

# UV and optical cross correlation studies of AGN at different time-scales

University of Crete

Department of Physics



Undergraduate Thesis

Langis Dimitrios Alkinoos  
Supervisor : Prof. Papadakis Iossif

September 2023

# Acknowledgments

I would like to express my sincere gratitude to the following individuals for their support and contributions to the completion of my undergraduate thesis.

- My supervisor, Professor Papadakis Iossif for his mentorship, patience and guidance throughout this project.
- My family and friends, for their understanding and encouragement during my academic journey.
- My godmother, for her spiritual guidance and psychological support throughout my undergraduate studies.
- The staff and faculty of the department of physics in the University of Crete, who assisted me in the procedure for the presentation of my thesis.

I deeply appreciate everyone who helped me achieve this milestone. Your support has been invaluable, and I am truly thankful.

# Contents

<b>Acknowledgments</b>	<b>1</b>
<b>1 Introduction</b>	<b>3</b>
1.1 Gravitational power . . . . .	4
1.2 AGN variability . . . . .	5
<b>2 Fourier Transform of Light Curves</b>	<b>7</b>
2.1 Definitions . . . . .	7
2.2 Fourier Decomposition of Fairall9 Light Curves . . . . .	9
2.3 Light Curve Interpolation . . . . .	10
2.4 The Longest Sampled Time-Scale . . . . .	12
<b>3 Cross Correlation Analysis</b>	<b>13</b>
3.1 Stationary Processes . . . . .	13
3.2 Cross Correlation Function . . . . .	13
3.3 Application in the case of the Fairall9 longest sampled variations . . . . .	15
<b>4 Data Analysis</b>	<b>18</b>
4.1 The sample of AGN and light curves . . . . .	18
4.2 Fourier Decomposition . . . . .	22
4.3 Cross correlation results . . . . .	25
<b>5 Summary and Conclusions</b>	<b>29</b>
<b>References</b>	<b>31</b>
<b>Appendix</b>	<b>32</b>

# Chapter 1

## Introduction

An active galactic nucleus (AGN) is an extremely bright compact central region of a galaxy, emitting substantial amounts of energy that spans across the entire electromagnetic spectrum. These objects have characteristics that indicate it would be impossible for their luminosity to be produced by stars. Moreover, AGN are the most luminous persistent sources of electromagnetic radiation in the universe, with typical bolometric luminosity around  $10^{42}$  to  $10^{48}$  ergs per second, with their visible emitting area being not much larger than the size of our solar system. For example, the bolometric luminosity of a quasar, which is the most luminous type of AGN, is  $\sim 10^{43}$  to  $10^{48}$  ergs per second, while the respective luminosity of their host galaxy is  $\sim 10^{42}$  to  $10^{44}$  ergs per second, meaning that in a lot of cases AGN outshine their host galaxy. It is estimated that around 10 – 20% of the galaxies in our local universe host an AGN. We classify an object as an AGN based on some specific observational characteristics which are listed below.

1) AGN emit approximately the same amount power over a broad range of wavelengths. The range of emission covers wavelengths from X-rays to even radio waves, as seen in the left panel of Fig.(1.1). AGN that show strong emission of radio waves are called radio-loud AGN, while the rest are called radio-quiet AGN.

2) The optical/UV spectrum increases towards shorter wavelengths. This is indicated by the label "Big blue bump" in the left panel of Fig.(1.1). This is a characteristic exclusive to AGN, for this behaviour at UV is not observed in normal galaxies, where the spectral energy distribution (SED) is dominated by the collective light of the stars. The emission towards UV in these galaxies decreases.

3) Presence of broad and narrow emission lines in the optical/infrared band, which arises from cold material close and distant to the active nucleus, respectively.

The large luminosity, in combination with the small emitting area implies that a very efficient mechanism is needed in order to produce this amount of energy per unit of time. Our current understanding is that AGN are powered by the accretion of matter onto a super massive black hole (SMBH) located at the central region of the host galaxy. The matter falling into the SMBH will form a disc around it, due to angular momentum, which dictates the trajectories of particles as they orbit the black hole. Interactions among these particles result in the redistribution of angular momentum, causing them to adopt a more stable and less collision-prone orientation, forming a disc. As the material falls towards the SMBH, potential energy is released due to the gravitational pull. If this energy can be converted to thermal energy and heat the material of the disc, then the disc can be very luminous in the optical/UV band. Below we give an example in order to understand the efficiency of this mechanism.

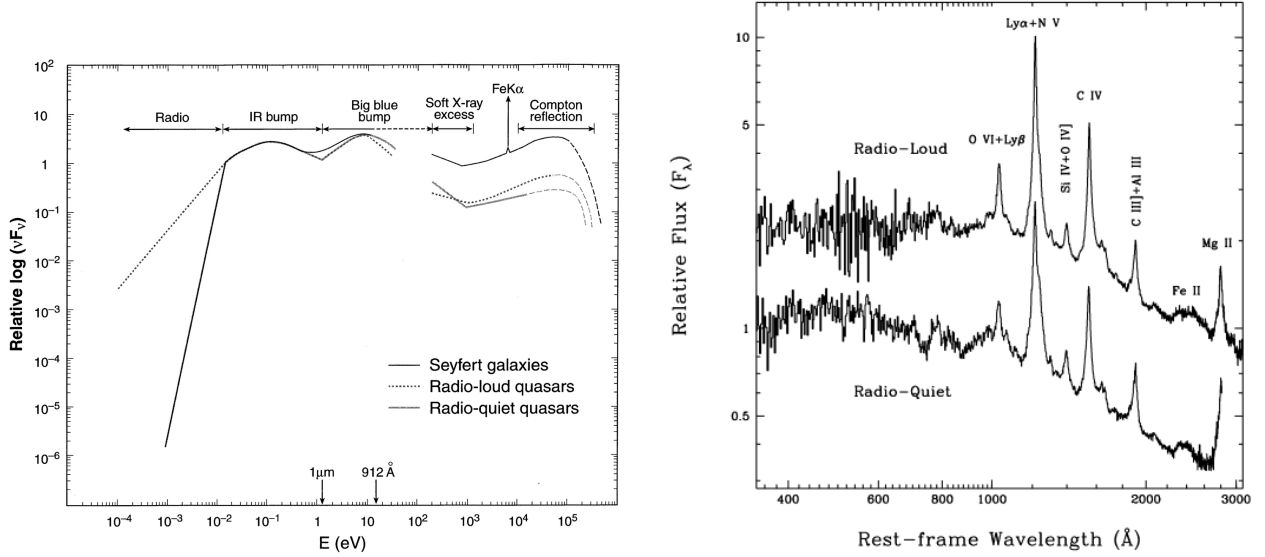


Figure 1.1: *Left panel*: Schematic representation of the SED of different types of AGN (figure taken from Koratkar and Blaes (1999)). *Right panel*: Extreme UV spectra of 60 radio-loud and 41 radio-quiet quasars. Radio-quiet quasar flux level is shifted down for display purposes (figure taken from Zheng et al (1997)).

## 1.1 Gravitational power

Let us assume a particle, say an electron with mass  $m_e$ , that falls from infinity to a distance  $r$  from a central SMBH, with mass  $M_{BH}$ . Then, the change in potential energy will be equal to :

$$\Delta U = -\frac{G \cdot M_{BH} \cdot m_e}{r}. \quad (1.1)$$

Let us assume a non-rotating SMBH, then the smallest possible stable circular orbit for the electron will be equal to three times the Schwarzschild radius,

$$R_S = \frac{2 \cdot G \cdot M}{c^2}. \quad (1.2)$$

We can substitute  $r = 3 \cdot R_S$  using eq.(1.2) to eq.(1.1), giving us :

$$\Delta U = -\frac{G \cdot M_{BH} \cdot m_e}{3 \cdot R_S} = -\frac{1}{6} \cdot m_e \cdot c^2. \quad (1.3)$$

The Virial theorem, which describes a system in equilibrium, states that half of the change in potential energy will be converted into kinetic energy and the other half will be converted to radiation by heating the material. Using the Virial theorem and (1.3), we finally get :

$$E_{heat} = -\frac{\Delta U}{2} = \frac{1}{12} \cdot m_e \cdot c^2. \quad (1.4)$$

The constant  $\frac{1}{12} \approx 0.08$  is the efficiency coefficient of the process, which indicates that the aforementioned mechanism is around 10 times more efficient than the one that produces energy in the core of main sequence stars (the efficiency coefficient there is  $\epsilon \approx 0.007$ ).

## 1.2 AGN variability

A defining property of AGN is flux variability. AGN show flux variations at all wavelengths, from radio to optical/UV and hard X-rays. Variations can be observed on timescales ranging from minutes, in X-rays, to years, in infrared. Both the amplitude and the characteristic frequencies increase with increasing energy. Variability provides information about the physical processes occurring around the SMBH at the center of an AGN. We can separate the variability into slow variations and fast variations. Slow variations refer to changes in the properties or behavior of AGN that happen over relatively long timescales, ranging from weeks to years or even decades, while fast variations refer to rapid and relatively short-term changes that occur on timescales ranging from minutes to hours or days. Moreover, the slow variations provide insights into the long-term evolution and underlying physical processes of AGN, while fast variations provide information about the physical processes occurring near SMBHs.

Accretion discs can vary on various timescales. The most important disc timescales are the following (taken from Czerny (2006)).

1) The dynamical timescale,  $t_{dyn} = \Omega_K^{-1} = \sqrt{\frac{G \cdot M}{r^3}} = 10^4 \cdot R_3^{\frac{3}{2}} \cdot M_8 (s)$ , where  $M$  is the BH mass,  $r$  the radius,  $R_3 = r/(3 \cdot R_S)$  and  $M_8 = M/(10^8 \cdot M_\odot)$ . This timescale describes various motions inside the system. These include the circular orbits motion, the local rotation in sync with the epicyclic frequency (for instance, when a magnetic loop emerges from the discs surface, its lower end becomes entangled on this timescale), the dynamical oscillations which occur perpendicular to the discs surface, the timescale needed to attain hydro-static equilibrium and the timescale for sound movement in the vertical direction of the disc.

2) The timescale of propagation of sound waves in the radial direction given by  $t_{sound-r} = t_{dyn} \cdot (\frac{r}{h_d})$ , where  $h_d$  is the thickness of the disc. Typically, the radius is greater than the thickness of the disc, meaning that  $t_{sound-r}$  is greater than  $t_{dyn}$ .

3) The thermal timescale, which is defined as the ratio of internal energy to the rate of cooling or heating. It is given by  $t_{th} = \alpha^{-1} \cdot t_{dyn}$ , where  $\alpha$  is the disc viscosity, introduced by Shakura & Sunyaev (1973). Assuming  $\alpha = 0.1$  as a value for the viscosity parameter, we get that  $t_{th}$  is 10 times greater than  $t_{dyn}$ . This timescale is neither dependant on the optical depth of the disc, nor the cooling mechanism, meaning the same thermal timescale is present for a cold optically thick disc and a hot optically thin disc.

4) The viscous timescale is established as the timescale governing the flow of mass, for instance locally, as the ratio of the radius to the radial velocity. It is given by  $t_{visc} = t_{th} \cdot (\frac{r}{h_d})^2$  assuming an  $\alpha$  disc model. If  $h_d \simeq 0.1 \cdot r$ , then  $t_{visc}$  is  $\simeq 100$  times longer than  $t_{th}$ .

Figure (1.2) shows a typical example of how AGN vary. Figure (1.2) shows the light curves of Fairall9, which is a typical AGN, in many wavebands (figure taken from Hernandez Santisteban et al (2020)). A light curve shows how the brightness of an object changes over time, by plotting the flux in the  $y$ -axis and the time of the observation in the  $x$ -axis. The light curves in Fig.(1.2) show data taken in a period of  $\sim 8$  months, almost daily, across multiple wavebands ranging from extreme UV to almost 10000 Å. The data were taken using satellites and various ground based telescopes. The observed light curves are plotted in the top of the left panel in Fig. (1.2), with the top maroon light curve having the longest wavelength. The rest are placed in order of decreasing wavelength, until the purple light curve, which corresponds to the shortest wavelength band.

The object is variable at all wavelengths and at all the sampled timescales. Moreover, we observe that the same variations are present in all the light curves. Therefore, it is interesting to study the correlation between the light curves, both at long and short timescales, and investigate if there are any delays. In Fig.(1.2), in the top of the left panel, the dashed lines show the best-fit parabola fitted to the observed light curves. These parabola fit the data relatively well and are indicative of the long term variations in Fairall9 at the different wavebands. Furthermore, the

black open diamonds found on the parabola show the minimum flux in each light curve. Long term variations are detected first in the longer wavelength bands and propagate to the shorter wavelength bands. We observe that a delay is present among the minimum of the parabola (negative time lags). These delays, with respect to the shortest wavelength, are displayed in the top right panel in Fig.(1.2).

The fast variations on short timescales are studied by removing the parabolic fit from the original data. These curves are shown in the lower part of the left panel in Fig.(1.2). We can quantitatively determine the delay between those curves, using methods that will be described in Chapter 3, to determine the time lags. The bottom right panel in Fig.(1.2) shows the time lags as a function of wavelength. The time lags were determined by using the W2 waveband light curve as reference. The time lag plot shows that variations are detected first in the shorter wavelength bands and propagate to the longer wavelength bands (positive time lags). Lags increasing with wavelength are indicative of X-ray illumination of the disc. The behaviour of the fast variations is typical in AGN, but delays between the slow variations is not (i.e. negative time lags).

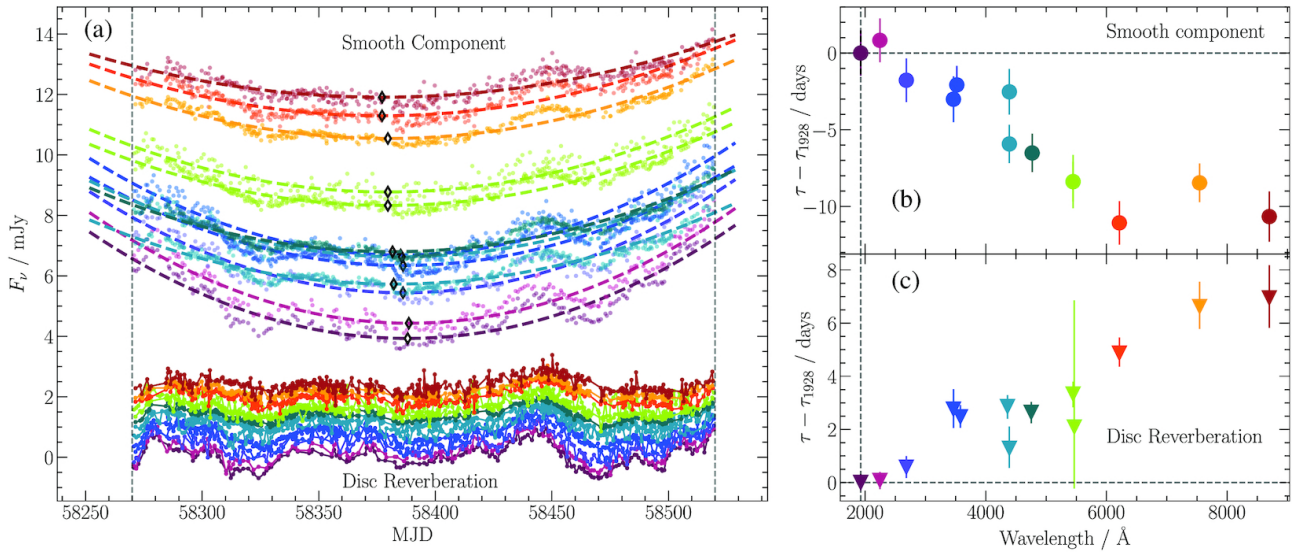


Figure 1.2: *Left panel:* The Fairall9 light curves for wavelengths ranging from  $\sim 1900\text{\AA}$  to  $\sim 10000\text{\AA}$  (upper part, from bot to top with increasing wavelength). The dashed lines are indicative of the best-fit parabola fitted to the observed data. The black open diamonds show the minimum flux of the best-fit parabola. The lines in the lower part of the panel represent the residual light curves, taken from the subtraction of the aforementioned parabola from the original data. *Right panel:* Top panel shows the delays between the long-term variations, with respect to the W2 parabola, as a function of wavelength. The bottom panel shows the time lag spectrum of the fast variations, quantitatively measured by using the W2 as the reference for the computation of the time lags of the other wavebands. Figure taken from Hernandez Santisteban et al (2020).

In this work, our first objective is to quantitatively investigate the phenomena of the negative time lags present in the slow variations of Fairall9. To do that, we are going to determine the longest sampled timescale by applying the discrete Fourier transform to the light curves across all the wavebands. Then, we will calculate the delays between the wavebands in a quantitative way by applying methods that are described below. Our second objective is to repeat the same analysis using data for many AGN, in order to study the behaviour of their slow variations and see whether the presence of negative time lags in the long term variations in AGN is common or not.

# Chapter 2

## Fourier Transform of Light Curves

### 2.1 Definitions

Let us assume a variable process, for example light emission from an AGN. This process will be continuous however, in practice, we observe astronomical objects over a given period of time. This results in a number of observations, at different times, which we call the "light curve" of the observed object. Let us assume a collection of observations,  $x(t)$ , which were performed at equidistant times so that  $t = k \cdot \Delta t$ ,  $k = 0, 1, \dots, N-1$ , and  $\Delta t$  is the time separation between successive observations. We also assume that  $N$  is an even number. The discrete Fourier transform of the light curve (i.e. of the  $N$  data points) is defined as follows

$$X_n = \sum_{k=0}^{N-1} x(k \cdot \Delta t) \cdot e^{\frac{-2\pi i k n}{N}} = \sum_{k=0}^{N-1} x(k \cdot \Delta t) \cdot [\cos(\frac{2\pi k n}{N}) - i \cdot \sin(\frac{2\pi k n}{N})] = U_n - i \cdot V_n, \quad (2.1)$$

where  $n = -\frac{N}{2}, \dots, \frac{N}{2} - 1$ ,  $U_n = \sum_{k=0}^{N-1} x(k \cdot \Delta t) \cdot \cos(\frac{2\pi k n}{N})$  and  $V_n = \sum_{k=0}^{N-1} x(k \cdot \Delta t) \cdot \sin(\frac{2\pi k n}{N})$ .  $X_n$  are the Fourier components of the light curve. They are a complex number, with real and imaginary part  $U_n$  and  $V_n$  respectively. If we apply the inverse discrete Fourier transform, then we can recover the data points  $x(t_k = k \cdot \Delta t)$ , from  $X_n$  as follows:

$$\begin{aligned} x(k \cdot \Delta t) &= \frac{1}{N} \cdot \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} X_n \cdot e^{\frac{2\pi i k n}{N}} = \frac{1}{N} \cdot \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} X_n \cdot e^{\frac{2\pi i n k \Delta t}{N \Delta t}} = \\ &= \frac{1}{N} \cdot \sum_{n=-\frac{N}{2}}^{\frac{N}{2}-1} X_n \cdot e^{2\pi i t_k f_n}, \end{aligned} \quad (2.2)$$

where  $f_n = \frac{n}{N \cdot \Delta t}$ . We show in the Appendix that eq.(2.2) holds. The equation above can be rewritten as follows:

$$x(t_k) = \frac{1}{N} \cdot [X_0 \cdot e^{2\pi i t_k 0} + X_{-\frac{N}{2}} \cdot e^{2\pi i t_k f_{-N/2}} + \sum_{n=-\frac{N}{2}+1}^{-1} X_n \cdot e^{2\pi i t_k f_n} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{2\pi i t_k f_n}]. \quad (2.3)$$

According to eq.(2.1),  $X_{-n} = X_n^*$ , for all  $n$ . Therefore,



$$\begin{aligned}
x(t_k) &= \frac{1}{N} \cdot [X_0 + X_{\frac{N}{2}}^* \cdot e^{-2\pi i t_k f_{N/2}} + \sum_{n=1}^{\frac{N}{2}-1} X_{-n} \cdot e^{-2\pi i t_k f_n} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{2\pi i t_k f_n}] = \\
&= \frac{1}{N} \cdot [X_0 + X_{\frac{N}{2}}^* \cdot e^{-i\pi k} + \sum_{n=1}^{\frac{N}{2}-1} X_{-n} \cdot e^{-2\pi i t_k f_n} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{2\pi i t_k f_n}].
\end{aligned} \tag{2.4}$$

The equation above shows that the light curve can be decomposed into  $\frac{N}{2} + 1$  sinusoids, with frequency  $f_n$ ,  $n = 0, 1, \dots, \frac{N}{2}$ . The first sinusoid,  $S_0(t)$ , is a constant, equal to the light curve mean:

$$S_0(t) = \frac{1}{N} \cdot X_0 = \frac{1}{N} \cdot \sum_{k=0}^{N-1} X(k \cdot \Delta t). \tag{2.5}$$

The sinusoid at frequency  $f_{\frac{N}{2}} = \frac{1}{2\Delta t}$  is equal to:

$$S_{\frac{N}{2}}(t_k) = \frac{1}{N} \cdot X_{\frac{N}{2}}^* \cdot e^{-i\pi k} \tag{2.6}$$

We note that  $\cos(-\pi k) = \cos(\pi k)$  and  $\sin(\pi k) = 0$ , for all  $k$ , while  $V_{n=\frac{N}{2}} = 0$ . Therefore eq.(2.6) becomes:

$$S_{n=\frac{N}{2}} = \frac{1}{N} \cdot [U_{n=\frac{N}{2}} \cdot \cos(\pi k)]. \tag{2.7}$$

The sinusoids at frequencies  $f_n$ , with  $n = 1, 2, \dots, \frac{N}{2} - 1$ , are given by the last two terms in eq.(2.4).

$$\begin{aligned}
S_n(t_k) &= \frac{1}{N} \cdot [\sum_{n=1}^{\frac{N}{2}-1} X_{-n} \cdot e^{-2\pi i t_k f_n} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{2\pi i t_k f_n}] = \\
&= \frac{1}{N} \cdot \left\{ \sum_{n=1}^{\frac{N}{2}-1} (U_n + i \cdot V_n) \cdot [\cos(2\pi f_n t_k) - i \cdot \sin(2\pi f_n t_k)] + \right. \\
&\quad \left. + \sum_{n=1}^{\frac{N}{2}-1} (U_n - i \cdot V_n) \cdot [\cos(2\pi f_n t_k) + i \cdot \sin(2\pi f_n t_k)] \right\} = \\
&= \frac{1}{N} \cdot \left\{ 2 \cdot \sum_{n=1}^{\frac{N}{2}-1} U_n \cdot \cos(2\pi f_n t_k) + 2 \cdot \sum_{n=1}^{\frac{N}{2}-1} V_n \cdot \sin(2\pi f_n t_k) - \right. \\
&\quad \left. - i \cdot \sum_{n=1}^{\frac{N}{2}-1} U_n \cdot \sin(2\pi f_n t_k) + i \cdot \sum_{n=1}^{\frac{N}{2}-1} V_n \cdot \cos(2\pi f_n t_k) + \right. \\
&\quad \left. + i \cdot \sum_{n=1}^{\frac{N}{2}-1} U_n \cdot \sin(2\pi f_n t_k) - i \cdot \sum_{n=1}^{\frac{N}{2}-1} V_n \cdot \cos(2\pi f_n t_k) \right\} \Rightarrow \\
S_n(t_k) &= \frac{2}{N} \cdot \left\{ \sum_{n=1}^{\frac{N}{2}-1} U_n \cdot \cos(2\pi f_n t_k) + \sum_{n=1}^{\frac{N}{2}-1} V_n \cdot \sin(2\pi f_n t_k) \right\}
\end{aligned} \tag{2.8}$$

Equation above shows that each of the  $S_n(t)$  sinusoids is the sum of cosine and sine functions whose amplitude is determined by the real and imaginary parts of  $X_n$ , as determined by eq(2.1).

## 2.2 Fourier Decomposition of Fairall9 Light Curves

As an example, we present in this section the results from the application of the discrete Fourier transform analysis using real data. We will study the light curves of Fairall9, a Seyfert 1 galaxy taken from Hernandez Santisteban et al (2020). Fairall 9 was observed for approximately 10 months, from Modified Julian Date (MJD) 58250 to 58550 with *Swift* satellite and ground based telescopes. The *Swift* Observatory is a satellite that was launched on the 20<sup>th</sup> of November 2004 and carries three telescopes: the X-ray telescope (XRT), the Ultraviolet/Optical Telescope (UVOT) and the Burst Alert Telescope (BAT). In this work we will be using data collected from the UVOT telescope. The telescope has a diameter of 30 cm and a field of view of 17 arcmin<sup>2</sup>. It is equipped with 6 filters, namely UVW2, UVM2, UVW1, U, B and V, with effective wavelengths 1928Å, 2246Å, 2600Å, 3465Å, 4392Å, and 5468Å respectively. Hernandez Santisteban et al also presented observations of Fairall9 taken from the Las Cumbres Observatory (LCO). The LCO is a network of astronomical observatories with 7 sites, all of which are professional astronomical observatories. For the observation of Fairall9, LCO used 7 filters, namely u', B, g', V, r', i' and z', with effective wavelengths 3580Å, 4393Å, 4754Å, 5469Å, 6204Å, 7698Å and 9665Å respectively. Figure (2.1) shows the observed W2 and g-band light curves. We will demonstrate the application of the discrete Fourier transform method using these light curves and we will decompose them into various sinusoids.

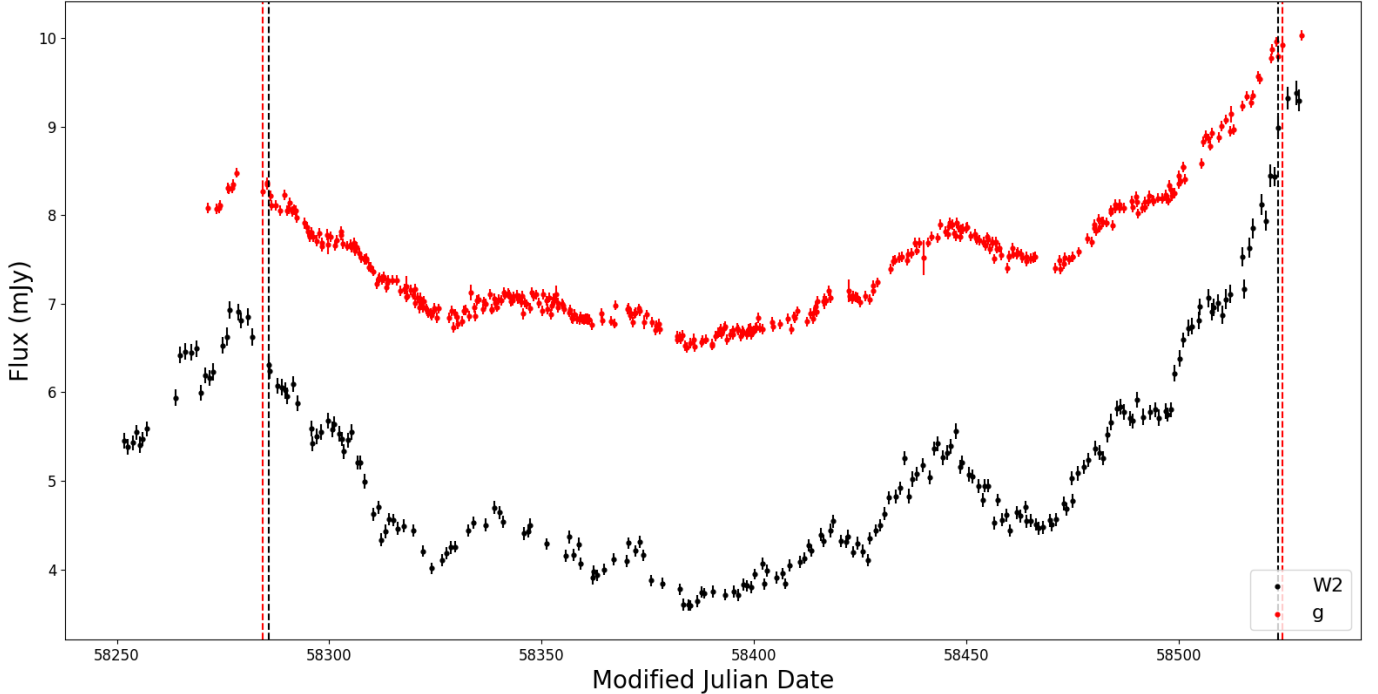


Figure 2.1: *Swift* W2 and LCO *g*-band light curves for Fairall9 (black and red dots, respectively). X-axis units are in "Modified Julian Date". MJD is calculated by subtracting 2,400,000.5 days from the Julian date. Julian date is counted as the number of days since noon on January 1<sup>st</sup>, 4712 BC. The vertical dashed lines indicates the start and finish of the observations we considered.

Fairall9 was observed using the UVOT W2 filter 219 times, from MJD 58251.66 to MJD

58528.27, for a duration of  $\sim 277$  days. It was also observed in the  $g$ -band by the LCO 403 times in  $\sim 258$  days, from MJD 58271.43 to MJD 58529.03. Figure (2.1) shows the resulting W2 and  $g$ -band light curves. In this figure, the vertical dashed line indicates the starting date and finishing date of the light curves that we used in our analysis ( $t_{start,W2} = 58285.736$  MJD,  $t_{start,g} = 58284.394$  MJD,  $t_{finish,W2} = 58523.507$  MJD and  $t_{finish,g} = 58524.426$  MJD). The data before and after the dashed lines were discarded because there were large gaps in the LCO observations, and the cadence of the *Swift* observations was low. After  $t_{start}$  the average time between the *Swift* observations and the LCO observations is  $\Delta t_{W2,obs} = 1.24$  days and  $\Delta t_{g,obs} = 0.62$  days respectively, while the number of points is  $N_{W2,obs} = 197$  and  $N_{g,obs} = 395$  respectively.

## 2.3 Light Curve Interpolation

We are going to use the light curves plotted in Fig.(2.1) to compute the discrete Fourier transformation and hence determine their Fourier components. If we wish to use eq.(2.1), then obviously the light curve must be evenly sampled. This is not the case with the light curves plotted in Fig.(2.1) even after  $t_{start}$ . Using the observed light curves, we need to create evenly sampled ones. Using linear interpolation we can easily achieve this.

Let us assume a continuous function  $y = f(x)$ , and two points  $f(x_1) = (x_1, y_1)$  and  $f(x_2) = (x_2, y_2)$ . Suppose we need to find the values of the function for several points  $x_i$  in between  $x_1$  and  $x_2$ . As a first approximation, we can assume a straight line which connects the two points. Given the  $x_i$ 's, we are able to compute  $y_i$ 's as long as we determine the equation of a straight line that connects  $(x_1, y_1)$  and  $(x_2, y_2)$ . This can be done as follows:

$$y = a + b \cdot x = \begin{cases} y_1 = a + b \cdot x_1 \\ y_2 = a + b \cdot x_2 \end{cases} \quad (2.9)$$

Using eq. (2.9) we get the following equation for  $a$  and  $b$  :

$$a = y_1 - x_1 \cdot \left( \frac{y_1 - y_2}{x_1 - x_2} \right), \quad b = \left( \frac{y_1 - y_2}{x_1 - x_2} \right). \quad (2.10)$$

By substituting  $a$  and  $b$  from eq.(2.10) in eq.(2.9) we finally get :

$$y(x) = y_1 + (x - x_1) \cdot \left( \frac{y_1 - y_2}{x_1 - x_2} \right). \quad (2.11)$$

The values  $y(x)$ , using (2.11), will be close to  $f(x)$ , for any  $x$  between  $x_1$  and  $x_2$  as long as the distance between  $x_1$  and  $x_2$  is small in comparison to the curvature of  $y(x)$ . Below we describe how we created an evenly sampled light curve by applying linear interpolation to the observed light curve.

We set the bin size,  $\Delta t_{even}$ , of the evenly sampled light curves to be equal to the mean time interval between the observations in the light curves ( $\Delta t_{W2,obs}$  and  $\Delta t_{g,obs}$ ). The first point in the evenly sampled light curves will be the actual observation at  $t_{start,W2}$  and  $t_{start,g}$  (these are for example the times indicated by the vertical lines in Fig.(2.1)). The number of points will be equal to  $N_{even}$ , such that  $N_{even} \cdot \Delta t$  is just above the final observation. In our case

$N_{W2,even} = 197$  and  $N_{g,even} = 395$ . We will compute the flux  $f_{even}(t_{even,i})$  using eq.(2.11) and the actual observations that are just before and after  $t_{even,i}$  where  $t_{even,i} = i \cdot \Delta t$ ,  $i = 1, \dots, N_{even}$ . We also add a Gaussian noise to each  $f_{even}(t_{even,i})$  point, by adding a random Gaussian number of mean zero and variance  $\sigma_{PN}^2$  which should be representative of the variance that is introduced to the light curve by the experimental Poisson noise. To achieve this, we computed the mean value of the error squared of the original observations,  $\overline{\sigma^2} = \sum_{i=1}^{N_{obs}} \sigma_i^2 / N_{obs}$ , where  $\sigma_i$  is the error of the flux of each observation, and we set  $\sigma_{PN}^2 = \overline{\sigma^2}$  i.e. we assume that the mean error squared represents the variance introduced by using the observational noise to the light curve. So at each point,  $i \cdot \Delta t_{even}$ , the final flux estimation is equal to  $f_{even}(i \cdot \Delta t_{even}) + N(0, \sigma_{PN}^2)$ , where  $N(0, \sigma_{PN}^2)$  represents the random Gaussian variable with mean zero and variance  $\sigma_{PN}^2$ .

Figure (2.2) shows a part of the observed and the interpolated W2 light curves. The two appear to be in good agreement. As long as there are no major gaps between the observations, we expect the interpolated evenly sampled light curve to be close to the observed one. If this is the case then we expect the number of points of the interpolated and the observed light curve to be approximately the same. Indeed, we get  $N_{W2} = 197$  and  $N_g = 395$  both in the observed and the interpolated light curves. Moreover, the points of the new light curves should be close to the observed light curve points. In our case, the mean time difference between  $i \cdot \Delta t_{even}$  and  $t_{obs,i}$  or  $t_{obs,i+1}$  (depending which is closer to  $i \cdot \Delta t_{even}$ ) is 0.42 days for the W2 band and 0.31 days for the  $g$ -band light curves. The result is notably smaller than the time scales where significant variations appear in the observed light curves, which is of the order of a few days. For this reason, assuming a linear trend between successive points in the observed light curves was a reasonable conclusion. We conclude that, using linear interpolation, results in evenly sampled light curves which are very close to the original ones, for all the sources we are going to study in this work. Figure (2.2) shows a part of the observed and the interpolated W2 light curves.

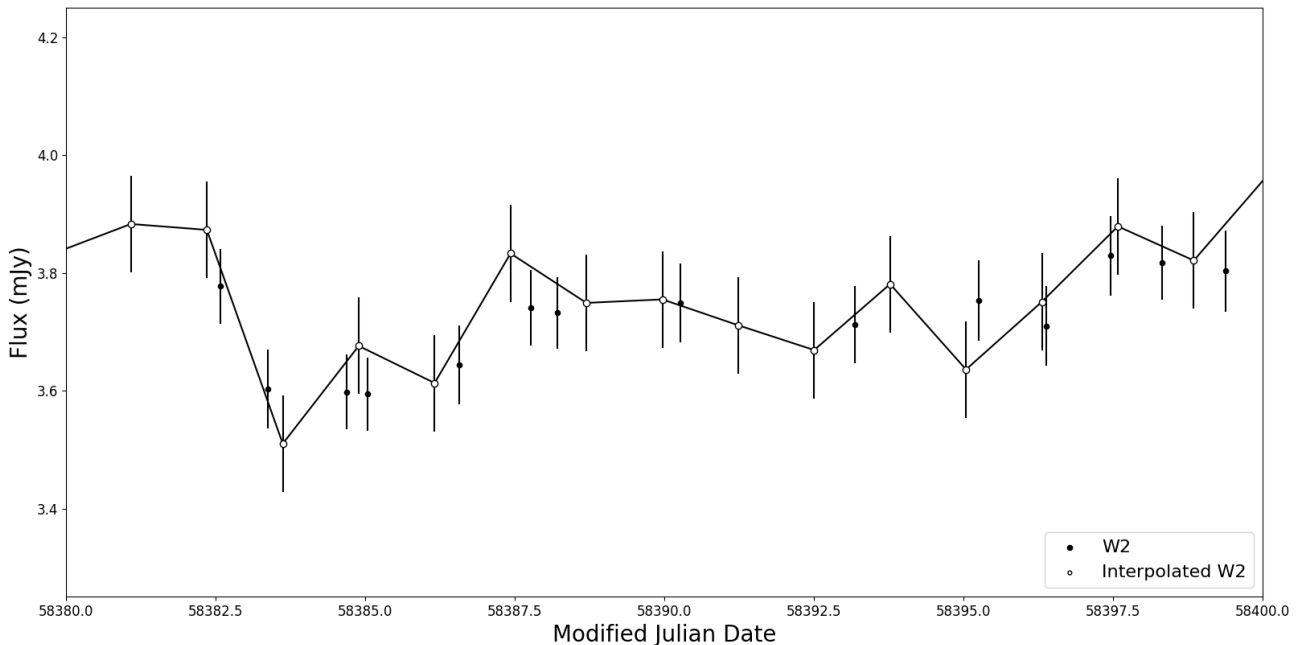


Figure 2.2: Part of the observed and interpolated light curves of the W2 band (filled and open circles respectively). The continuous line connects the interpolated data.

## 2.4 The Longest Sampled Time-Scale

Having applied the interpolation method to the Fairall9 light curves, we are now able to use the discrete Fourier transform in order to decompose them to various components. In order to demonstrate how this works, we are going to present the results for the W2 light curve. Using eq.(2.5), eq.(2.7) and eq.(2.8), we can compute the sinusoids that contribute to the observed variations at any frequency,  $f_n = \frac{n}{N \cdot \Delta t}$ . In Figure (2.3), the filled circles represent the evenly sampled W2 light curve, which we constructed through linear interpolation from the observed light curve, as explained in the preceding section. The solid line in the top of this plot shows  $S_{n=0} + S_{n=1}$ , i.e. the sum of the sinusoids with frequencies  $n = 0$  and  $n = 1$ . The frequency  $f_{n=1} = \frac{1}{N \cdot \Delta t}$  is the longest sampled frequency in the light curve. It corresponds to the longest variation that is fully sampled by the light curve.

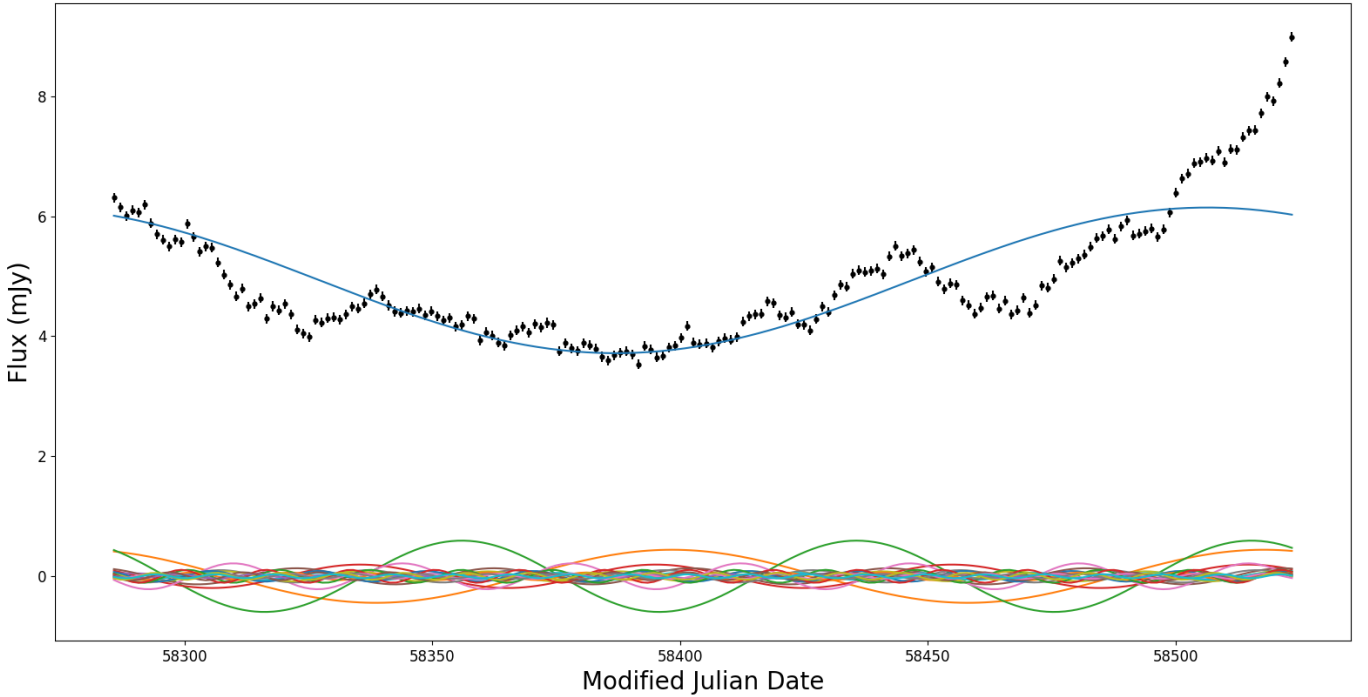


Figure 2.3: The constructed (evenly sampled) W2 light curve, together with the sinusoids for frequencies  $f_{n=0}$  and  $f_{n=1}$  (top line in the plot) and  $f_{n=2, \dots, 20}$  (shown by the lines in the bottom of the plot).

The lower portion of the plot displays sinusoidal curves corresponding to frequencies ranging from  $f_{n=2}$  up to  $f_{n=20}$ . They show the sampled variations (around zero) of frequencies which are increasing from  $\frac{2}{N \cdot \Delta t}$  up to  $\frac{20}{N \cdot \Delta t}$ . As it is shown in the plot, the amplitude of these variations decreases with increasing frequency. It is evident that by summing all the sinusoids,  $S_n(t)$ ,  $n = 0, \dots, \frac{N}{2}$ , we can reconstruct the observed light curve.

# Chapter 3

## Cross Correlation Analysis

### 3.1 Stationary Processes

A random (or stochastic) process is generally a time-varying (or space-varying) quantity,  $X(t)$ , where a precise value for  $X(t)$  at a time  $t$  cannot be determined theoretically. Instead there is a range of possible values with an associated probability distribution describing the relative likeliness of each possible value. Let us assume for example the observation of the solar flux at a given time. If we repeat the observation at a later time we will almost certainly get a different value, despite the fact that the solar flux is constant, even if we use the same equipment and exposure time. This happens because there are a lot of random parameters that affect both the emission of light from the Sun's photosphere and the detection of light from the detector. For example, excited atoms or free electrons have a certain probability of light emission in the Sun's photosphere. Therefore we do not expect exactly the same number of photons emitted in the photosphere per unit time. Furthermore, there is a non-zero probability of scattering or getting absorbed, as light is traveling through the photosphere (as well as the Earth's atmosphere) towards the observer. Therefore we do not expect exactly the same number of solar photons arriving on the detectors surface per unit time. Regarding the detector, there is also a probability associated with the way the light is going to interact with the material. Because of that, for a given number of photons aiming at the detector, we do not expect that the detector will always detect the same percentage of them. Therefore, if we observe the Sun at different periods, many times, we will detect different numbers of photons. The collection of all these possible records is called the "ensemble" and each particular record which results from observations is called a "realization".

A random process is stationary, when the statistical properties of the process (like the mean and variance) do not change over time. Basically, if we take any realization of a stationary process and divide it up into a number of time intervals, the various sections of the realization should look mostly the same.

### 3.2 Cross Correlation Function

Let us assume two stochastic processes  $X_1(t)$ ,  $X_2(t)$  where  $t = 0, \pm 1, \pm 2, \dots$ . We say that  $X_1(t)$ ,  $X_2(t)$  is a stationary bivariate process if their mean ( $\mu_1$ ,  $\mu_2$ ) and variances ( $\sigma_1^2$ ,  $\sigma_2^2$ ) are not functions of time, meaning they do not change if time is shifted. If we want to describe the correlation structure between these two processes, we need to introduce the cross-covariance function ( $CCVF$ ) which is defined as follows :

$$CCVF(\tau) = E\{[X_1(t) - \mu_1] \cdot [X_2(t + \tau) - \mu_2]\}, \quad (3.1)$$

where  $\tau$  is the lag (say the time difference between two observations), and  $E[f(t, \tau)]$  denotes the expectation value of  $f(t, \tau)$ . We can also introduce the cross-correlation function ( $CCF$ ), which is the normalized version of eq.(3.1), given by:

$$CCF(\tau) = \frac{CCVF(\tau)}{\sigma_1 \cdot \sigma_2}. \quad (3.2)$$

Let us now assume two time series  $x_{1,i}$  and  $x_{2,i}$ , which are the result of observing the realization of two processes, namely  $X_1(t)$  and  $X_2(t)$ , at discrete times  $t_i$  ( $i = 1, \dots, N$ ) with equal sampling ( $\Delta t = t_{i+1} - t_i$ ). Then, we can estimate  $CCF(t)$  as follows :

$$\hat{CCF}(\tau_k) = \frac{(\frac{1}{N}) \cdot \sum_{i=1}^{N-k} (x_{1,i} - \bar{x}_1) \cdot (x_{2,i+k} - \bar{x}_2)}{[(\frac{1}{N}) \cdot \sum_{i=1}^N (x_{1,i} - \bar{x}_1)^2]^{1/2} \cdot [(\frac{1}{N}) \cdot \sum_{i=1}^N (x_{2,i} - \bar{x}_2)^2]^{1/2}}, \quad (3.3)$$

where  $\bar{x}_1$ ,  $\bar{x}_2$  are the sampled mean values of  $x_{1,i}$  and  $x_{2,i}$ , and  $\tau_k = k \cdot \Delta t$ ,  $k = 0, 1, \dots, N - 1$ . Using eq.(3.3) we can estimate the cross correlation between two light curves when they are evenly sampled, have the same bin size,  $\Delta t$ , and they have the same number of points. However, in many cases, the aforementioned conditions are not fulfilled. In practice, most of the times, observations result in light curves which are not evenly sampled. Furthermore, light curves in various wave bands may not have the same bin size and the same number of points. A situation like this is shown in Fig.(3.1).

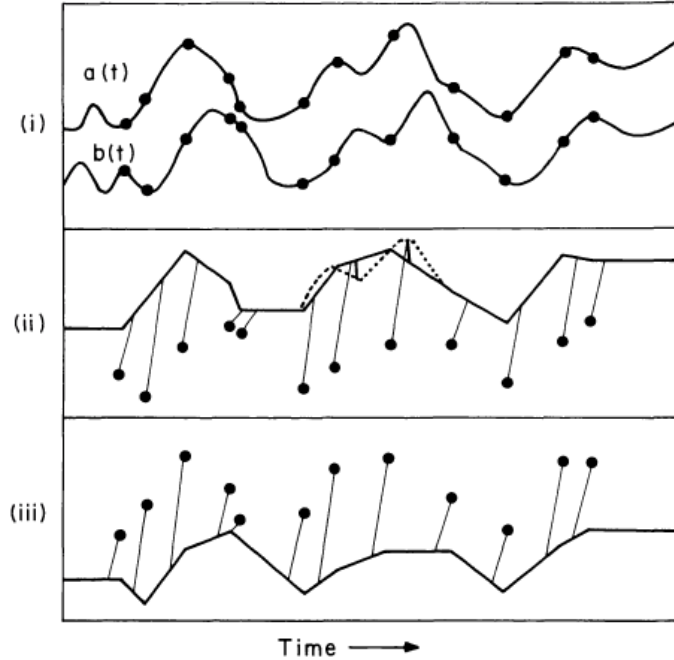


Figure 3.1: *Top*: One realization of two continuous random processes  $a(t)$  and  $b(t)$  (plotted by solid lines), which were observed in 12 different nights (resulting in the observations,  $A(t_i)$  and  $B(t_i)$ , which are shown by the filled dots). We observe that  $b(t)$  is intrinsically delayed relative to  $a(t)$ . *Middle* : Approximation of  $a(t)$ , say  $a_{inter}(t)$  achieved by linear interpolation between  $A(t_i)$ 's. The 12  $B(t_i)$ 's are correlated with the interpolated points of  $a_{inter}(t)$  at times  $\tau$ . *Bottom* : Similarly to the middle panel, the plot shows how the 12  $A(t_i)$ 's are correlated with the interpolated points of  $b_{inter}(t)$  at times  $-\tau$ . Figure taken from C. Martin Gaskell and Bradley M. Peterson (1987).

The continuous lines  $a(t)$  and  $b(t)$  in the top panel of this figure show one realization of two, continuous, random processes. As we can see, the two processes are obviously correlated, but with a delay (process  $b(t)$  is delayed with respect to process  $a(t)$ ). The filled dots indicate 12 different observations of  $a(t)$  and  $b(t)$  (say  $A(t_i)$  and  $B(t_i)$ ) from different nights.

The solid line in the middle panel shows the line that is produced assuming a linear interpolation between observations  $A(t_i)$ . Before and after the first and last observations the interpolated light curve is equal to  $A(t_1)$  and  $A(t_{last})$ . The solid line in the bottom panel shows the results from the linear interpolation between the points in light curve  $B(t_i)$ . Before and after the first and last observations the interpolated light curve is equal to  $B(t_1)$  and  $B(t_{last})$ . Suppose now we want to compute the  $CCF$  between  $a(t)$  and  $b(t)$  at a lag  $\tau$ . We can achieve that by using  $A(t_i)$ ,  $B(t_i)$ , the respective interpolated light curves and eq.(3.3) as follows. Firstly, we can use the  $B(t_i)$  observations, the points  $A(t_i + \tau)$  from the interpolated  $A(t)$  light curve, and eq.(3.3) to compute  $\hat{CCF}_{BA}(\tau)$ . Then, we can also use the  $A(t_i)$  observations, the points  $B(t_i - \tau)$  from the interpolated  $B(t)$  light curve, and eq.(3.3) to compute  $\hat{CCF}_{AB}(-\tau)$ . Gaskell and Peterson (1987) suggested that we should use as a final estimate of the cross correlation between the two light curves,  $\hat{CCF}(\tau)$ , the mean of  $\hat{CCF}_{BA}(\tau)$  and  $\hat{CCF}_{AB}(-\tau)$ .

### 3.3 Application in the case of the Fairall9 longest sampled variations

As an example, we are going to apply the aforementioned methods to compute the time delays between the longest time scale sinusoids in the light curves of AGN at different wavelengths.

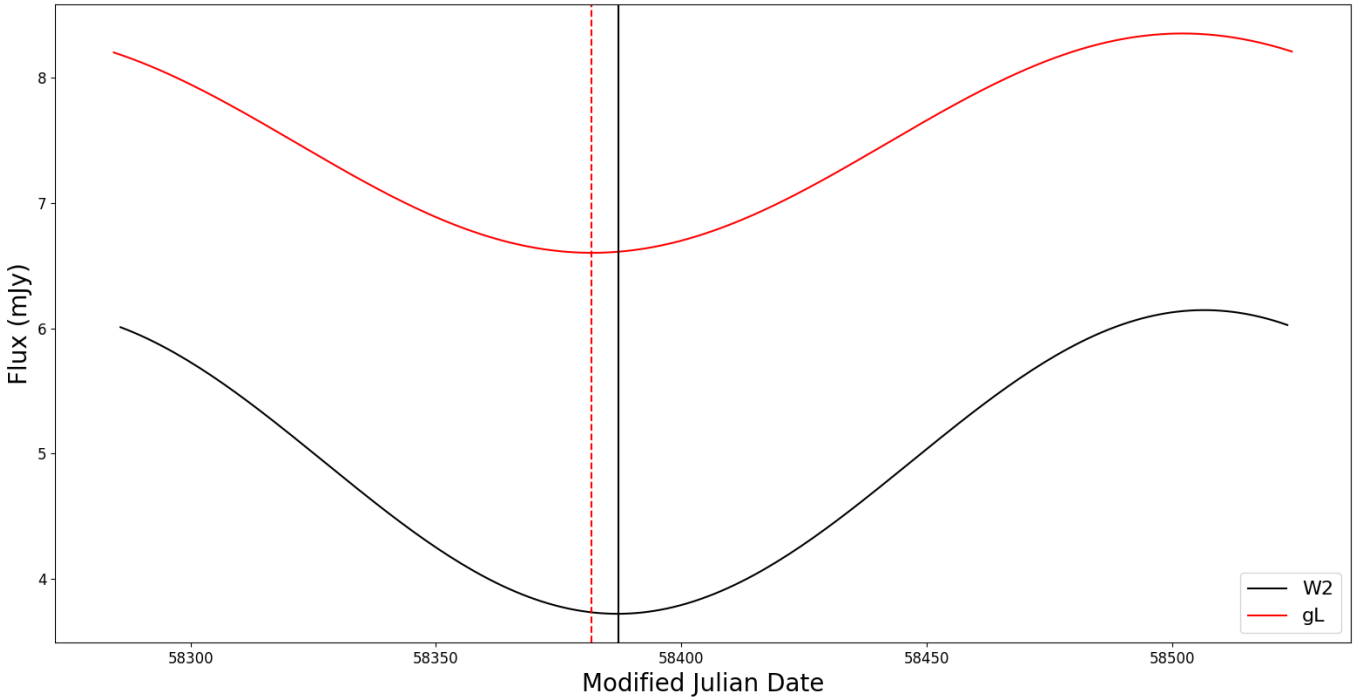


Figure 3.2: The longest variability components ( $n = 1$  and  $n = 0$ ) in the W2 and  $g$ -band light curves of Fairall9.

Figure (3.2) shows the sum of the longest variability components ( $S_{n=0} + S_{n=1}$ ) in the W2 and  $g$ -band light curves of Fairall9. Our objective is to investigate if there are any delays between



these components. The vertical black solid line and the red dashed line show the lowest values of the W2 and  $g$ -band components respectively. It is clear that the minimum flux in the W2 component is delayed with respect to the minimum flux in the  $g$ -band component. If we want to quantify this delay we need to compute the  $CCF$  between the components plotted in Fig.(3.2).

In order to study the correlation between the two variability components shown in Fig.(3.2) we are going to use a python program called “PyCCF”<sup>1</sup>. The code works like the method described in section (3.2) (and Peterson et al (1998)). It uses linear interpolation on two unevenly sampled light curves, it cross correlates the light curves and calculates the peak and the lag centroid of the cross-correlation function. In our case, the sinusoids are evenly sampled by construction. However, the longest variability components shown in Fig.(3.2) do not have the same number of points in both bands, because the original W2 and  $g$ -band light curves did not have the same number of points. That being the case, we need to apply the interpolation step in the PyCCF code, by determining an interpolation bin size. We set the bin size at  $\Delta t_{interp} = 0.5$  days, which is slightly smaller than both  $\Delta t_{W2,obs}$  and  $\Delta t_{g,obs}$  (as suggested by the code instructions).

The program computes the  $CCF(\tau)$  at each lag  $\tau$ , twice, as it is explained in section 3.2. PyCCF later determines the time-lag at which  $CCF(\tau)$  is maximum, say  $CCF_{max}(\tau)$  and then calculates the average centroid time-lag,  $\tau_{cent,obs}$ , as the mean of the time-lags for which  $CCF(\tau) \geq 0.8 \cdot CCF_{max}(\tau)$ . The left panel at Fig.(3.3) shows the  $CCF$  between the longest time scale sinusoids in the W2 and  $g$ -band light curves. The  $CCF$  is defined in such a way so that, when the W2 component precedes the  $g$ -band component,  $CCF_{max}$  will appear at a positive lag. The vertical line indicates  $\tau_{max}$ , i.e. the time lag where the  $CCF$  maximum appears. In this case,  $\tau_{max} = -5$  days. The dashed horizontal line shows all the  $CCF$  values which are 80% of  $CCF_{max}$ . The range that is used to compute  $t_{cent}$  is between the two time lags that the dashed line is intercepted with the  $CCF$  curve.

The code also performs a Monte Carlo experiment to improve the accuracy in determining the time lag between the two components and to determine its corresponding error. It creates  $N$  ( $=10000$  in our case) pairs of W2 and  $g$ -band variability components by randomly sampling the original curves (shown in Fig.(3.2)). For each of these pairs of synthetic components, the program computes a  $t_{cent,synth}$ . In this way, the code creates a distribution of centroid time lags. This distribution is shown in the right panel of Fig.(3.3). For the final estimation of the time lag between the curves shown in Fig.(3.2),  $t_{cent}$ , the code uses the median of the  $t_{cent,synth}$  distribution shown in the right panel of Fig.(3.3).

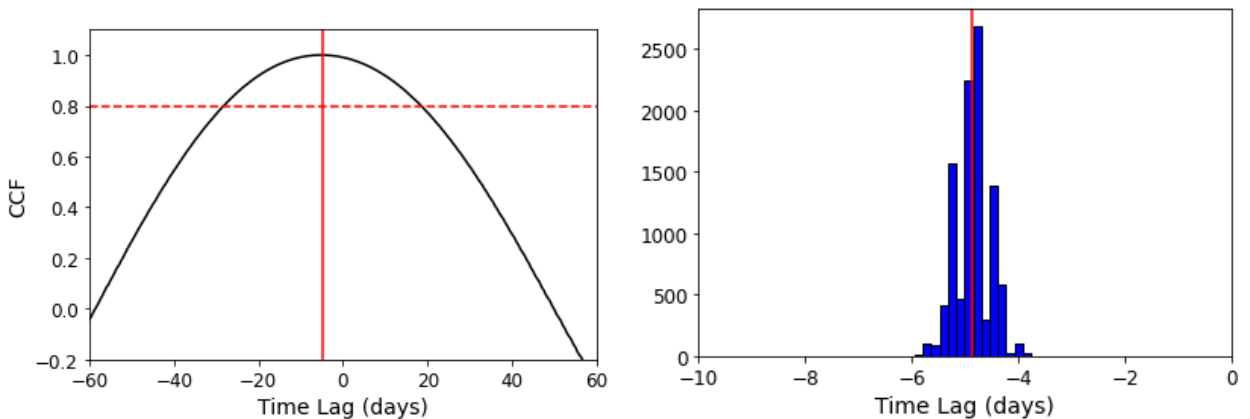


Figure 3.3: *Left panel:* dotted horizontal line indicates the  $0.8 \cdot CCF_{max}$ . Vertical solid line shows the time lag at  $CCF_{max}$ . *Right panel:* Monte Carlo  $t_{cent,synth}$  distribution. Vertical line indicates the median  $t_{cent}$ .

Furthermore, the positive error of the final  $t_{cent}$  is calculated by subtracting the median of the  $t_{cent,synth}$  distribution from their 90<sup>th</sup> percentile score, while the negative error of the final  $t_{cent}$

<sup>1</sup>For details visit: <http://ascl.net/code/v/1868>

is calculated by subtracting the 10<sup>th</sup> percentile score of the Monte Carlo iteration time lags from their median score. Through this process we get the time lag of the *g*-band light curve relative to the W2 band, which will be set as the reference band in all the objects we will be using for our analysis. In this particular example the time lag was found to be: -4.8(+0.3, -0.5) days.

# Chapter 4

## Data Analysis

### 4.1 The sample of AGN and light curves

Our objective in this work is to compute the longest variability component in the light curves of as many AGN as possible across multiple wavebands, to determine the delays between these components and study their dependence on wavelength. Table (4.1) lists the AGN selected for our analysis, ordered by decreasing black hole mass. These objects were chosen due to their extensive observations using both the *Swift* satellite and various ground-based telescopes, spanning a significant period of time across multiple wavelengths. The data are accessible in existing literature and can be used for determining the delays between the longest variability components, as described in the previous Chapters.

The first column of Table (4.1) lists the name of the objects we are going to study, namely Fairall9, Mrk509, NGC5548, NGC4151 and Mrk142. The Fairall9 data were taken from Hernandez Santisteban et al (2020) and the Mrk142 data were taken from Cackett et al (2020). The Mrk509, NGC4151 and NGC5548 data were taken from Edelson et al (2019) and Fausnaugh et al (2016) (NGC5548 data for the *i*-band data). The second column in Table (4.1) lists the redshift and the black hole mass of the objects in our sample. Redshifts are taken from NED<sup>1</sup>. The BH masses for Fairall9, Mrk509 and Mrk142 were taken from Bentz & Katz (2015), while the BH masses for NGC4151 and NGC5548 were taken from Bentz et al (2022) and Horne et al (2021) respectively. The third column lists the waveband, while the fourth column shows the number of the observations in the respective light curve. In the fifth and sixth columns, the date range and the duration,  $T$ , of the light curves are listed respectively, while in the seventh column we list the average time between observations,  $\Delta t_{obs}$ .

Figures (4.1) to (4.5) show the light curves for each source across various wavelengths. The vertical dashed lines in Figs. (4.1), (4.3), and (4.5), indicate the starting point ( $t_{start}$ ) and ending point ( $t_{stop}$ ) of the data used in our analysis. We opted not to utilize the entire duration of the light curves for these objects due to substantial gaps between observations in at least one band, either before  $t_{start}$  or after  $t_{stop}$ . We also wish to keep the duration of the light curves, in the various wavebands, to be roughly equal. The values listed in columns 4 through 7 in Table (4.1) correspond to the light curves that fall within the dashed boundaries seen in Figs. (4.1), (4.3), and (4.5). In the 6<sup>th</sup> column, the numbers in parenthesis indicate the light curve duration in the rest frame of each source, i.e.  $T_{rf} = T/(1+z)$ .

---

<sup>1</sup>The NASA/IPAC Extragalactic Database (NED) is funded by the National Aeronautics and Space Administration and operated by the California Institute of Technology. <https://ned.ipac.caltech.edu/>

Table 4.1: BH mass and light curve properties of the objects we study in this work.

Object	$M_{BH}$ ( $M_{\odot}$ ) and redshift	Band	N	Date Range (MJD)	T (days)	$\Delta t_{obs}$ (days)
Fairall9	$2 \cdot 10^8$ $z = 0.047$	W2	193	58285.736 - 58523.507	237.8(227.1)	1.24
Fairall9		W1	191	58285.734 - 58523.505	237.8	1.26
Fairall9		US	201	58284.936 - 58523.505	238.6	1.2
Fairall9		BL	418	58284.383 - 58524.413	240	0.58
Fairall9		gL	393	58284.394 - 58524.425	240	0.62
Fairall9		VL	418	58284.414 - 58524.416	240	0.58
Fairall9		rL	387	58284.417 - 58524.428	240	0.63
Fairall9		iL	388	58284.420 - 58524.430	240	0.63
Fairall9		zL	375	58284.423 - 58524.432	240	0.65
Mrk509	$1.1 \cdot 10^8$ $z = 0.034$	W2	233	57829.857 - 58102.491	272.6(263.6)	1.17
Mrk509		W1	228	57829.852 - 58102.486	272.6	1.2
Mrk509		U	245	57829.853 - 58102.488	272.6	1.11
Mrk509		B	245	57829.854 - 58102.489	272.6	1.11
Mrk509		V	238	57829.859 - 58102.493	272.6	1.15
NGC5548	$7.7 \cdot 10^7$ $z = 0.017$	W2	258	56705.983 - 56829.920	123.9(121.8)	0.5
NGC5548		W1	261	56705.979 - 56829.918	123.9	0.48
NGC5548		U	267	56705.980 - 56829.918	123.9	0.47
NGC5548		B	271	56705.981 - 56829.919	123.9	0.46
NGC5548		V	263	56705.986 - 56829.922	123.9	0.48
NGC5548		i	137	56705.610 - 56829.409	123.8	0.93
NGC4151	$1.7 \cdot 10^7$ $z = 0.0033$	W2	251	57438.045 - 57507.289	69.2(69.0)	0.28
NGC4151		W1	268	57438.040 - 57507.284	69.2	0.26
NGC4151		U	310	57438.042 - 57507.286	69.2	0.22
NGC4151		B	311	57438.043 - 57507.288	69.2	0.22
NGC4151		V	303	57438.046 - 57507.291	69.2	0.23
Mrk142	$2 \cdot 10^6$ $z = 0.044$	UVW2	131	58486.067 - 58580.231	94.2(90.2)	0.72
Mrk142		UVW1	133	58486.062 - 58580.226	94.2	0.71
Mrk142		U	132	58486.063 - 58580.228	94.2	0.71
Mrk142		B	145	58486.064 - 58580.229	94.2	0.65
Mrk142		g	215	58486.018 - 58580.454	94.4	0.44
Mrk142		r	211	58486.017 - 58580.456	94.4	0.45
Mrk142		i	210	58486.018 - 58580.458	94.4	0.45

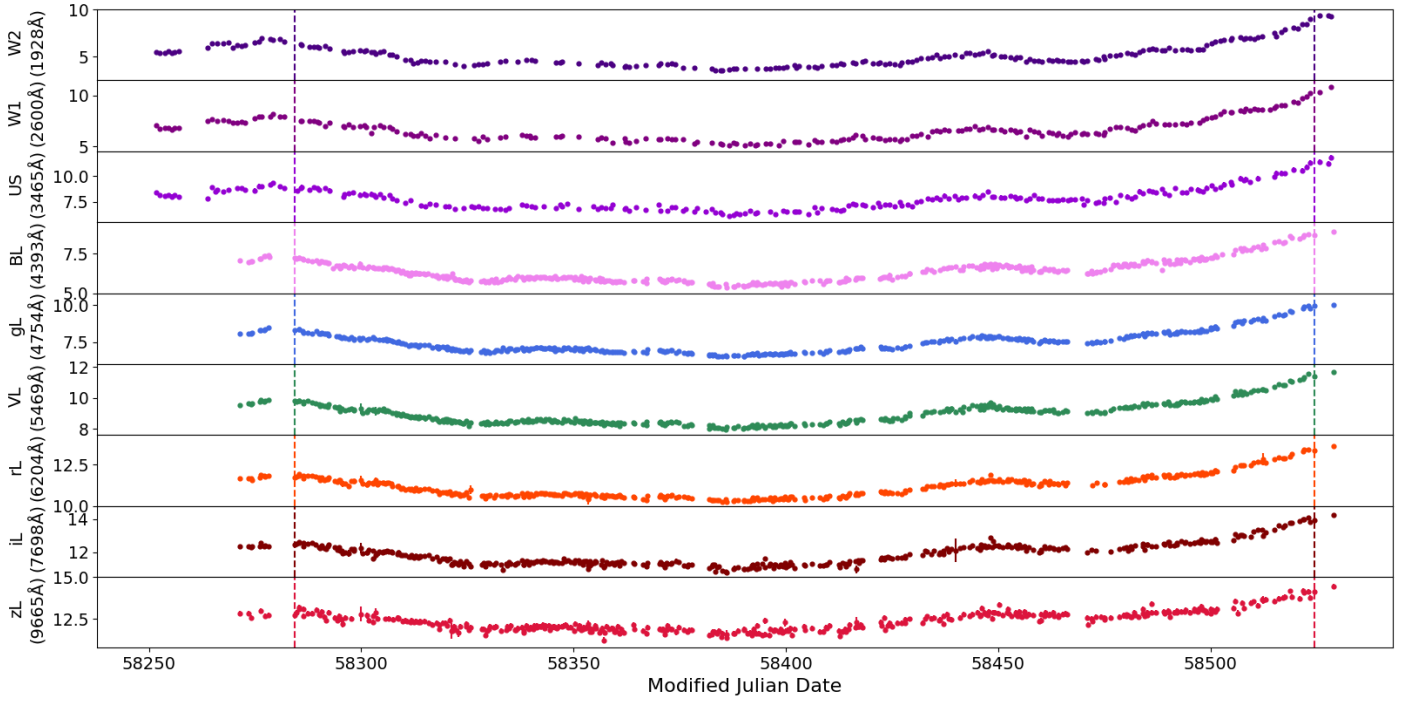


Figure 4.1: The Fairall9 light curves. Vertical dashed lines indicate the start and finish of the data we will be using in our analysis. Light curves are ordered by increasing wavelength from top to bottom.

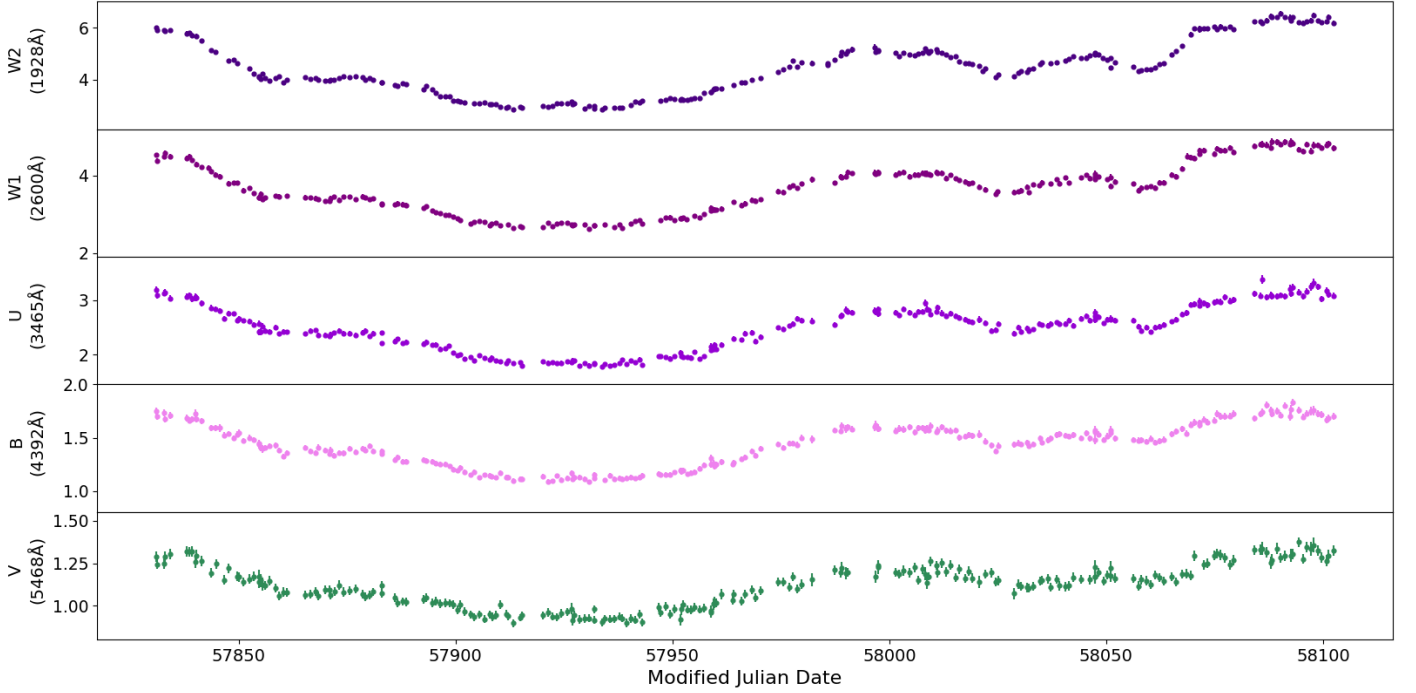


Figure 4.2: Same as in Fig. (4.1), but for Mrk509.

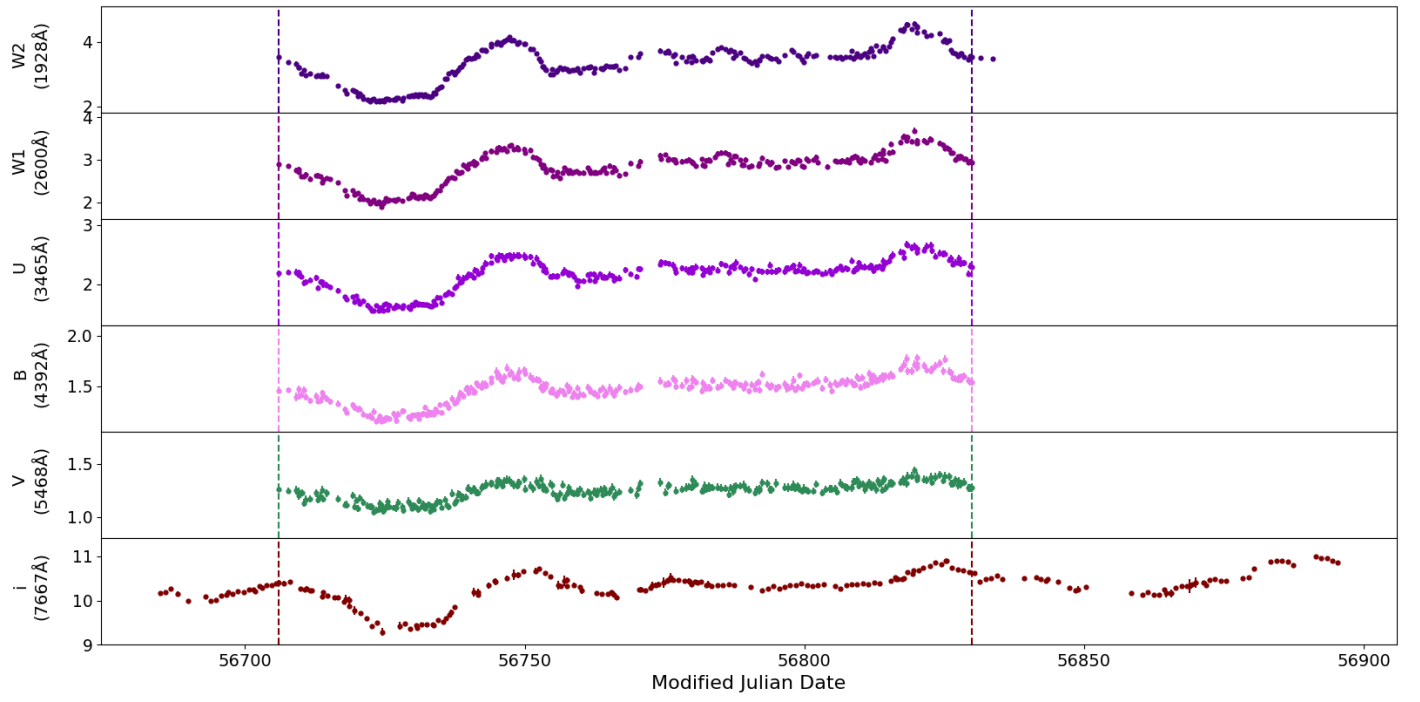


Figure 4.3: Same as in Fig. (4.1), but for NGC5548.

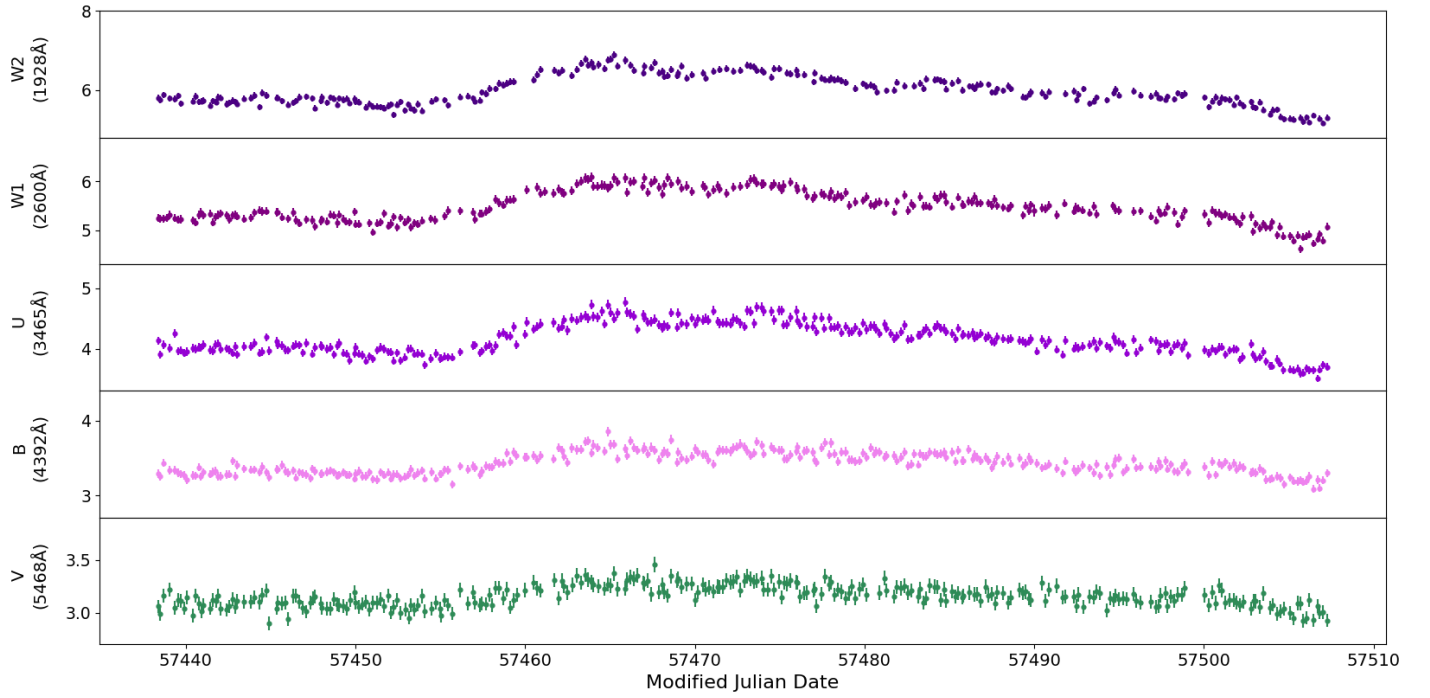


Figure 4.4: Same as in Fig. (4.1), but for NGC4151.

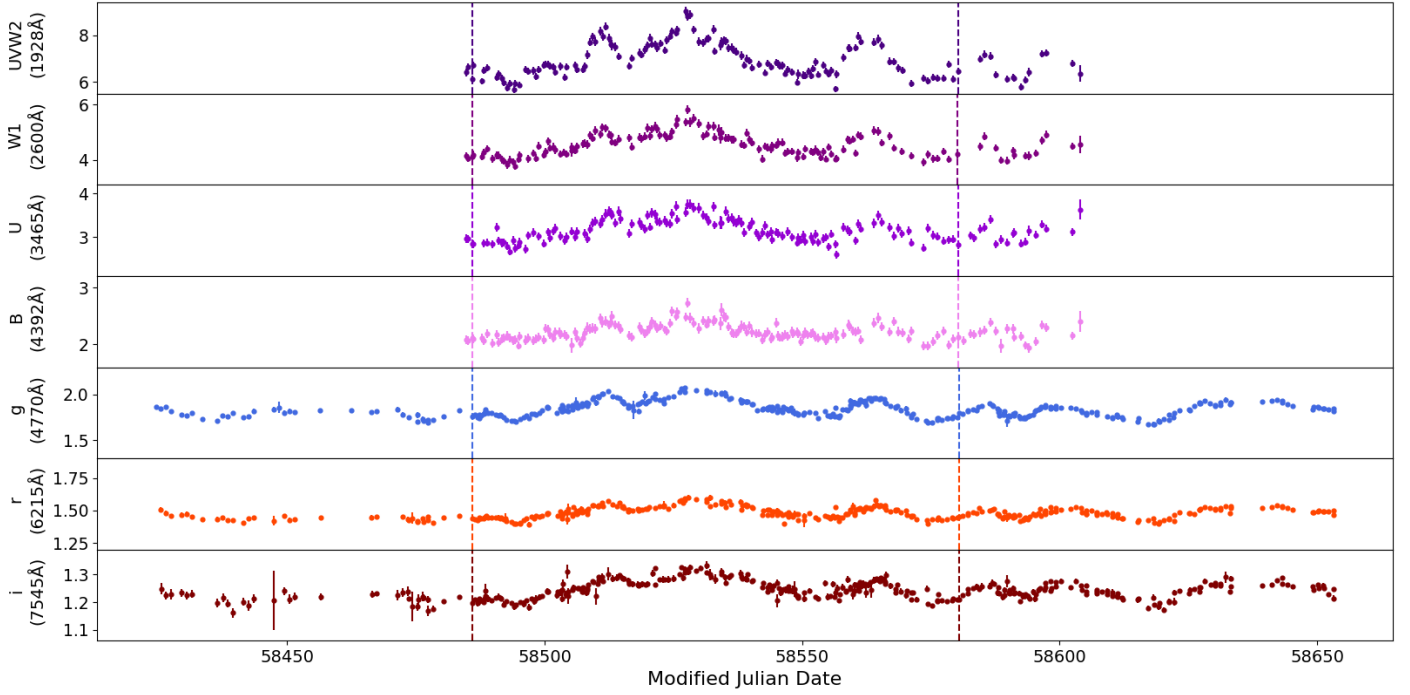


Figure 4.5: Same as in Fig. (4.1), but for Mrk142.

## 4.2 Fourier Decomposition

We used the light curves plotted in the figures of the previous section and we computed their longest variability components. The original light curves are not evenly sampled. In order to apply the discrete Fourier transform method described in Chapter 2, we need to produce evenly sampled light curves using linear interpolation, as we explained in section 2.3. We will then use these evenly sampled light curves to compute the longest variability components in the light curves using eq.(2.8) for  $n = 1$ .

Figures (4.6) through (4.10) show the sum of  $S_{n=0} + S_{n=1}$ , for all the sources listed in Table (4.1). In the case of Mrk509, NGC5548 and NGC4151 we plot the longest variability components at all wavebands. However, for the remaining 2 sources, Fairall9 and Mrk142, we plot them in a limited number of bands which cover the full frequency range of the available wavebands.

The dashed vertical lines indicate the minimum or maximum flux of the longest variability component in the longest wavelength band, located in the bottom of each figure. In the case of Fairall9, Mrk509 and NGC5548, the minimum is well defined and the vertical lines can help us make a first estimation about the delays between the longest variability components at different wavelengths. In the case of NGC4151 and Mrk142 the minimum is not well defined. For that reason, the dashed vertical line indicates the maximum flux value of the component in the longest wavelength band.

In Fairall9, these lines show that the minimum flux appears at later times in the bands with shorter wavelength. This implies that a delay is present between the variations of the variability components, which increases with decreasing wavelength. However, this is not the case for the other sources. The clearest example is Mrk142, where the opposite trend is obvious. There is also a delay between the longest sinusoids in the light curve, but in this particular source, the variations initially appear at shorter wavelengths and then propagate to longer wavelengths.

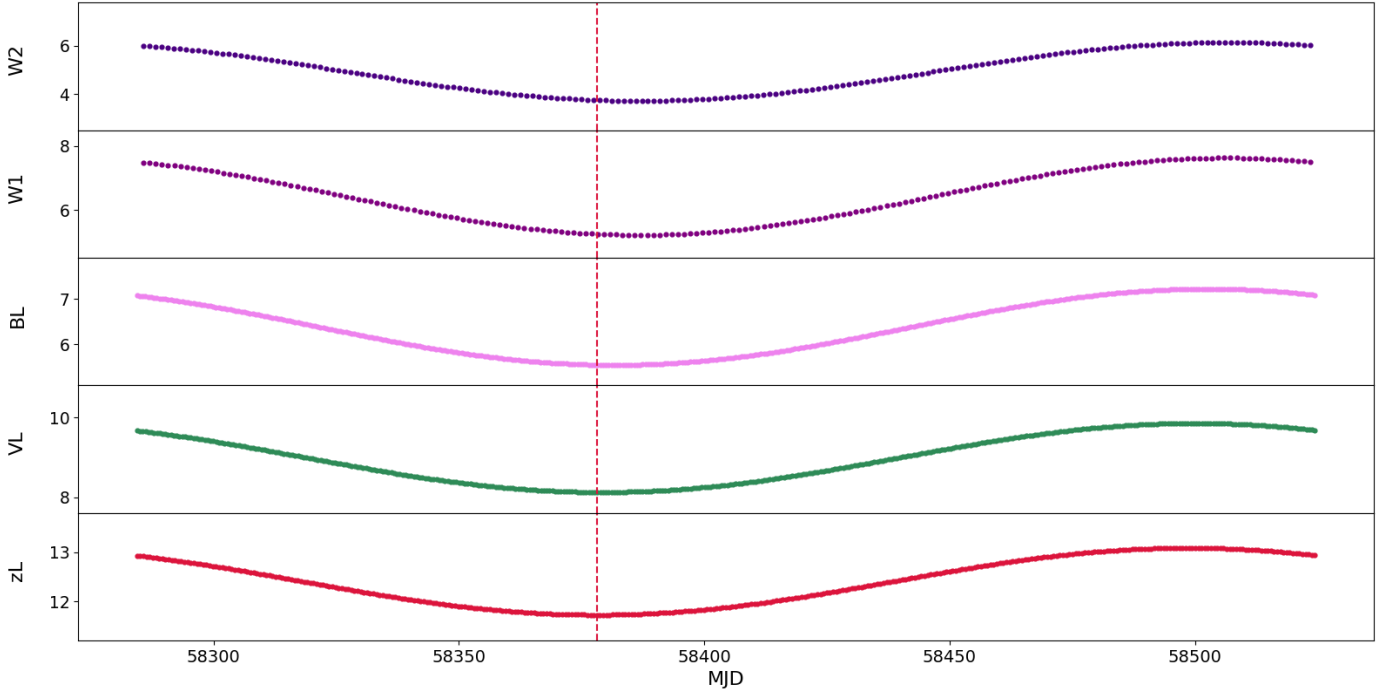


Figure 4.6: The longest variability components (i.e.  $S_{n=0} + S_{n=1}$ ) in the Fairall9 light curves, at various wavelengths. The vertical dashed line indicates the minimum flux value of the variability component in the longest wavelength band ( $z$ -band).

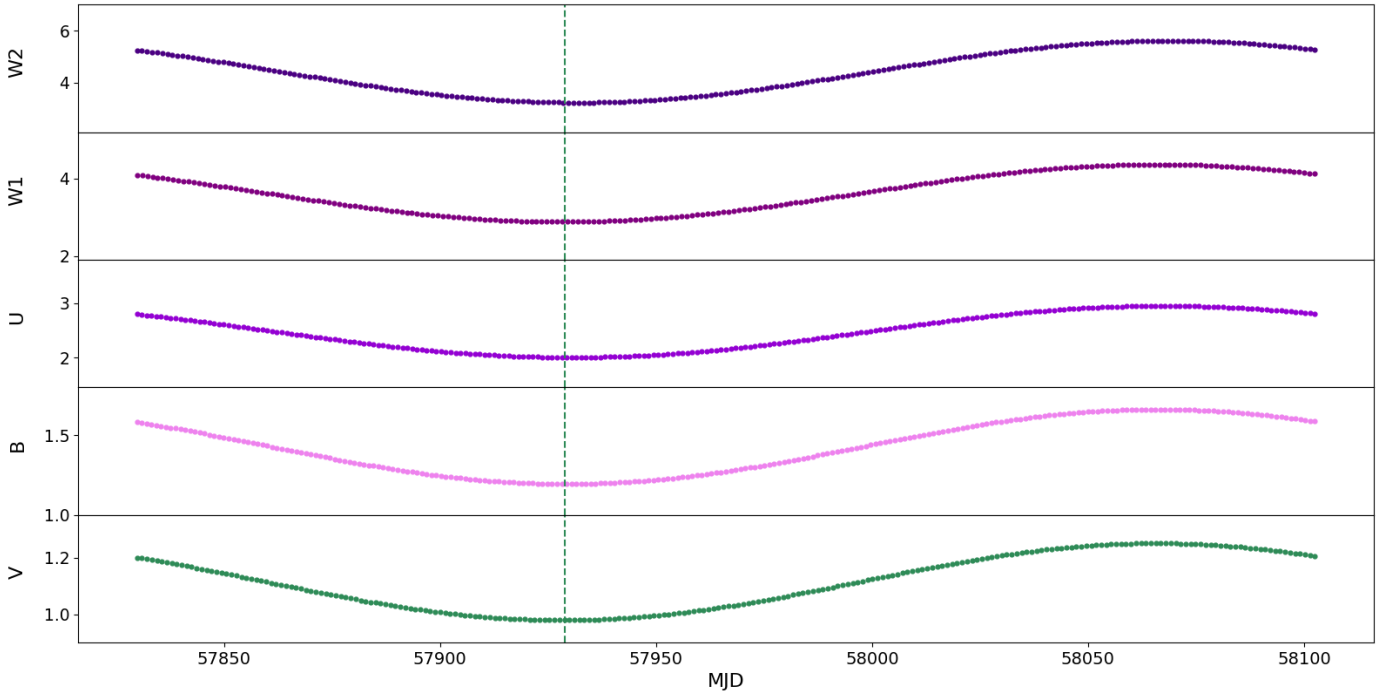


Figure 4.7: Same as in Fig. (4.6), but for Mrk509.



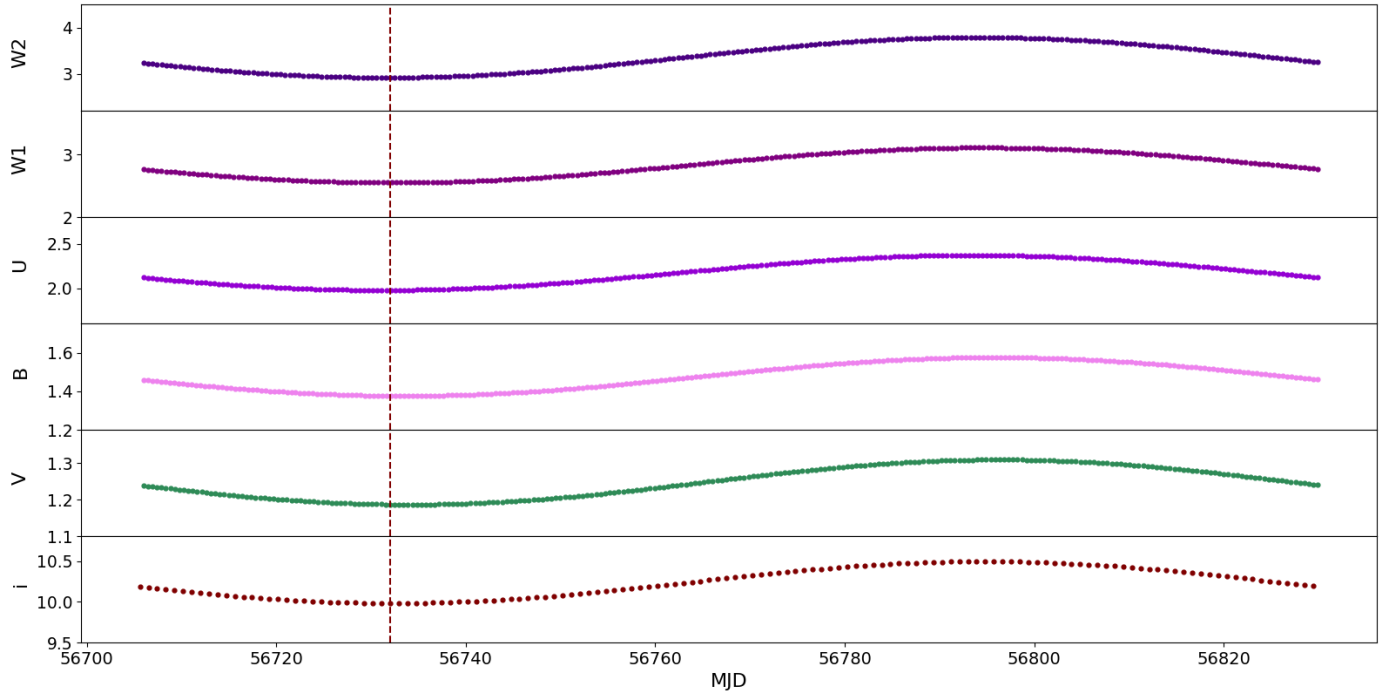


Figure 4.8: Same as in Fig. (4.6), but for NGC5548.

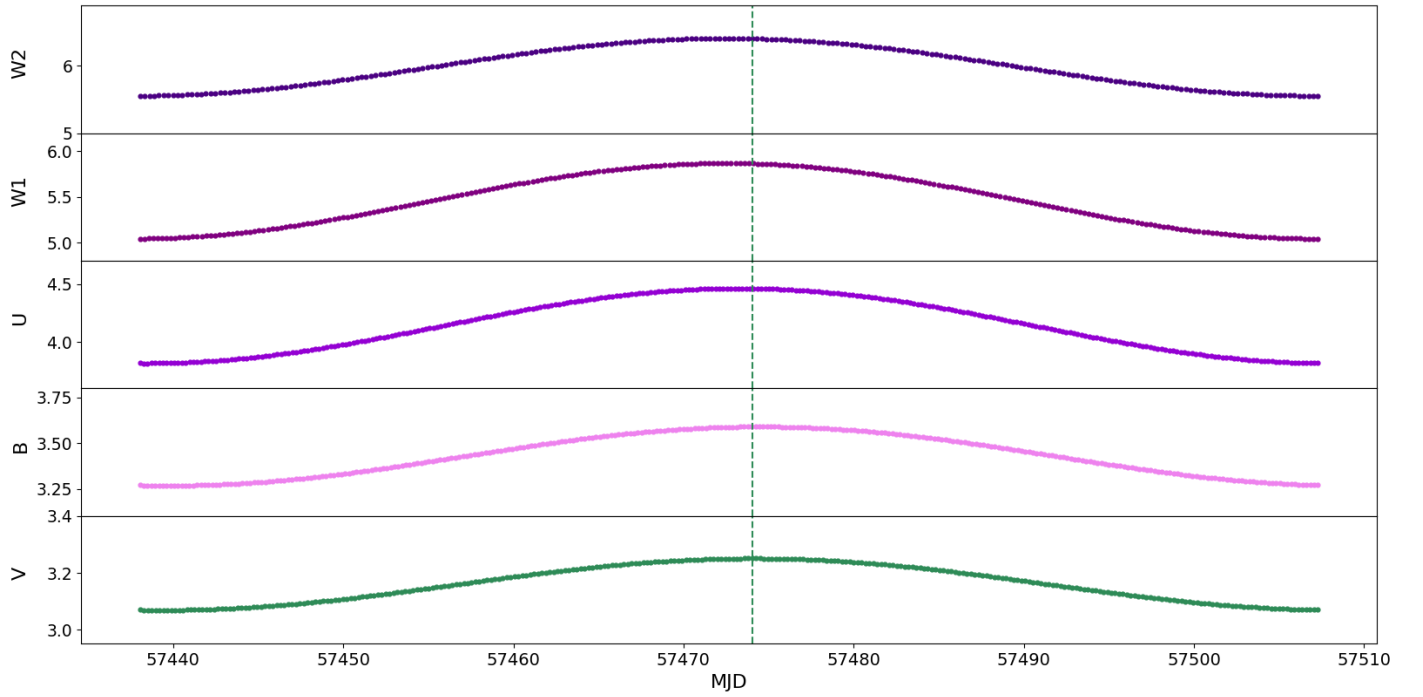


Figure 4.9: Same as in Fig. (4.6), but for NGC4151. Vertical lines in this case indicate the maximum flux of the plotted variability component in the longest wavelength band (V-band)

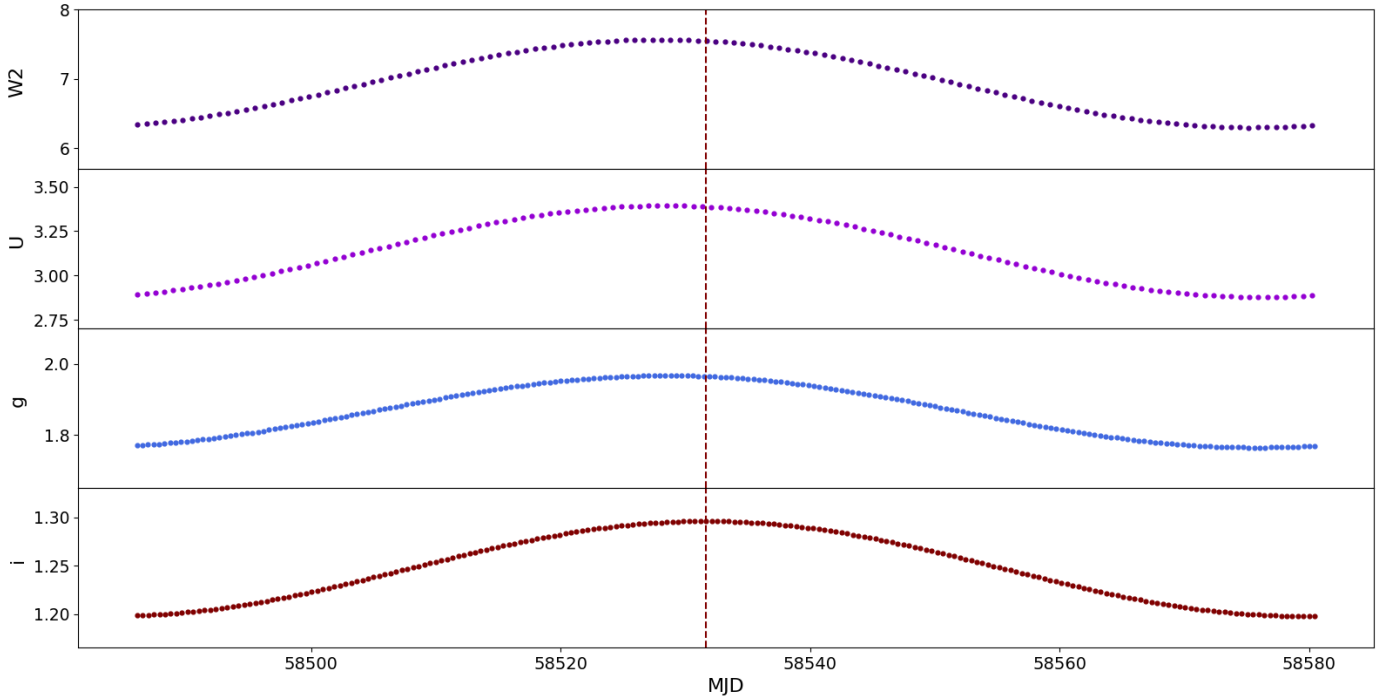


Figure 4.10: Same as in Fig. (4.9), but for Mrk142.

### 4.3 Cross correlation results

In order to quantify the delays between the variability components plotted in Figs. (4.6)-(4.10), we compute their *CCF*, as described below. We use the "PyCCF" code, as discussed in section 3.3, in order to compute the time lags between the longest variability components in the various wavebands, for all sources in the sample. The time lag is computed with respect to the W2 band, which was chosen as the reference band for all sources. Our cross-correlation results are presented in Table (4.2). In the first column of the Table we list the effective wavelength for each waveband. The first and second column of each source list the  $\tau_{cent}$  of the *CCF* between the longest variability component in that band and in W2, and the quantity  $V_{amp} = \frac{F_{max} - F_{min}}{F_{mean}}$ , where  $F_{max}$ ,  $F_{min}$  and  $F_{mean}$  are the maximum, minimum and mean flux of the longest variability component. This quantity does not have units and is a measure of the variability amplitude, in percentage, around the mean flux of each component. It is evident that the variability amplitude of each component is decreasing as the wavelength increases in all sources.

Figures (4.11)-(4.14) show a plot of  $\tau_{cent}$  versus  $\lambda$ . In all these plots, the horizontal line is drawn to indicate  $\tau_{cent} = 0$ . We keep the  $x$ -axis scale the same in all plots to emphasize that Fairall9 and Mrk142 have the widest wavelength coverage. It is clear that in Fairall9, variations in the longest component initially propagate from the longest wavelength to shorter wavelengths. The delays are of the order  $\sim$ days between W1 and W2, and  $\sim$ 8 days between the  $z$ -band and W2. However, this trend is observed only in Fairall9. In NGC5548, NGC4151 and Mrk142, the delays increase with increasing wavelength. The situation in Mrk509 is a bit unclear. The delays appear to be negative, but roughly constant at all wavelengths.

Nevertheless, the comparison drawn from the delays-versus-wavelength plots shown in Figs.(4.11)-(4.13) may be misleading. More specifically, the rest frame frequency of the longest variability

component in the light curve of Fairall9,  $f_{1,rf}$ , is not the same in all the sources. Infact,  $f_{1,rf,F9} = 4.4 \cdot 10^{-3} d^{-1}$  is lower than  $f_{1,rf}$  in NGC5548, NGC4151 and Mrk142 ( $f_{1,rf,NGC5548} = 8.2 \cdot 10^{-3} d^{-1}$ ,  $f_{1,rf,NGC4151} = 1.4 \cdot 10^{-2} d^{-1}$  and  $f_{1,rf,Mrk142} = 1.1 \cdot 10^{-2} d^{-1}$ ). However, frequencies in AGN scale inversely proportional to BH mass. Hence, for a meaningful comparison between delays detected in different sources, we must examine delays for variability components with frequencies similar to  $f_{1,rf,F9}$ , taking into account the differences in BH masses between the sources. The ratio of the Fairall9 BH mass over the BH mass of the other sources in the sample is :  $\sim 1.8$ ,  $\sim 2.6$ ,  $\sim 11.8$  and  $100$ , for Mrk509, NGC5548, NGC4151 and Mrk142 respectively. This implies that we should compare delays between variability components whose frequency is  $f_{comp,Mrk509} \sim 1.8f_{1,rf,F9}$ ,  $f_{comp,NGC5548} \sim 2.6f_{1,rf,F9}$ ,  $f_{comp,NGC4151} \sim 11.8f_{1,rf,F9}$  and  $f_{comp,Mrk142} = 100f_{1,rf,F9}$ . These frequencies correspond to the variability components with  $n = 2$ ,  $n = 1$ ,  $n = 3$ , and  $n = 40$  for these objects, respectively. We computed the delays between those components in the various wavebands and W2. The results are listed in Table (4.2), with the numbers in parenthesis, in the first column of each source, except Mrk142. In the case of Mrk142, the  $n = 40$  component has a low amplitude and is dominated by variations caused by the Poisson noise process. For this reason, we are unable to determine the delays between the intrinsic variability components with  $n = 40$  in this particular source.

Figure (4.14) shows the  $f_{1,rf}$  time lags for Fairall9, as well as the  $f_{comp}$  time lags for Mrk509, NGC5548 and NGC4151. Clearly, when we compare the delays between components with equivalent frequencies in the various sources we do not detect the same trend. Contrary to Fairall9, the time lags in the other three sources appear to increase with longer wavelengths.

Table 4.2: Time lag results for Fairall9, Mrk509, NGC5548, NGC4151 and Mrk142

$\lambda$ (Å)	Fairall9		Mrk509		NGC5548		NGC4151		Mrk142	
	$\tau_{cent}$ (Days)	$V_{amp}$	$\tau_{cent}$ (Days)	$V_{amp}$	$\tau_{cent}$ (Days)	$V_{amp}$	$\tau_{cent}$ (Days)	$V_{amp}$	$\tau_{cent}$ (Days)	$V_{amp}$
1928		0.49		0.54		0.26		0.14		0.18
2600	$-0.2 \pm 0.5$	0.38	$-3.0^{+0.5}_{-0.9}$ (2.0 $\pm$ 0.5)	0.41	$0.6 \pm 0.4$	0.20	$0.2 \pm 0.1$ (1.5 $\pm$ 0.1)	0.15	$0.0^{+0.4}_{-0.2}$	0.19
3465	$-3.7 \pm 0.5$	0.32	$-1.0^{+0.5}_{-0.9}$ (6.0 $\pm$ 0.4)	0.38	$-0.4^{+0.2}_{-0.4}$	0.18	$0.9^{+0.2}_{-0.1}$ (1.2 $\pm$ 0.1)	0.15	$0.2^{+0.2}_{-0.4}$	0.16
4393	$-4.8^{+0.3}_{-0.5}$	0.26	$-2.5^{+0.9}_{-0.5}$ (6.5 $\pm$ 0.4)	0.33	$2.2 \pm 0.4$	0.13	$2.3 \pm 0.1$ (-1.1 $\pm$ 0.1)	0.09	$3.4 \pm 0.2$	0.10
4754	$-4.8^{+0.3}_{-0.5}$	0.23	-	-	-	-	-	-	$0.4^{+0.2}_{-0.4}$	0.11
5469	$-7.3 \pm 0.5$	0.19	$-3.5^{+0.9}_{-0.5}$ (8.0 $\pm$ 0.4)	0.24	$2.6 \pm 0.4$	0.10	$1.4 \pm 0.10$ (1.1 $\pm$ 0.1)	0.06	-	-
6204	$-6.0^{+0.5}_{-0.3}$	0.15	-	-	-	-	-	-	$1.4 \pm 0.2$	0.07
7698	$-11.2^{+0.5}_{-0.3}$	0.13	-	-	$1 \pm 0.4$	0.05	-	-	$3.2 \pm 0.2$	0.08
9665	$-8.3^{+0.3}_{-0.5}$	0.11	-	-	-	-	-	-	-	-

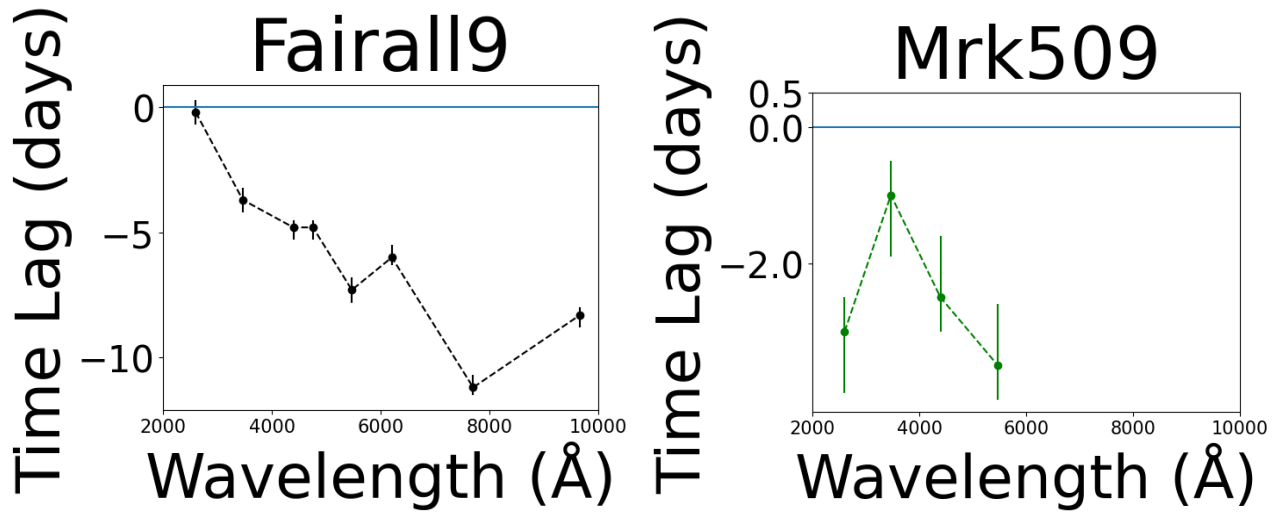


Figure 4.11: *Left panel:* The  $\tau_{cent}$  versus  $\lambda$  plot for Fairall9. *Right panel:* Same as in the left panel, but for Mrk509.

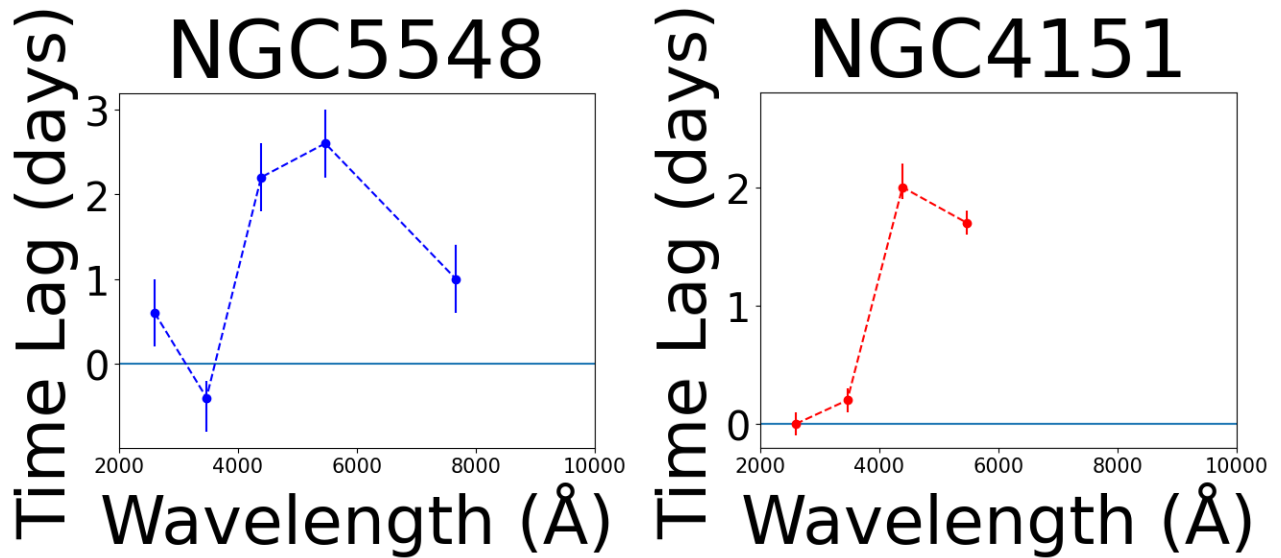


Figure 4.12: Same as in Fig.(4.11), but for NGC5548 and NGC4151.

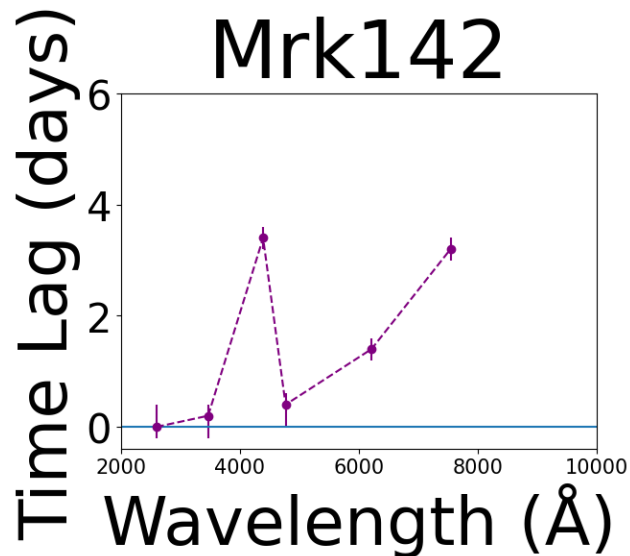


Figure 4.13: Same as in Fig.(4.11), but for Mrk142.

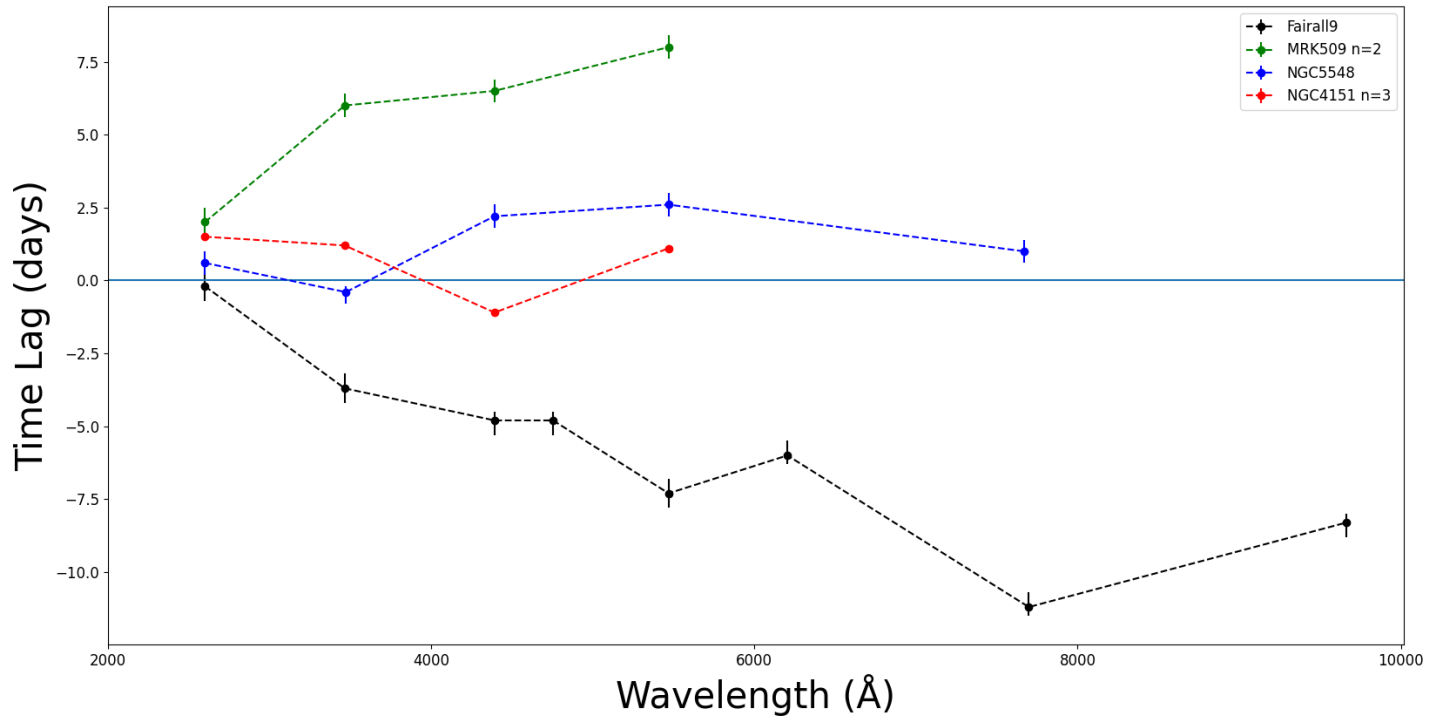


Figure 4.14: Same as in Fig.(4.11), but for all sources (except Mrk142), where  $\tau_{cent}$  is the delay between variations in variability components with the same frequency, normalized to BH mass, which is equal to  $f_{1,rf,F9}$ .

# Chapter 5

## Summary and Conclusions

Our main objective in this work was to compute the longest variability component in the light curves of as many AGN as possible, across multiple wavebands and determine the delays between these components, while studying their dependence on wavelength. In order to achieve this, we used archival data of 5 sources, namely Fairall9, Mrk509, NGC5548, NGC4151 and Mrk142, that were available in the literature. These objects have been observed regularly by both the *Swift* satellite and various ground based telescopes, over a long period of time, across multiple wavebands. Therefore, the available light curves can be used in our analysis, by applying the discrete Fourier transform and calculating the delays between the longest variability components by cross-correlating them. To accomplish this, we followed the steps summarized below.

1) We created evenly sampled light curves for the aforementioned objects in all the wavebands, by using linear interpolation, as described in Chapter 2. We also added a random number to the new, evenly sampled light curves, in order to further simulate the observational noise of a real data set.

2) We performed a Fourier analysis, using the discrete Fourier transform, and determined the longest time-scale Fourier components in each waveband, for all the sources in our sample. To create these components, we used the new, evenly sampled light curves and applied the discrete Fourier transform and then the inverse discrete Fourier transform, but only for the longest sampled frequency, that corresponds to  $n = 1$  as described in detail in Chapter 2. The longest variability components for each source are plotted in Chapter 4, in Figs.(4.6)-(4.10).

3) We cross-correlated the longest variability component in the W2 waveband with all the other wavebands and determined the delays between them, as explained in Chapter 3. The resulting time lags are listed in Table(4.2) and the delay-vs-wavelength figures are shown in Figs.(4.11)-(4.13). We plotted the delays-vs-wavelength and found that for the longest variability component of Fairall9, the results are in agreement with Hernandez Santisteban et al (2020).

4) We computed the maximum variability amplitude of the long variability components. The results are listed in Table (4.2) in the second column of each source. The variability amplitude of these components decreases as wavelength increases in all sources.

5) We computed the delays for the components of the sources with the same rest frame frequency, when normalized by the BH mass. We should normalize frequencies to  $M_{BH}$ , if indeed frequencies in AGN scale inversely proportional to BH mass. The results are listed in Table (4.2) in the second column of each source, inside the parenthesis, and are plotted in Fig.(4.14). We observe that the results do not change after comparing the delays between the W2 waveband and the other wavebands in the same timescale by normalizing their rest frame frequencies. Fairall9 is the only source that exhibits variations that propagate from longer to shorter wavelengths.

Hernandez Santisteban has proposed that the long term time lags in Fairall9 may be due to accretion rate fluctuations in the accretion rate that travel inward from the outer accretion disc toward the inner region. In fact, for a given BH mass, the accretion rate determines the temperature and, consequently, the luminosity of the disc. Therefore, if  $\dot{m}$  varies, at a radius  $R$ , the locally emitted flux from the disc at that radius should also exhibit variations. If the  $\dot{m}$  variations propagate inwards, then variations of the same frequency should appear at all radii, but with a time delay. This delay would be equal to the time it takes for the  $\dot{m}$  variations to travel from the outer radius  $R_{out}$  to the inner radius  $R_{in}$ .

If  $\dot{m}$  variations that propagate through the disc are responsible for the variations that we observe in Fairall9, then we would expect to observe similar behaviour in other AGN. However, when we study the delays between components with the same time-scales by normalizing to the BH mass of each source, we do not detect timelags that behave like the ones in Fairall9. Moreover, the case of Mrk142 is quite important. When we normalize to BH mass, the length of Mrk142's light curves are  $\sim 40$  times longer than that for Fairall9, meaning that the latter's delay behaviour is not due to it being observed for a relatively longer period of time. In Mrk142, as well as the other 3 sources, we see delays towards longer wavelengths, which is what we would expect in the case of X-ray illuminated discs.

Nevertheless, in order to adequately investigate the case of Fairall9, it would seem that we need new data, in order to determine whether this is a transient phenomena or permanent. If it were transient, it could be due to a pressure variation that propagates inwards the accretion disk, caused by either an orbiting star or some other object that affects the accretion disk.

# References

- Bentz, Misty C and Sarah Katz (2015). *PASP* 127, 67.
- Bentz, Misty C, Peter R Williams, and Tommaso Treu (2022). *ApJ* 934, 168.
- Cackett, E. M. et al. (2020). *ApJ* 896, 1.
- Czerny, B. (2006). *ASPC* 360, 265.
- Edelson, R. et al. (2019). *ApJ* 870, 123.
- Fausnaugh, M. M. et al. (2016). *ApJ* 821, 56.
- Gaskell, C. M. and B. M. Peterson (1987). *ApJS* 65, 1.
- Hernández Santisteban, J. V. et al. (2020). *MNRAS* 498.4, 5399.
- Horne, Keith et al. (2021). *ApJ* 907, 76.
- Koratkar, Anuradha and Omer Blaes (1999). *PASP* 111, 1.
- NASA/IPAC Extragalactic Database (NED). URL: <https://ned.ipac.caltech.edu/>.
- Peterson, Bradley M et al. (1998). *PASP* 110, 660.
- Shakura, N. I. and R. A. Sunyaev (1973). *A&A* 24, 337.
- Sun, Mouyuan, CJ Grier, and BM Peterson (2018). *Astrophysics Source Code Library*, 1805.  
URL: <http://ascl.net/code/v/1868>.
- Zheng, W. et al. (1997). *ApJ* 475, 469.



# Appendix

$$\begin{aligned}
x(\kappa \cdot \Delta t) &= \frac{1}{N} \cdot \sum_{n=0}^{N-1} X_n \cdot e^{\frac{2\pi i k n}{N}} = \frac{1}{N} \cdot \sum_{n=-\frac{N}{2}+1}^{\frac{N}{2}} X_n \cdot e^{\frac{2\pi i k n}{N}} = \\
&= \frac{1}{N} \cdot [X_0 \cdot e^{\frac{2\pi i k 0}{N}} + X_{\frac{N}{2}} \cdot e^{\frac{2\pi i k \frac{N}{2}}{N}} + \sum_{n=-\frac{N}{2}+1}^{-1} X_n \cdot e^{\frac{2\pi i k n}{N}} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{\frac{2\pi i k n}{N}}] = \\
&= \frac{1}{N} \cdot [X_0 + X_{\frac{N}{2}} \cdot e^{\pi i k} + \sum_{n=1}^{\frac{N}{2}-1} X_{-n} \cdot e^{\frac{-2\pi i k n}{N}} + \sum_{n=1}^{\frac{N}{2}-1} X_n \cdot e^{\frac{2\pi i k n}{N}}].
\end{aligned} \tag{5.1}$$

We will substitute  $X_n$  using eq.(2.1) with new parameter  $k'$

$$\begin{aligned}
x(\kappa \cdot \Delta t) &= \frac{1}{N} \cdot \left[ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot e^{\frac{-2\pi i k' 0}{N}} + \left( \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot e^{\frac{-2\pi i k' \frac{N}{2}}{N}} \right) \cdot e^{-\pi i k} + \right. \\
&+ \sum_{n=1}^{\frac{N}{2}-1} \left( \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot e^{\frac{2\pi i k' n}{N}} \right) \cdot e^{\frac{-2\pi i k n}{N}} + \sum_{n=1}^{\frac{N}{2}-1} \left( \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot e^{\frac{-2\pi i k' n}{N}} \right) \cdot e^{\frac{2\pi i k n}{N}} \Big] = \\
&= \frac{1}{N} \cdot \left[ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) + \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot e^{\pi i (k-k')} + \right. \\
&+ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \sum_{n=1}^{\frac{N}{2}-1} e^{\frac{-2\pi i (k-k') n}{N}} + \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \sum_{n=1}^{\frac{N}{2}-1} e^{\frac{2\pi i (k-k') n}{N}} \Big] = \\
&= \frac{1}{N} \cdot \left[ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) + \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot \cos(\pi i (k-k')) + \right. \\
&+ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \left[ \left( \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i (k-k') n}{N}} \right) - 1 \right] + \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \left[ \left( \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i (k-k') n}{N}} \right) - 1 \right] \Big] = \\
&= \frac{1}{N} \cdot \left[ - \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) + \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot \cos(\pi i (k-k')) + \right. \\
&+ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \left( \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i (k-k') n}{N}} + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i (k-k') n}{N}} \right) \Big] = \\
&= \frac{1}{N} \cdot \left[ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot (-1 + \cos(\pi i (k-k'))) + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i (k-k') n}{N}} + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i (k-k') n}{N}} \right].
\end{aligned} \tag{5.2}$$

We observe that the last term in the equation above has two finite geometric sums and can be written as follows,

$$\sum_{n=0}^{\frac{N}{2}-1} [e^{\frac{\pm 2\pi i(k-k')}{N}}]^n = \frac{1 - [e^{\frac{\pm 2\pi i(k-k')}{N}}]^{\frac{N}{2}}}{1 - e^{\frac{\pm 2\pi i(k-k')}{N}}}. \quad (5.3)$$

For  $k' \neq k$  and using eq.(5.3), the last term in eq.(5.2) becomes

$$\begin{aligned} & \frac{1 - [e^{\frac{-2\pi i(k-k')}{N}}]^{\frac{N}{2}}}{1 - e^{\frac{-2\pi i(k-k')}{N}}} + \frac{1 - [e^{\frac{2\pi i(k-k')}{N}}]^{\frac{N}{2}}}{1 - e^{\frac{2\pi i(k-k')}{N}}} = \\ &= \frac{(1 - e^{-\pi i(k-k')}) \cdot (1 - e^{\frac{2\pi i(k-k')}{N}}) + (1 - e^{\pi i(k-k')}) \cdot (1 - e^{\frac{-2\pi i(k-k')}{N}})}{(1 - e^{\frac{-2\pi i(k-k')}{N}}) \cdot (1 - e^{\frac{2\pi i(k-k')}{N}})} = \\ &= \frac{1 - \cos(\frac{2\pi(k-k')}{N}) - \cos(\pi(k-k')) + \cos(\pi(k-k')) \cdot \cos(\frac{2\pi(k-k')}{N})}{1 - \cos(\frac{2\pi(k-k')}{N})} = \\ &= 1 - \cos(\pi(k-k')). \end{aligned}$$

Now if we use the above we get that

$$\begin{aligned} & -1 + \cos(\pi(k-k')) + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i(k-k')n}{N}} + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i(k-k')n}{N}} = \\ &= -1 + \cos(\pi(k-k')) + 1 - \cos(\pi(k-k')) = 0. \end{aligned}$$

For  $k' = k$  we get that :

$$\begin{aligned} & -1 + \cos(\pi i(k-k)) + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i(k-k)n}{N}} + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i(k-k)n}{N}} = \\ &= -1 + 1 + \frac{N}{2} + \frac{N}{2} = N. \end{aligned}$$

So we conclude that

$$-1 + \cos(\pi i(k-k')) + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i(k-k')n}{N}} + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i(k-k')n}{N}} = \begin{cases} 0, & \text{for } k' \neq k \\ N, & \text{for } k' = k \end{cases} = \delta_{k',k}. \quad (5.4)$$

Using eq.(5.4), we can rewrite eq.(5.2) as follows

$$\begin{aligned} x(\kappa \cdot \Delta t) &= \frac{1}{N} \cdot \left[ \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot (-1 + \cos(\pi i(k-k'))) + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{-2\pi i(k-k')n}{N}} + \right. \\ &\quad \left. + \sum_{n=0}^{\frac{N}{2}-1} e^{\frac{2\pi i(k-k')n}{N}} \right] = \sum_{k'=0}^{N-1} x(\kappa' \cdot \Delta t) \cdot \delta_{k',k} = x(\kappa \cdot \Delta t). \end{aligned} \quad (5.5)$$