(Re) Calculation of the Milky Way's rotation speed using the hydrogen line

Valentino Carneiro

In 1944, van de Hulst predicted the existence of a line, in the 21cm band (approximately 1420.4Mhz), of neutral hydrogen (HI) throughout our galaxy that, until then, was not considered a spiral. In 1951, the first attempts were made to receive this signal, which led to the mapping of the HI of our galaxy, revealing its spiral structure. The objective of this work was to redo these observations, calculating the speed of rotation of the galaxy as a function of the distance to its center. This was made with data that was collected using a DIY antenna and a circuit to amplify and filter the received signal. The signal was then converted to digital using a SDR module, trying as much as possible to reduce, whenever possible, the production costs. This work aims to show that it is possible for anyone, even on a low budget, to discover lots of information about our galaxy.

The end result of this work. It certainly wasn't easy, but it was possible to get there.

Quick note before moving on

This document is not intended to be totally scientific. The main goal is to present what was built and done, why, and what steps were made during the process, like a report. For this reason is shown much more information than a typical research paper and the structure is different. For instance, this report has many parts, as it is possible to see in the index, each one describing the various steps of the project. For more information (and images/videos) about the construction visit the following website, where this project was posted:
www.instructables.com/Automated-Radio-Telescope

Index

- About the signal to be received
    - Description of the frequency that needs to be detected, and why it was emitted
- Construction of the antenna and waveguide
    - Self explanatory, with some math to correctly understand where the antenna should be placed
- Construction of the horn
    - Measures and format of the pyramidal horn
- Rotation system
    - Description on why it was made, and how
- Amplification circuit
    - LNA concept explanation, DIY production of lnas, switches and filters
- Software Defined Radio (SDR)
    - What it is and which one was used
- GNU Radio
    - Explanation of this fantastic piece of software
- Elementary graph analysis
    - The 1st calibrated graph obtained. Analysis and explanation
- Data collection - Angles
    - The angles of observation that were used and why
- Doppler effect
    - How it is possible to calculate the relative velocity of the hydrogen
- Trigonometry and tangent points
    - Describes what part of the collect radiation is useful and should be used in the Doppler Effect formula
- Observation data
    - The results obtained and the math done to come to a conclusion
- Conclusion
    - Comparison to previous and more “scientific” observations
- References
    - Because this project would not be anything if they did not exist

About the signal to be received

Hydrogen is the most common element in the universe, so its mapping reveals important information about our galaxy (rotation speed, shape, density, etc.). One of the ways to detect neutral hydrogen (HI), at temperatures between 100 to 3000K, is through the reception of radiation that it emits, when a spin occurs flip transition. Both the electron and the hydrogen proton have their own rotation, and they can rotate in the same directions or in opposite directions. The state in which both rotate in parallel has more energy and, when the spin the electron changes, energy is released in the form of radiation, at (approximately) 1420.4 Mhz [FIG.1]. This is called spin-flip transition. Although this transition is rare, high amounts of gas, like the gas present in the arms of our galaxy, will emit radiation constantly. As the wavelength of this signal is about 21cm, it can pass through space dust more easily than visible light, offering a better description of the galaxy. Through the Doppler effect it is then possible to calculate the relative speed of this gas (it is expected that the frequency received may deviate by up to 2 Mhz from the value emitted), and using some trigonometry, one can calculate the speed of the arms in relation to the center of the galaxy.

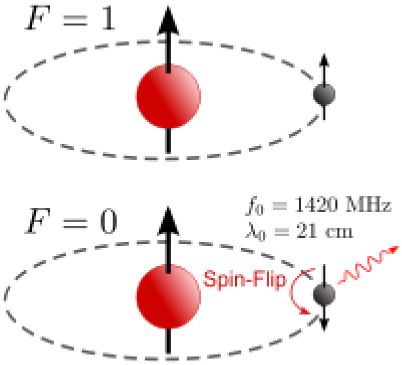


Figure 1

Construction of the antenna and waveguide

In order to receive a signal that is in the band of 1420.4 ± 2 Mhz, was built an antenna, as well as a pyramidic horn, and a waveguide, using aluminum foil. The antenna is basically a 5.25 cm long copper wire, corresponding to ¼ of the length from where the signal we want to receive, blocking higher frequency values. This was placed in the center of the waveguide, which has a rectangular opening, measuring 17.0 ± 0.5 cm by 10.4 ± 0.5 cm, being connected to an SMA input. The length value affects the electric fields that only move across the antenna. This way, these fields end up corresponding to the antenna polarization, and the width (10.4 cm) has no effect in the signal reception. The length of the waveguide has a great role in filtering the radiation that can be received, with a cut-off frequency of around 891Mhz, and a band of frequencies from ≈1.1Ghz to ≈1.7Ghz. These values fit the desired frequencies, 1420.4 ± 2 Mhz. Lastly, but by no means less important, we need to calculate how far from the bottom the antenna is placed. Once the the distance between the power peaks inside the waveguide is different from the free space (≈21cm) due to the angles (which are different from 0) in the contours of the waves, we need to calculate the wavelength inside the guide [1]:

$$\lambda = \frac{c}{v}\left(1 - \left(\frac{vc}{v}\right)^{2}\right)^{-1/2}$$

Being vc = 891Mhz and v = 1420.4Mhz, this value will be equal to 27 cm, so ¼ will be 6.8 cm, which is the distance from the antenna to the bottom of the waveguide where we should place the antenna.

Construction of the horn

To amplify the signal was also built, with aluminum foil, one pyramidic horn, connected to the waveguide (in front of it), as it is possible to see in the figure 2 (in the next page). The measures used were 5 provided by the DSPIRA project, designed for this specific use, with some changes made by myself in edges to allow fixing the entire system. To complete the horn, and in order to cover any imperfections, aluminum tape was used. Using CST software (because of lack of equipment to properly test it) it was possible to calculate the gain obtained for this specific antenna. The values calculated for the desired frequency vary between 14 and 17dB, values that already take in account building imperfections. Figure 3 shows the final result.

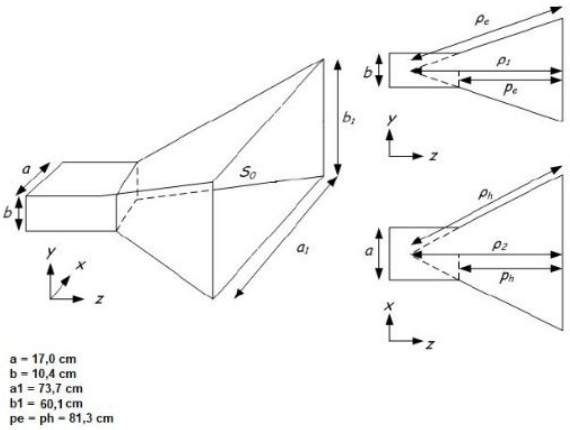


Figure 2

Figure 3

Rotation system

In order to automate the data collection process, a mechanized rotation system was created, that allows through a computer to change the latitude and longitude for which the radio telescope is pointing at. Using 2 motors (one for the xy axis and the other for the z axis) it is possible to point the device towards a semi-sphere that corresponds to all the sky. The coordinates to which it needs to point (latitude and longitude galactic) are introduced, and converted to a number of degrees that must be rotated, using the python library Astropy.coordinates. This is then sent to an arduino, which will rotate both the xy and z axis. The control of position of the xy axis is done through an accelerometer that integrates the received values, calculating the rotated angle and its used a motor from a drilling machine (and a crazy system to control it, because the controllers that were available were not capable of handling so much current (some of them ended up being destroyed)). As they are fast movements, little error is introduced, and it is not cumulative, since the system has a reference point, where it started. Image 4 shows the final result of this part of the rotation system. The Z axis control is done using a gyroscope. Using some fishing line, a stepper motor, controlled with a L298N, pulls the antenna, changing its angle, until reaching the desired value, when it stops. It was also built a platform where the antenna will stay, as is possible to see in the image 5. It is worth saying that this one of the hardest parts of the work, since controlling the drilling machine´s motor was simply too hard, and the motor sometimes would go faster than usual, others would not even move. A really headache.

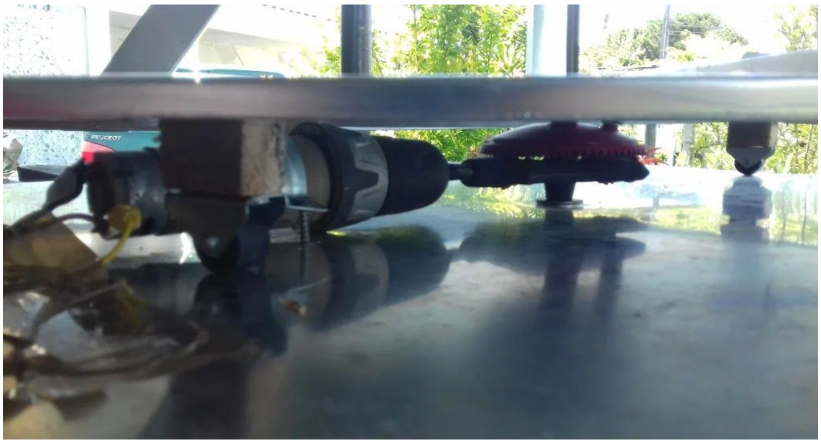


Image 4

Image 5

Amplification circuit

It is not possible to detect the signal only with the antenna connected to the SDR (the device that will capture the signal) since it adds a lot of noise to the system, becoming imperceptible. Thus, it is necessary to previously increase the signal strength, as well as filter it, to get actual results. In the DIY community, are used low noise amplifiers (LNA), devices that amplify the signal without adding too much noise. The cost of these is indeed high, as well as the filters, and it was decided that they were going to be built in the old fashioned way of DIY. .The signal, before being received by the SDR, follows this route route:

Horn > Waveguide> Antenna> Switch> LNA # 1> Filter> LNA # 2> SDR

The switch, the LNAs, as well as the filter were designed on a PCB with 4 layers using chips from Minicircuits. The first step is the Switch, which allows, using an arduino connected to the computer, to change the input signal for the first LNA, in order to use a resistance of 50 Ohm (resistance of all devices used in the system) as a control group, without the need for disconnect the antenna each time an observation is made. For this,it was used a VSW2-33-10W + module. The 2 LNAs are based on the same chip: PMA2-43LN +. These present a noise coefficient [2] of 0.45dB and a gain of 21.6dB, introducing little noise to the system. Finally, the filter used was the BFCN-1445 +, in which only frequencies between 1420-1470Mhz suffer little reduction, with a loss by insertion of 2.23dB, at the frequency of 1420.00 Mhz. A total of 4 PCB boards were created (respecting the rules of construction of radio frequency PCBs in order to obtain the best results), with 2 SMA connectors each. It is assumed that each SMA leads to a loss of 0.06dB in strength of the signal. However, this value is almost irrelevant, since all SMA connectors together have a total loss of 0.5dB, negligible value, when compared to the gain caused by LNA.

In conclusion:

$$Gain_{Circuit} = 41\ dB \text{ and } Gain_{System} = 55\ - 58\ dB$$

The noise coefficient can then be calculated using the Friis formula for noise:

$$F_{Total} = F1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + ... + \frac{F_n - 1}{G_1 G_2 ... G_{n-1}}$$

where values are given through ratios (not decibels). F represents the noise coefficient of each component of the system and G the gain the component. We can see now why it is possible to use a SDR to capture the signal, even having a noise coefficient of 4-7dB. As can be seen from the expression, only the first amplifier has a relevant role in the total noise coefficient, in this case the LNA#1. For this system:

$$F_1 = F_2 = 0.56\ dB = 10^{0.53/10} = 1.14 \text{ and } G_1 = G_2 = 21.6\ dB\ = 10^{21.6/10} = 145$$

Using the Friis formula:

$$F_{Total} = 1.14 = 1\ dB$$

Since these values are from the manufacturer (and due to the lack of equipment to test each chip) these values are expected to be “ideal” . Thus, the realistic total noise coefficient is between 1dB and 1.4dB or a ratio of 1.26 to 1.38, values from similar works. Its important to note that, although the system reduces the signal strength by 26-38% because of the noise, it ends up being amplified more than 300.000 times. Unfortunately, it was not possible in this work to do performance tests on the circuit built and, because of this, only theory can function as a support, at least in this part of the project.

Software Defined Radio (SDR)

As already highlighted, the collection and digital processing of data will be done through an SDR module, in this work an RTL-SDR. This small gadget features a frequency band between 10Mhz and 1.7Ghz (although a downconverter can also be used for higher frequencies), with a 50 Ohm impedance and a 2.4Mhz bandwidth. This value is too low for the desired observations (we try to analyze them in a bandwidth of 4Mhz) and, therefore, for each analyzed position of the sky it is necessary to make two iterations of data collection in order to solve this problem. Since it is expected the emitted radiation to be constant, this won't cause any problem because the interval between both iterations is less than 5 seconds, and therefore the motion of the Earth will not be a problem as well. The SDR has the job of receiving the signal and then convert it to digital, where it can be analyzed.

GNU Radio

GNU Radio is an open-source tool that allows, through flowcharts, receive, filter, calibrate and even make calculations with the signal it receives, which makes it ideal for this type of application. The flowchart used, spectrometer_w_cal.grc [3], was created by the DSPIRA project, but some parameters were modified. This flowchart has several parts that will be explained. First. it is necessary to take into account the sampling rate, whose maximum module used is 3.2MS / s. However, this value is very unstable and could compromise the fidelity of the observations. Thus, 2.6 MS / s was used, the maximum stable value for the RTL-SDR. The internal gain of the module has been set to 0. Now, about the flowchart in specific: after the signal is received, there are two blocks that will allow to analyze the signal. The first is a spectrometer, that has the function of dividing the power received in small “boxes”, per unit of frequency. So, the more energy is received for a given frequency in a given range the longer the peak will be when plotted. Then, there is a polyphase filter block that works as support for the previous one, improving results [4]. For this step, it is used a Fast Fourier Transform (FFT), which transforms the time function into a frequency function, since the received signal is considered to be constant over time. The FFT size was changed from 4096 to 1024 “boxes”, to avoid adding an error due to noise. Thus, the data will be divided into ≈ 2 KHz intervals. There is also a Complex to Magnitude ^ 2 block, which multiplies the FFT data by their complex conjugates, converting them to IR. The last part of this same block allows the cancellation of more noise, since it calculates the average signal of several iterations, integrating them. The time chosen for each observation was 2 seconds. The flowchart has many other blocks that are not going to be explained here, since most of them is just there to support the main one, already explained.

To facilitate the autonomous data collection, in addition to the observation window being removed, it was enabled the setting to save the data in a .csv file, for future automatic analysis and calculation of the relative velocity of the gas mass. Here is the already calibrated result of an observation (Image 6). The antenna was pointed at the following galactic coordinates: latitude = (-6 ± 1) degrees and longitude = (151 ± 1) degrees.

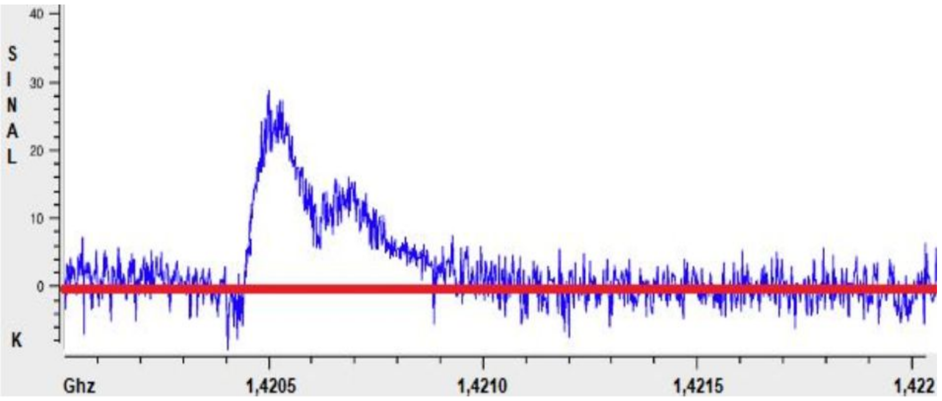


Image 6

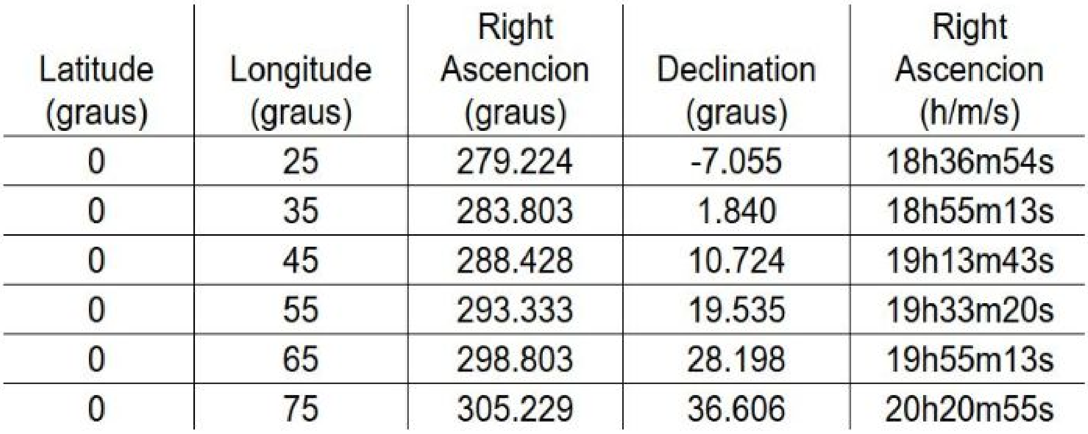
One of the aspects of this graph to be highlighted is the ordinate axis, whose scale is in Kelvin. Since FFT describes the power per unit of frequency (W / Hz) this would be expected to be the scale of the results. However, in astronomy, it is used a temperature scale, in Kelvin. More specifically, the temperature that a resistance (in this case 50 Ohm) would generate if a power equal to the same value that the antenna is receiving was applied to this resistance [5].

Elementary graph analysis

The first thing one can notice right away when looking at the previous graph is that there is a peak, and right after another, with less amplitude. In addition, it is possible to observe that the signal was blue-shifted,so this mass is coming closer to us. But, going back to the two peaks, such suggests that there is an enormous amount of gas, then a space with no HI, and again more gas. As this continues to happen for other longitudes, it is possible to say that the galaxy has a structure of Curved “arms” with at least 2 of these: the Perseus one and the one where we are. If the galaxy was elliptical or irregular, there would be countless peaks in the graph.

Data collection - Angles

Before developing this part, it is important to mention some important aspects of the observation: it was assumed that the movement of the gas mass is uniform, as already mentioned, that the energy emitted over time is also constant, and (in order to simplify trigonometry), that the galaxy is on a two-dimensional plane, hence latitude = 0º is used. In addition, although the trigonometric methods described below can be used between -90º and 90º longitudes, some problems reduce the possible angles of longitude of observation from ≈25º up to ≈75º. Since we are in the northern hemisphere, the angles less than 0º were excluded. Furthermore, close to the center of the galaxy, it is no longer possible to consider that the movement of the gas masses is uniform circular. For angles close to 90º, because of the distance from the planet Earth and the HI masses is smaller, the observed deviation is not due to the movement of the mass of gas as a whole, but rather to differences in the speed that are more local and specific. The following chart shows, for each longitude (reinforcing again, in the galactic coordinate system), the values of right ascension. (both in degrees and in hours / minutes / seconds) and the declination of the observed longitudes.



Latitude (graus)	Longitude (graus)	Right Ascencion (graus)	Declination (graus)	Right Ascencion (h/m/s)
0	25	279.224	-7.055	18h36m54s
0	35	283.803	1.840	18h55m13s
0	45	288.428	10.724	19h13m43s
0	55	293.333	19.535	19h33m20s
0	65	298.803	28.198	19h55m13s
0	75	305.229	36.606	20h20m55s

Doppler effect

Since the frequency that HI emits is accurately known (1,420,405,751.7667 ± 0.0009 Hz to be more specific), the Doppler effect can be used to calculate the relative velocity of the gas mass in relation to the planet Earth. For this, it is used the formula presented next, where f_0 is the emitted frequency and c the speed of light in the vacuum.

$$\Delta f = \frac{\Delta v}{c} f_0$$

Trigonometry and tangent points

As seen in the graphic presented before, a wide variety of frequencies have shifted (both red to blue), which means that different gas masses are being observed at different speeds. As the goal is to compute the speed as a function of the distance to the galaxy core, this becomes problematic. However, for angles between -90º and 90º and if the angular velocity does not increase more than a linear function of distance (which explains why one can´t use this method for longitudes near the center of the galaxy) one can determine the frequency that should be used. The next image (Image 7) explains the concept. The point with higher relative speed (higher observed deviation) is what is aligned with our line of sight, that is, the point of tangency in relation to the Sun (B, in the image). This happens because the more / less close these points are to us, the smaller the Doppler effect will be, since the relative speed in relation to us will be less. Furthermore, the point of tangency is the closest point to the center of the galaxy, in our line of sight. This distance can then be calculated with simple trigonometry. The speed from the point of tangency to the center of the galaxy corresponds to the sum of its speed in relation to the Sol (which is calculated using the Doppler effect) with the velocity of rotation of the Sun in relation to the center of the Milky Way.

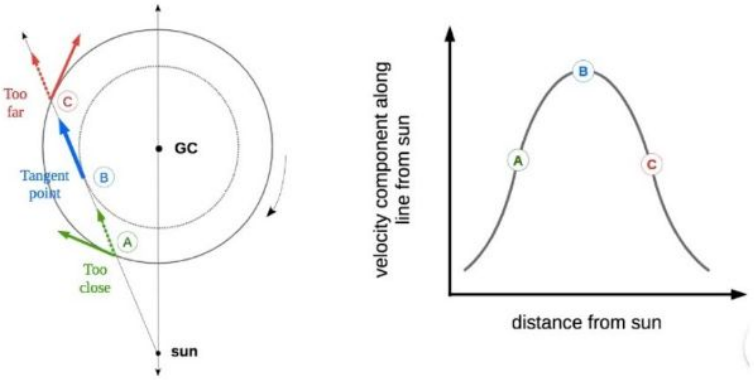


Image 7

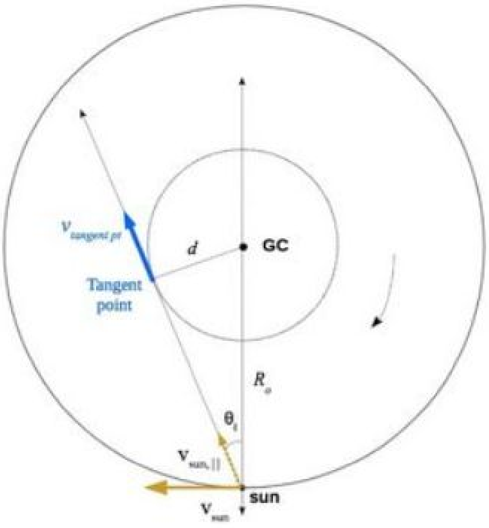


Image 8

As it is possible to see in Image 8 (at left),

$$v_{tangentpoint/galaxycore} = v_{observed} + v_{sun} * sin(\theta_l)$$

Also,

$$d = R_0 * sin(\theta_l)$$

Unfortunately, the speed of the sun is a data for which it was not possible to obtain any reference that had studied this value. So, it was assumed that $v_{sun} = 200\ km/s$. [6] Furthermore, it was also considered that the distance from the Sun to the center of the galaxy is constant throughout the observations with a value of 8.17 kpc.

Observation data

For control, the data collection was switched from the antenna to the 50 Ohm resistance, which, as expected, did not produce any results. Only noise. The following table shows the obtained results:

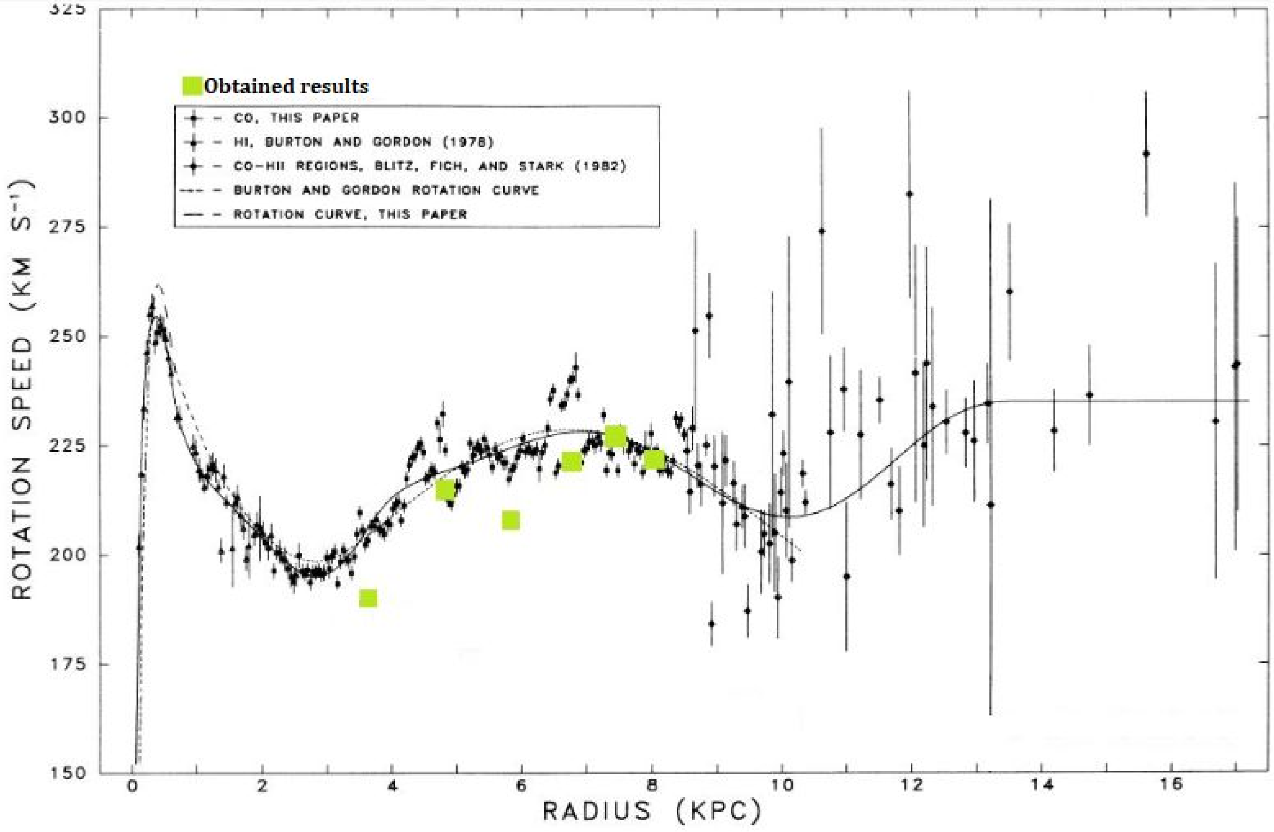
Longitude (degrees)	Distance to the galaxy core (kpc)	Relative velocity (km/s)	Sun´s velocity in relation to the tangent point (km/s)	Velocity of the mass of gas to the galaxy core (km/s)
0	0	0	0	0
25	3.5	115.5	85	201
35	4.7	108.4	115	223
45	5.8	69.8	141	211
55	6.7	67.2	163	230
65	7.4	49.1	181	230
75	7.9	22.2	193	215

Although these results may seem satisfactory, no account was taken of the speed of the Earth in relation to the Sun, which is about 30km/s, leading to a big error in the final results. Recalling the graph obtained, the first peak observed was due to gas present near the solar system, in our arm of the galaxy. This way, it can be affirmed that the speed of the Sun relative to this gas will be relatively similar (although not completely). Using the Doppler effect, again, it is possible to calculate the relative speed of the Earth to the Sun and subtract the result obtained previously. The new table looks like this:

Longitude (degrees)	Velocity of the mass of gas to the galaxy core (km/s)	Velocity of the Earth in relation to the Sun (km/s)	Adjusted velocity (km/s)
0	0	0	0
25	201	16	185
35	223	15	208
45	211	13	198
55	230	11	219
65	230	8	222
75	215	5	210

Conclusion

The results are shown below, in a graph [7], along with some of the studies already done by the masterminds of research in this topic. As already mentioned, the values obtained have a huge error, of about ± 30 km / s, but it was possible to reduce this value. Because of the lack of means to obtain better results about the rotation of the earth in relation to the Sun, it is impossible to obtain with certainty the margin of error for each observation. Which is lower than 30 km / s that's very likely. Although other errors can also alter the results (when using the Doppler effect in FTT, an error of ± 2 km / s is shown, for example), these, in the end, turn out to be irrelevant compared to the error presented by the movement of the Earth.



References

- Image 1 is from the wikipedia page on the hydrogen line
- Figure 2 adapted from Guardiola, Marta (2012) : “Fabrication and Measurement of Homemade Standard Antennas”
- Images 7 and 8 are from the DSPIRA project and are available online on their page wvurail.org/dspira
- [1] Essential Radio Astronomy, Cap 3.4 - Waveguides
- [2] Values from the manufacturer at the frequency of 1.5Ghz
- [3] Available for download on the DSPIRA project homepage
- [4] Edmund Lai, “Practical Digital Signal Processing”
- [5] Essential Radio Astronomy, Cap 3.6
- [6] Not very ortodox, but the only value found was in the NASA´s website solarsystem.nasa.gov/solar-system/sun/in-depth/
- [7] Prof. Dale E. Gary, NJIT, Astrophysics II: Lecture #19
- Ewan, H. I.; Purcell, E. M. (1951) "Observation of a line in the galactic radio spectrum”
- KTonix, “Using Two Tri-Axis Accelerometers for Rotational Measurements”
- Helmuth Hellwig; et al. (1970). "Measurement of the unperturbed hydrogen hyperfine transition frequency"
- Eric Trumbauer and Sahar Khashayar, “Welcome to...The Astronomy Zone!”