DOI: 10.54254/2753-8818/18/20230450

Separating 21-cm M31 signal from background

Wenlan Miao

Sichuan University, Chengdu, China, 610065

miaowl12@163.com

Abstract. By investigating 21 cm radio wave emitted by Hydrogen gas in Milky Way galaxy we can determine the intensity map of M31(a.k.a.,"Andromeda"). The author applied the data sets collected by 18m radio telescope pointing at different locations of M31 and using FFT, double subtraction method and matplotlib to visualize the data in the term of power spectral density-Doppler velocity spectrum. Then the author compared the spectrum and used the structure of M31 to explain the difference between the on-source spectrum and background spectrum.

Keyword: radio astronomy, andromeda, galaxy structure, spectrum analysis.

1. Introduction

Hydrogen gas is an important component of Milky Way galaxy which makes up 10 % of the mass of our galaxy. 21 cm corresponds to 1420.4 MHz radio signal emitted by neutral hydrogen. The signal can be detected by radio telescope and by analyzing the data from radio telescope, we can get the trajectories of these hydrogen gases and deduce the structure of the galaxy.

M31 (a.k.a., “Andromeda”) is a spiral galaxy with a massive disk like structure located in the direction of Andromeda, with the Messier catalog number M31 and the New General Catalog number NGC 224[1], is 15200 light years in diameter [2], 2.54 million light years from Earth. It is the closest large galaxy to the Milky Way [3]. It is expected to collide with Milky Way and form a giant elliptical galaxy [4] or a large lenticular galaxy [5]. Also, it is one of the most distant objects for naked eye to see.[6]

It is possible to detect 21-cm signals from M31, because it is large enough to produce a 21-cm signal that can be detected with the 18-m dish. The signal is relatively weak, however, so special care must be taken to separate it from the background noise. TLM-18 was applied to observe M31 galaxy. It is a Parabolic Dish Antenna which can be operated from the control console, or remotely via the Internet. Now the TLM-18 is being operated remotely from Princeton University. It has the radius of 18m and can observe 21-cm signal from hydrogen.

The Doppler effect formula is introduced to convert the observed frequency to Doppler velocity. The formula is

$$f' = \left(1 + \frac{v}{c}\right) f \tag{1}$$

where f’ is observed frequency, f is the frequency of the wave the hydrogen gas emitted, v represented the approaching velocity of the stars, and c is the speed of light. When v is positive the stars are moving toward the detector, when v is negative the stars are moving away from the detector.

DOI: 10.54254/2753-8818/18/20230450

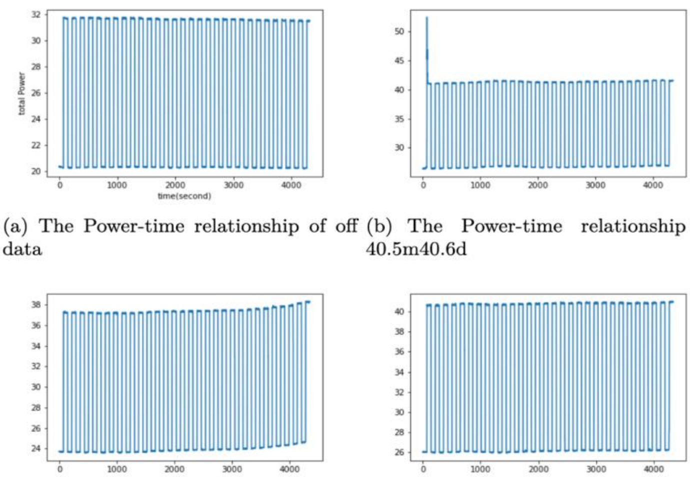
2. Code development

2.1. Data format

In order to extract clear signals, "double subtraction" approach was introduced. In particular, a subtraction step will consist of subtracting the signal from the electronic noise to obtain a spectrum with only the signal. The second subtraction will consist of comparing two noise subtracted spectra: one taken with the telescope pointed at M31 and the other taken with the telescope pointed at the area of the sky that is slightly removed from M31. There are three sets of data taken with the telescope pointed at M31 from different direction. They were called "on-source" data and were labeled as "40.5m40.6d", "42.5m41.25d", "45.0m42.0d". The substrings represent the right ascension and declination of the direction in which the telescope is pointed. The ch1 substring refers to channel being used. For "40.5m40.6d", it represents that telescope pointed to the direction with $0^{h}40.5^{m}$ right ascension and 40.6° declination. For "42.5m41.25d", it represents that telescope pointed to the direction with $0^{h}42.5^{m}$ right ascension and 41.25° declination. For "45.0m42.0d", it represents that telescope pointed to the direction with $0^{h}45^{m}$ right ascension and 42° declination. The data taken with the telescope pointed at the area of the sky that is slightly removed from M31 was called "off-source" data, and was labeled as "offdata". The telescope is pointed at a region of the sky slightly different from M31.

The data were taken a noise source was turned on and off once every ten seconds or so. The noise source generated broadband "white noise".

2.2. Ploting the total power from the receiver as a function of time



(a) The Power-time relationship of off data
(b) The Power-time relationship of 40.5m40.6d
(c) The Power-time relationship of 42.5m41.25d
(d) The Power-time relationship of 45m42.0d

Figure 1. The power-time relationship of data sets.

The noise source is turned on and off every ten seconds or so during the data acquisition process. This noise source produces broadband "white noise", which corresponds to a flat spectrum. To extract information from this spectrum, function ‘getSpectrum()’ was defined. It does not actually return a spectrum. Rather it returns a time series called totalPower by summing over all bins for each spectrum. Thus, total power from the receiver as a fuchtion of time can be plotted if it is coverted to numpy array. Since the values are very large, each element was multiplied by 10^{-6}. The plot is shown in figure 1, where X-axis means time in seconds and Y-axis represents the total power.

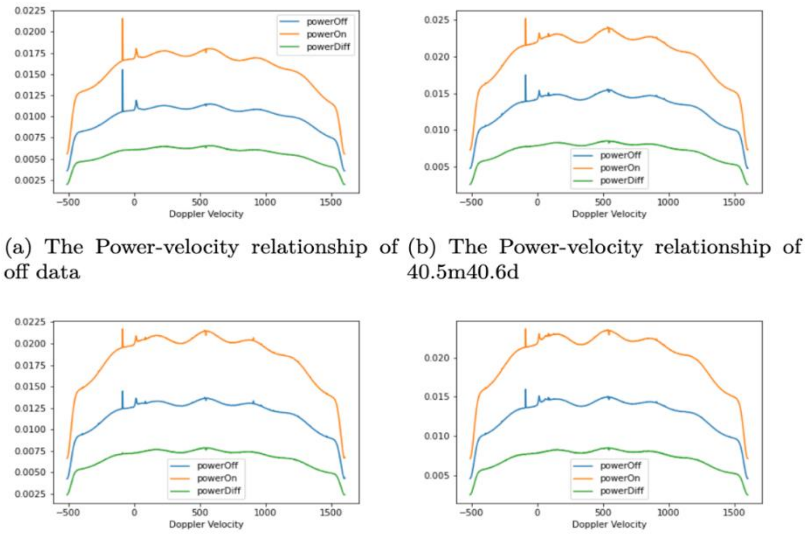
DOI: 10.54254/2753-8818/18/20230450

2.3. *Dividing the spectra into two classes*

Figure 1 shows that the points are generated in to groups, which means that there are two values of total power. The upper points occur when the noise generator is on and the lower points occur when the noise generator is off. Accordingly, two classes of spectra were divided by declaring a threshold value. Spectra with total power greater than the threshold are noise on and spectra with total power less than it are noise off.

However, A closer look reveals that there are some cases that fall somewhere in between. These situations occur when the noise generator switches from a power-on state to a powered-off state while acquiring the spectrum. These transition cases can be eliminated by developing two more detailed thresholds "HighThres" and "LowThres". When the spectra is lower than LowThres, it is considered as "powerOff" spectra; when the spectra was higher than HighThres, it is considered as "powerOn" spectra.

For Fig.1(a), the HighThres is 31, the LowThres is 21; for Fig.1(b), the HighThres is 40, the LowThres is 28; for Fig.1(c), the HighThres is 36, the LowThres is 26; for Fig.1(d), the HighThres is 40, the LowThres is 27. As a result, power on and power off spectras were divided. Next the average power of powerOn and powerOff spectrum can be calculated. Accordingly, the difference between powerOn and powerOff spectrum can be computed.



(a) The Power-velocity relationship of off data

(b) The Power-velocity relationship of 40.5m40.6d

(c) The Power-velocity relationship of 42.5m41.25d

(d) The Power-velocity relationship of 45m42.0d

Figure 2. The Power-velocity relationship of data sets.

Then three power-time relationship curves were plotted for each data set. Meanwhile, the data sets contained not only the information of time and power but also the frequencies. Each time has a frequency respectively. Using Doppler effect formula, the observed frequency were converted to Doppler velocity. Thus, power-time relationship diagrams were converted to power-frequency relationship diagrams. Figure 2 shows the results of this section, where the X-axis means Doppler velocity in km/s and the Y-axis means average power.

There are three curves for each diagram that are all similar in appearance. The top curve is the spectrum with the noise generator on, the middle spectrum is the spectrum with the noise generator off, and the lower spectrum is the difference. Note that the difference spectrum is relatively free of peaks. That is because any signal peaks present in the data are common to both the power on and power data and are therefore removed by the subtraction.

DOI: 10.54254/2753-8818/18/20230450

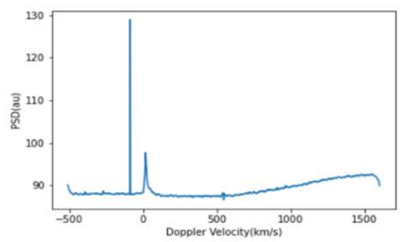
2.4. Calibrating the system

The lower curves of Fig.2 represent the difference between the powerOn and powerOff state which means the lower curves represent the response of the system to the noise generator (gain response). The white noise, however, has been removed by the subtraction step because the white noise distribute to same peak in powerOn and powerOff curve. As a result, by calculating the difference of two curves, the peaks are removed and the influence of white noise disappeared.

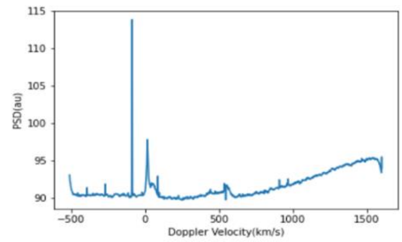
In this step, a function "smoothedPower" was defined to remove the fluctuation in gain response. The function applied savgol_filter from scipy to smooth the PowerDiff curve. The smoothed data set is called "smoothedPower" and it is an array. Then use the following equation to get the gain-corrected function that only contains signal.

$$\text{correctedSpectrum} = 50 * \frac{\text{powerOff}}{\text{smoothedPower}} \tag{2}$$

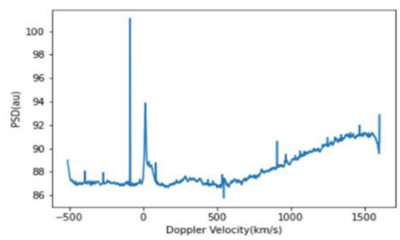
The result is shown in figure 3, where the X-axis means the Doppler velocity, the Y-axis means power spectral density.



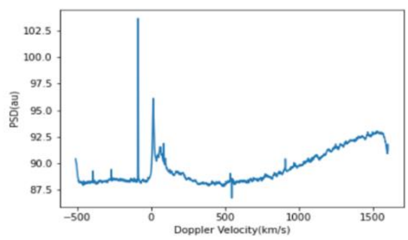
(a) The PSD-velocity relationship of off data



(b) The PSD-velocity relationship of 40.5m40.6d



(c) The PSD-velocity relationship of 42.5m41.25d



(d) The PSD-velocity relationship of 45m42.0d

Figure 3. The PSD-velocity relationship of calibrated data sets.

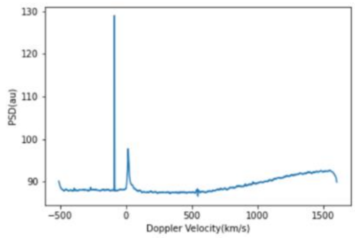
Ideally, the baseline of the spectrum should be flat over the entire range. This is almost correct for the range $-500 < v_D < 500\text{km/s}$, but for reasons not currently understood there is a gradual increase of PSD when $v_D > 600\text{km/s}$. However, the region of interest for M31 is approximately $0 < v_D < 600\text{km/s}$, so this does not affect the observation of M31. The value of the flat part is about 88 K, which is the effective temperature of the system. Fig.3(a) is background spectrum which was taken with the telescope pointing slightly away from M31. There are two peaks: the large sharp peak at $v_D = -100\text{km/s}$ is a spurious signal in the receiver chain and is a persistent feature in nearly all spectra. The broader peak near $v_D \approx 50\text{km/s}$ is a real 21 cm signal, presumably from the Milky Way.

It is obvious that the baseline value of Fig.3(b), Fig.3(c), Fig.3(d) is different from the baseline value of the background spectrum because drifts in the system lead to slight changes in its apparent noise level. Thus, further optimization is necessary. The way to modify the spectrum is to subtract the noisebase which is the average of samples in the range $-400 < v_D < 100\text{km/s}$. After the subtraction, the spectrum is successfully modified.

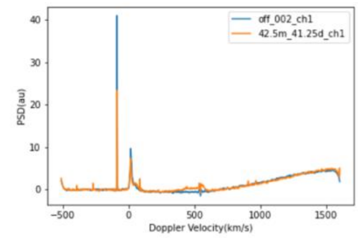
DOI: 10.54254/2753-8818/18/20230450

3. Results

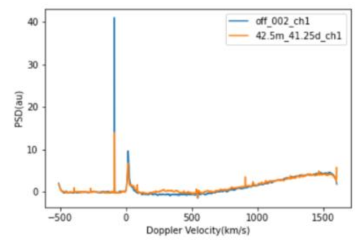
Plotting the modified spectrum of Fig.3(b), Fig.3(c), Fig.3(d) with the background spectrum in the same plot, the result is shown in figure 4, where the X-axis means the Doppler velocity, the Y-axis means power spectral density.



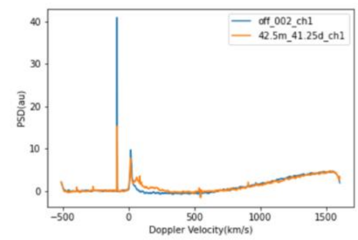
(a) The PSD-velocity relationship of off data



(b) The PSD-velocity relationship of 40.5m40.6d



(c) The PSD-velocity relationship of 42.5m41.25d



(d) The PSD-velocity relationship of 45m42.0d

Figure 4. The PSD-velocity relationship of final data sets.

The corrected spectrum should line up in the region $v_D < 100km/s$. Comparing the On-source signal (pointing at M31) with the off source signal (background spectrum), there are differences because the curves did not overlap. For higher values of v_D, the differences can be attributed mostly to signals coming from M31, which means the signal from M31 was successfully detected. However, since the signal from the Milky Way will change in the direction of the dish, the region around $v_D \approx 0$ will be affected by the Milky Way.

4. Conclusion

Comparing three On-source spectrum with each other, the M31 signal is slightly different. The structure of M31 contributes to these differences. M31 is a spinning disk whose angular size is 3∘ from the verticle scale that is larger than the angular resolution of TLM18(± 0.5 ∘). Stars are not uniformly distributed in M31, which means there are different amounts of stars in different locations. As a result, the PSD of three right ascension are different. It can be indicated that 45m42.0d had densest hydrogen gas. Moreover, M31 is a spinning disk, so the velocities of different locations on the disk are different. As a result, the peaks of PSD locate differently.

More investigation is needed to further the study of M31. The work this paper did only show the overall structure of M31. It is difficult to study the structure of M31,. Many scientists applied more advanced method to make detail measurements of rotational velocity of M31, and figured out its spiral structure. Walter Baade is the first man to thoroughly study the spiral arms of M31. His studies shown two tightly wound spiral arms [7]. However, the the Andromeda Galaxy is seen close to edge-on. After the rectification of the images of galaxy, the spiral structures were proposed as the alternative explanation [8].

DOI: 10.54254/2753-8818/18/20230450

References

[1] Results for Messier 31. NASA/IPAC Extragalactic Database

[2] De Vaucouleurs, Gerard; De Vaucouleurs, Antoinette; Corwin, Herold G.; Buta, Ronald J.; Paturel, Georges; Fouque, Pascal (1991) Third Reference Catalogue of Bright Galaxies

[3] Ribas, Ignasi; Jordi, Carme; Vilardell, Francesc; et al. (2005). "First Determination of the Distance and Fundamental Properties of an Eclipsing Binary in the Andromeda Galaxy". Astrophysical Journal Letters. 635 (1): L37–L40.

[4] "NASA's Hubble Shows Milky Way is Destined for Head-On Collision". NASA. 31 May 2012. Archived from the original on 4 June 2014. Retrieved 12 July 2012.

[5] Ueda, Junko; Iono, Daisuke; Yun, Min S.; et al. (2014). "Cold molecular gas in merger remnants. I. Formation of molecular gas disks". The Astrophysical Journal Supplement Series. 214 (1): 1

[6] "M 31, M 32 & M 110". 15 October 2016.

[7] Arp, Halton (1964). "Spiral Structure in M31". Astrophysical Journal. 139: 1045.

[8] Simien, François; Pellet, André; Monnet, Guy; et al. (1978). "The spiral structure of M31 – A morphological approach". Astronomy and Astrophysics.