situation we see a change in the spin angular momentum, $\Delta s = \pm 1$, but none in the orbital motion of the electron $\Delta l = 0$ [9]. "Highly forbidden" does not mean that this transition cannot occur, simply that it occurs at an extremely small transition rate of $2.9 \times 10^{-15} s^{-1}$. Between this and the mean lifetime of the excited degenerate state of around 10 million years [6], extremely large amounts of hydrogen are necessary to see emission. Fortunately large amounts are currently very abundant in our universe, making up 74% of the baryonic matter present [10]. Within our galaxy, hydrogen is commonly observed in astronomical settings such as gaseous clouds [11].

While on the theory side, spin angular momentum can explain the hyperfine splitting that we see as radiation from hydrogen, the actual 21-cm line was first detected as late as 1951 by Harold Ewen and Edward Purcell at Harvard University [1].
This radiation was first detected in the 1930s: astronomer Karl Jansky noted a background radio 'hiss', varying with the earth's rotation and appearing to be extraterrestrial in origin [12]. Initially suggestions were that this was solar in origin. After these were published in 1940, Jan Oort noted that if these were emission lines in the radio part of the spectrum, significant advances could be made in astronomy. He mentioned this to Hendrik van de Hulst who was the first in 1944, to predict that neutral hydrogen could produce radiation at a frequency of 1420 MHz due to the two closely spaced energy levels in the ground state of the hydrogen atom [11]. Once Ewen and Purcell detected the signal, van de Hulst, along with Oort and Muller, began making the first maps of the neutral hydrogen in the Galaxy, and for the first time revealed the spiral structure of the Milky Way [1].

3 Radio Telescope Build

Now we know what we want to detect and why, we can build a radio telescope fitted to this purpose. While neutral hydrogen radiates at 1.42 Ghz, maximum red and blue shifts require optimization for the frequency range of 1420.4 ± 2 MHz. The non-electronic components (waveguide and horn) must efficiently direct desired frequencies toward the antenna. The waveguide itself (an open waveguide) is not an effective energy collector as there is an impedance mismatch at the opening. Therefore the trumpet- or pyramidal-shaped guide, known as the horn, improves the efficacy, coupling the impedance between the waveguide and that of free space [13].

3.1 Waveguide

A vital component of the system is the waveguide, which is a hollow metal pipe used to carry radio waves. The direction of the antenna determines the polarization direction of the wave. Here, the electrical field is polarized parallel to the antenna; the intensity is zero on the surface of the metal waveguide and the maximum intensity is reached inside the waveguide. The form of the propagation mode is characterized by a 'waveguide wavelength'.

$$\lambda_G = \frac{\lambda}{\sqrt{1-(\frac{\lambda}{2a})^2}} = 27.0cm \tag{1}$$

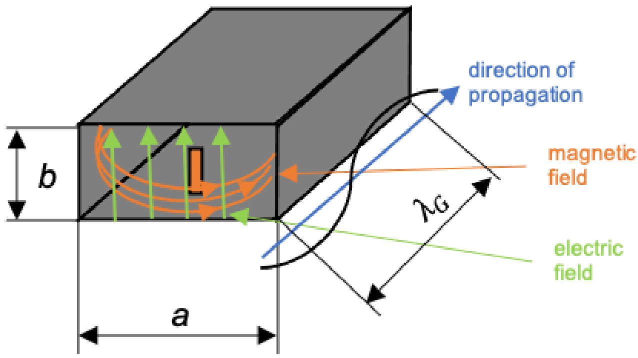

Figure 3. Electro-magnetic fields within the waveguide. (Figure created by student.)

The maximum wavelength that can propagate in the waveguide is known as the cutoff wavelength. It is defined in the direction of the second longest side a where $\lambda_{cutoff} = 2a = 33.4cm$. (Side b is required to be larger than a quarter wavelength $b > \frac{\lambda}{4}$.) This is equivalent to saying the waveguide acts as a high pass filter. Conveniently enough, a readily available rectangle 5-liter paint can with dimensions 16.7 cm x 10.5 cm x 23 cm can act as a functioning and practical waveguide.

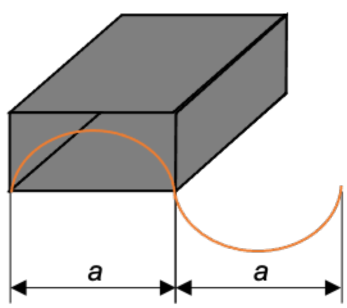

Figure 4. Horizontal waveguide length showing cutoff wavelength. (Figure created by student.)

As we want maximum signal within the antenna, it can be made from a length of copper wire ($\sim$ 14 gauge) cut to a quarter wavelength

$$l = \frac{\lambda}{4} = \frac{21.0cm}{4} = 5.25cm \tag{2}$$

And as the signal should be zero or dead at the walls, the positioning of the antenna can be determined from the guide wavelength.

$$d = \frac{\lambda_G}{4} = 6.75cm \tag{3}$$

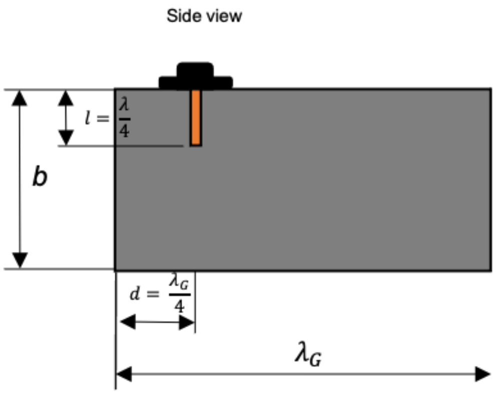

Figure 5. Dimensions and position of antenna as a function of λ and λ_G. (Figure created by student.)

3.2 Horn

The horn is used to direct the waves in a beam, and therefore is also known as a horn antenna. Most commonly made of sheet metal, any conductive surface will do as the desire is to funnel the radiation down to the highly conductive copper antenna, placed less than a quarter-wave above the bottom edge. For ease of use and transport, price and availability, aluminum-coated foam (polyisocyanurate) insulation boards are often employed. Horns are widely used as antennas at these ultra-high radio and microwave frequencies, above 300 MHz. The main concern with any antenna is gain, which is a measure of the ability of dual-port amplifier to increase the power or amplitude of an input sign to an output port. While the best gain is often prohibited by size, and in this case, by the fact that the horn antenna will have to connect with the rectangular wave guide, I used an online application at https://hornantennacalculator.blogspot.com/p/calculator.html to determine that our antenna gain is about 18 dB, which is respectable among amateur radio telescope builders [14]. This is fortunate as I was hampered by the need to construct inside and operate outside, therefore the horn had to pass through my largest doorway placing restrictions on width.

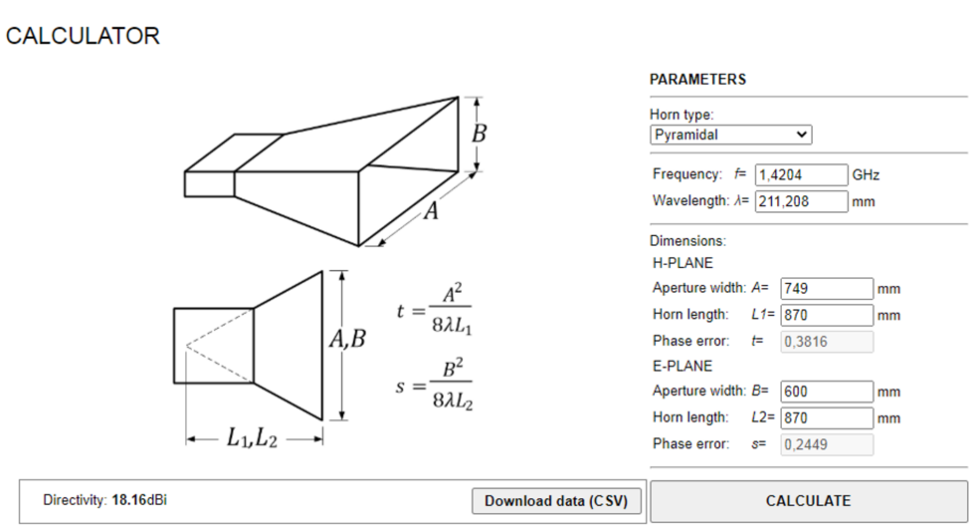

Figure 6. Example of results from https://hornantennacalculator.blogspot.com/p/calculator.html when wavelength and opening widths are used as parameters. Note that the horn's widest opening will be square (870 mm x 870 mm).

3.3 Electronics and Software

Once the antenna has been coupled to the electronic components via an SMA connecter, the signal received is very weak, practically indistinguishable from electronic noise. The aim is high gain in the electronics, while keeping the noise as low as possible in order to maximize the signal-to-noise ratio for a final gain of about 50 dB. The first component will then be a low noise amplifier (LNA) with a noise figure < 1 dB. This low-noise, slightly amplified signal can be passed through a general wideband amplifier. After this amplification stage, it is necessary to insert a band-pass filter centered at 1420 MHz in order to select only the portion of the spectrum that interests us. The actual receiver will be a software defined radio (SDR) module connected to the computer's USB port. (SDR refers to a radio communication system where components that have been traditionally implemented in hardware (e.g. mixers, filters, amplifiers, modulators/demodulators, detectors, etc.) are instead implemented by means of software on a personal computer or other system.) Airspy is an "all-in-one" SDR which is highly compatible with radio signal processing software [2].
So the essential electronic components needed are: LNA amplifier, pre-filtered at 1420MHz, with gain between 30 - 40 dB; wide band amplifier, with gain between 10 - 20 dB; band pass filter, with a passband of about 50MHz, centered at 1420MHz; Airspy R2 SDR receiver.

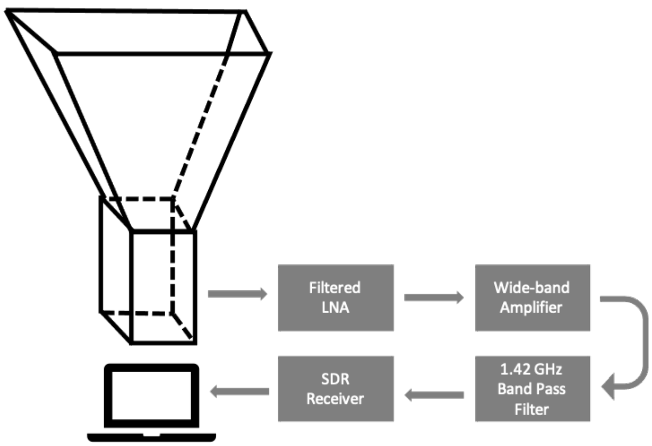

Figure 7. Schematic of the electronic components making the connection from antenna to computer system. (Figure created by student.)

To process the signal from the Airspy R2 SDR Receiver dongle, I installed a user-friendly open source application, Gqrx, that provides signal processing for implementing radio software projects on Unix [15].

Figure 8. Construction completed, electronics in place, ready for mounting outside.

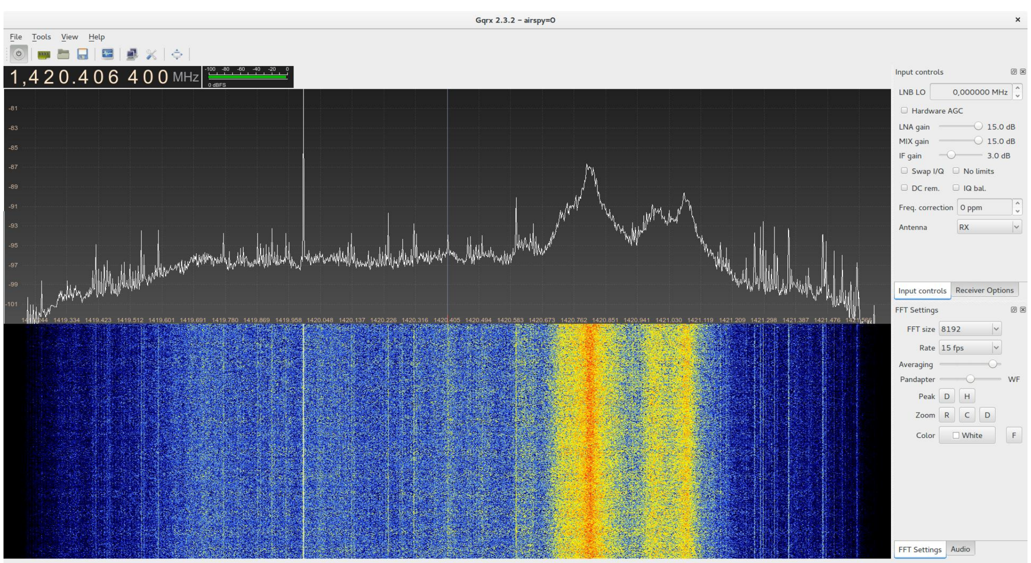

Figure 9. Gqrx signal readout centered at 1420 MHz.

4 Observations from the Radio Telescope

It is now possible to collect data from which there is much information to be extracted, such as this time averaged plot of the power received [16].

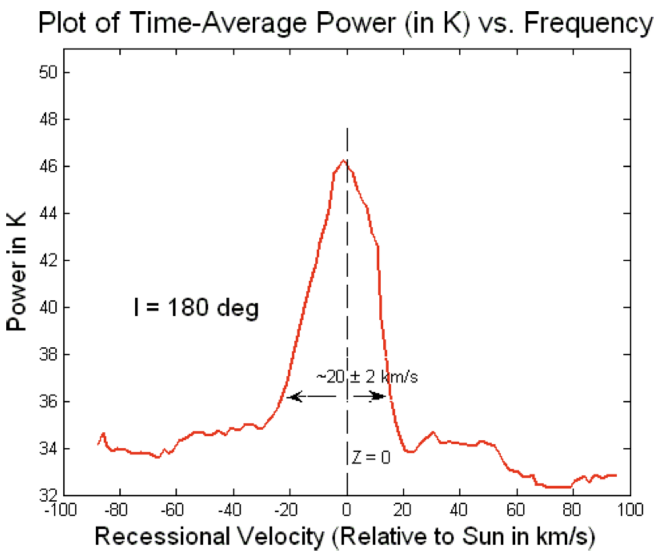

Figure 10. 21-cm spectrum observed at $l = 180°$, with velocity dispersion. (credit Liu, MIT, 2008)

Purcell used this set-up to measure the velocity of clouds of gas, abundant in the spiral arms of the Milky Way Galaxy. Mapping the locations of neutral hydrogen clouds, he found that the gas clouds tend to be distributed in clumps. When these clumps are mapped as a function of galactic longitude and distance (assuming a rotation curve for the galaxy), it is apparent that they lie along discrete spiral arms [17].
To calculate the relative speed between the observer (the radio telescope) and the hydrogen cloud, we use the relativistic Doppler shift formula [16], with $\beta = \frac{v_{rec}}{c}$, and obtain the recessional velocity, v_{rec}, by solving the equation for β.

$$v_{obs} = v_{source}\sqrt{\frac{1-\beta}{1+\beta}} \tag{1}$$

To find the rotational velocity of the source, we derive an equation from the relation of the orbital motion of local standard of rest, the galactic longitude (orientation of the telescope) l, and the Doppler-shifted recessional velocity found previously [16].

$$v_{rot} = v_{rec} + v_{LSR}sin(l) \tag{2}$$

This rotational velocity is fundamental to obtaining a radial distance with which to compile a map [17].

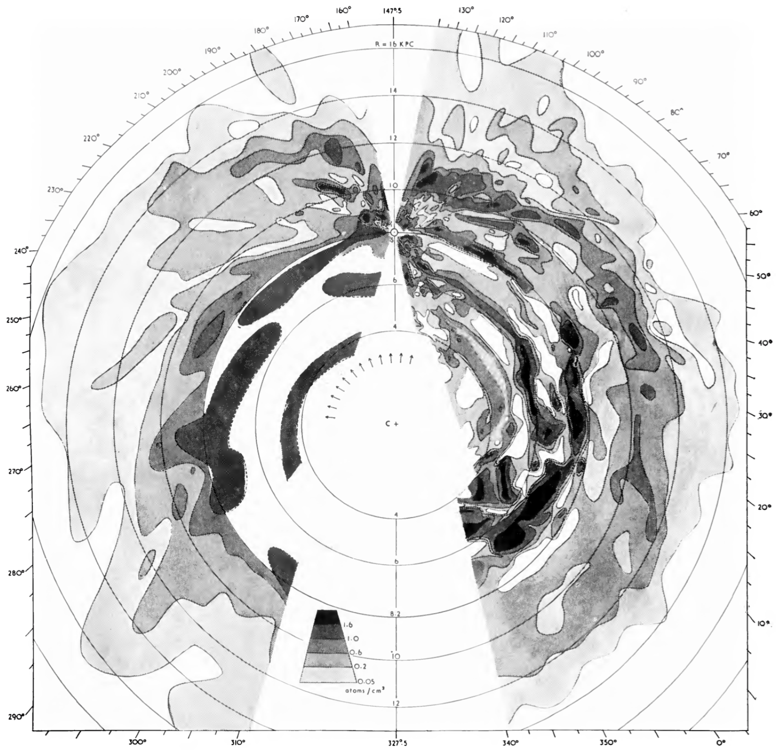

Figure 11. Jan Oort's mapping of the distribution of hydrogen in the Galactic System. (credit Oort, Keer, Westerhout, M.N.R.A.S., 1958)

This method of mapping gives rises to the same rotational galaxy curve that led Fred Hurt and Vera Rubin to propose that dark matter is present in our galaxy, as seen below [16].

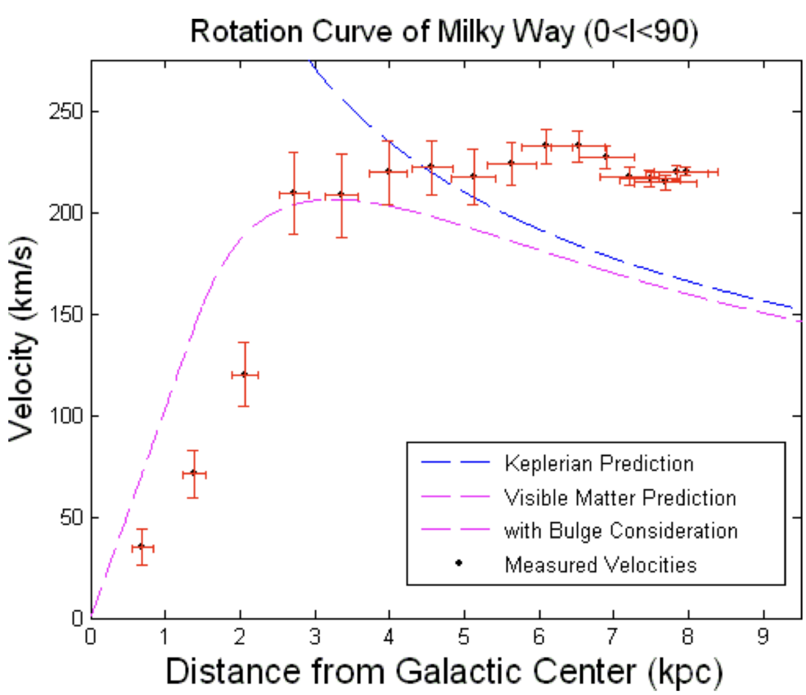

Figure 12. Rotation curve of the Milky Way. (credit Liu, MIT, 2008)

5 Conclusion

To understand why the galaxy is suffused with 21-cm wavelength radiation, we looked at the quantized energy levels of the most abundant element in the Universe. The details of the hyperfine structure of the 1s ground state of hydrogen explain the presence of emission lines at 21 cm or 1.42GHz. This requires understanding the interaction of the magnetic dipole moments of the proton and electron in the hydrogen atom. This simple quantum effect continues to provide radio astronomers, professional and amateur, and cosmologists with extraordinary amounts of information.

For the amateur, the self-built 21-cm radio telescope allows one to investigate the structure and rotation of the galaxy. All this is possible thanks to the availability of cheap LNAs and filters and the development of SDR technology for data acquisition and post-processing. With this equipment, our position in the galaxy and the rotational velocity of any point in our galaxy can be calculated.

This project is phenomenal in that it takes the builder from the exploration of the quantum realm of spins, hyperfine splitting and magnetic dipole moments, through rudimentary electronic amplification and signal processing, (very) small-scale construction and engineering of radio telescopes and then on to probing the masses of hydrogen clouds that swirl around our galaxy. This exploration shows how advanced and valuable amateur radio astronomy can be.

References

[1] K.D. Stephan. "How Ewen and Purcell discovered the 21-cm interstellar hydrogen line". In: *IEEE Antennas and Propagation Magazine* 41.1 (1999), pp. 7–17. DOI: 10.1109/74.755020.

[2] David Schneider. *Track the Movement of the Milky Way With This DIY Radio Telescope*. URL: https://spectrum.ieee.org/geek-life/hands-on/track-the-movement-of-the-milky-way-with-this-diy-radio-telescope. (accessed: 17 May 2021).

[3] Peter Gallagher. *PY 3004 - Atomic and Molecular Spectroscopy, Lecture 7*. [Online; accessed 17-May-2021]. 2008. URL: https://www.tcd.ie/Physics/people/Peter.Gallagher/lectures/PY3004/PY3004_lecture7_notes.pdf.

[4] Wikipedia contributors. *Hyperfine structure — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-May-2021]. 2021. URL: https://en.wikipedia.org/wiki/Hyperfine_structure.

[5] Wikipedia contributors. *Degenerate energy levels — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-May-2021]. 2021. URL: https://en.wikipedia.org/wiki/Degenerate_energy_levels.

[6] Crystal Moorman. "Spin Flip Transition of Hydrogen in Astrophysics". In: *Physics Department, Drexel University* (2010). URL: http://www.physics.drexel.edu/~bob/TermPapers/Moorman_QuantumPaper.pdf.

[7] Wikipedia contributors. *Proton magnetic moment — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-May-2021]. 2021. URL: https://en.wikipedia.org/wiki/Proton_magnetic_moment.

[8] N. Sirica and R. Van Wesep. *UNIT 4: Fine and hyperfine structure of the hydrogen atom*. URL: http://aesop.phys.utk.edu/ph522/ch4.pdf. (accessed: 17 May 2021).

[9] D.J. Griffiths. "Introduction to Quantum Mechanics, 3rd ed." In: Cambridge, 2018. Chap. 11.3.3.

[10] Contributing authors. "Astrophysics, Enhanced ed." In: OpenStax, 2018, p. 31.

[11] Wikipedia contributors. *Hydrogen line — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-May-2021]. 2021. URL: https://en.wikipedia.org/wiki/Hydrogen_line.

[12] Richard G. Strom. "How was the Atomic HI ($\lambda = 21cmline) inSpaceDiscovered$". In: International Journal of Modern Physics: Conference Series 23 (2013), pp. 472–477. DOI: 10.1142/S2010194513011781.

[13] *Horn Antenna for the 21cm Neutral-Hydrogen Line*. 2020. URL: https://physicsopenlab.org/2020/07/20/horn-antenna-for-the-21cm-neutral-hydrogen-line/.

[14] Wikipedia contributors. *Gain (electronics) — Wikipedia, The Free Encyclopedia*. [Online; accessed 17-May-2021]. 2020. URL: https://en.wikipedia.org/wiki/Gain_(electronics).

[15] Cleo Lepart. *DIY Low Noise Radio Telescope*. 2021. URL: https://www.instructables.com/DIY-Low-Noise-Radio-Telescope/.

[16] Lulu Liu and Chris Chronopoulos. "The Hydrogen 21-cm Line and Its Applications to Radio Astrophysics". In: *MIT Undergraduate* (2008). URL: http://web.mit.edu/lululiu/Public/8.14/21cm/21cm.pdf.

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