Determination of the Hubble constant from the strong lensing system B0218+357

Dissertation
zur
Erlangung des Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Haniyeh Mahmoudian
aus
Shiraz, Iran

Bonn 2013

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter:	Dr. Olaf Wucknitz
2. Gutachter:	Prof. Dr. Peter Schneider
Tag der Promotion:	08.05.2013
Erscheinungsjahr:	2013

Abstract

Gravitational lensing and particularly strong lensing provide variety of subjects to study such as galaxy evolution, substructure detection in galaxies or cluster of galaxies. Also strong lensing enables us to derive the Hubble constant. The advantage of strong lensing to find this cosmological parameter is its measurement on cosmic scale.

The strong lensing system B0218+357 is an isolated system and thus not significantly influenced by external shear. In addition, the redshift information of the lens and source and time delay between the images are available. These parameters make B0218+357 a promising system to determine the Hubble constant. Although radio observations provide accurate measurements of the positions of the double image, previous attempts on these data to derive the Hubble constant were not satisfactory because the lens galaxy cannot be observed in the radio.

The most important observational effect in optical astronomy is the spreading of the light ray distribution of the source reaching the CCD detector due to the design of the instrument and the atmosphere, the so called PSF. The information lost in images of the CCD cameras can be recovered with dithering. In this technique, the exposures are taken with sub-pixel shifts in order to extract structures on scales smaller than a pixel. The standard method to combine such exposures, Drizzle, reverses the shifts and rotations between the exposures and corrects for the geometric distortion caused by the instrument design. Then this method averages these exposures to produce the combined image. This produces good results but is not optimal for cases in which preservation of the PSF is required.

In B0218+357, with the small separation between the two images, precise subtraction of PSF is essential. Therefore, the combination process of the data should not add additional PSF to the system. In this work, an alternative method based on direct fitting with a least-squares approach is developed to combine the exposures of this system taken by the ACS/WFC detector of the HST. To have a unique solution, a smoothing constraint is also included in this method. This method has the ability of working with arbitrary rotations, shifts and dither pattern. The correction for the geometric distortion and flagging pixels affected by cosmic rays are included. To have a higher resolution, the pixels of the output result of this method has smaller size in comparison to the original exposures taken from the Hubble Space Telescope.

To find the relative positions of the images with respect to the lens galaxy, the combined image of each observed visit of B0218+357 is fitted with two PSF components and a Sersic

profile with additional parameters for spiral arms. With those positions and the assumption of an isothermal model for the lens potential, we are able to determine the Hubble constant to be $70 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with fitting the spiral arms and $66 \pm 4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the case of leaving the arms unfitted. The error bars presented here are rms scatters between the Hubble constant values derived from each visit. There might be additional systematic errors as well.

In the optical data of B0218+357, one of the images (image A) suffers from extinction due to a giant molecular cloud in the line of sight which causes a systematic shift of its positions. To have more precise results, in the next step we use the positions of image A from radio observations. In this approach, the obtained value for the Hubble constant for the case not fitting the spiral arms changes to $76 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and when we fit the spiral arms we derived the value of $79 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the Hubble constant. These results are consistent with previous results but rules out others.

To Mami, Babi, Anna, Azi and Eli

Acknowledgements

First and foremost, I have my sincere gratitudes for Dr. Olaf Wucknitz for his guidance, motivations and patience. I thank Dr. Richard Porcas for his insightful comments and nice teasing. I'm thankful to Prof. Dr. Schneider for his helpful comments. My words can not appreciate Dr. Filomena Volino, Zeinab, Zahra, Sahar and Maryam whom were besides me in this journey both in good and bad times and accepted me exactly as who I am. I thank Dr. Dominique Sluse for being a nice officemate and a colleague. I would like to thank Prof. Dr. Mansouri and Dr. Sadooghi for their tolerance in listening to my long grumble in my trip back home. I'm grateful for the chance of studying in AIfA and experiencing the life in Bonn with all the problems that I faced and to be part of the enjoyable informal lens seminars.

Last but not least, I would like to express my sincere gratitude to my mother and my family whom without their continuous support I could not finish this dissertation.

Contents

1

Introduction

The first step in modern cosmology as it is known today was made by Einstein while he introduced Special and General Relativity in 1905 and 1915. By introducing General Relativity as a new theory for gravity (which in the weak field limit leads to Newtonian gravity), he opened a new window in our knowledge about the universe.

Einstein (1916) completed his calculation and published his first results on General Relativity. At that time, the accepted model for describing the universe emphasized that the universe is static and without any dynamic. To allow his solution to be static, Einstein introduced a constant parameter there, the so-called '*cosmological constant*' Λ.

Shortly after that, Schwarzschild (1916) introduced one of the first solutions of Einstein's field equations. He assumed a perfectly spherical, stationary star with mass M surrounded by vacuum (empty space) and solved the field equations to find the space-time curvature in the exterior of the star. His solution is

$$\mathrm{d}s^2 = \left(1 - \frac{2G\,M}{c^2\,r}\right) c^2\,\mathrm{d}t^2 - \left(\frac{1}{1 - \frac{2G\,M}{c^2\,r}}\right) \mathrm{d}r^2 - r^2\,(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2). \tag{1.1}$$

This metric is a space-time with singularity at the centre of the mass.

Later, Willem de Sitter (1917) solved Einstein's field equations for a universe without matter. He assumed the universe is made of a cosmological constant acting like a repulsive force Λ. His result is a spatially flat universe which expands exponentially due to the repulsive force. But a decade had to pass till the expansion of universe was accepted.

Edwin Hubble (1929) published a paper about the dynamics of the universe. By observing galaxies, Hubble discovered that they are radially receding from us with a velocity proportional to their distance. His measurements led to a value of $500\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the expansion rate

of the universe (recent observations provide the value between 65 to 80 km s^{-1} Mpc^{-1} for the expansion rate or the Hubble constant). This discovery was predicted separately by Lemaitre (1927) and Friedmann (1922), from their solution of Einstein's field equations. Their model also predicted that the early universe was in a very dense state which is now called the Big Bang.

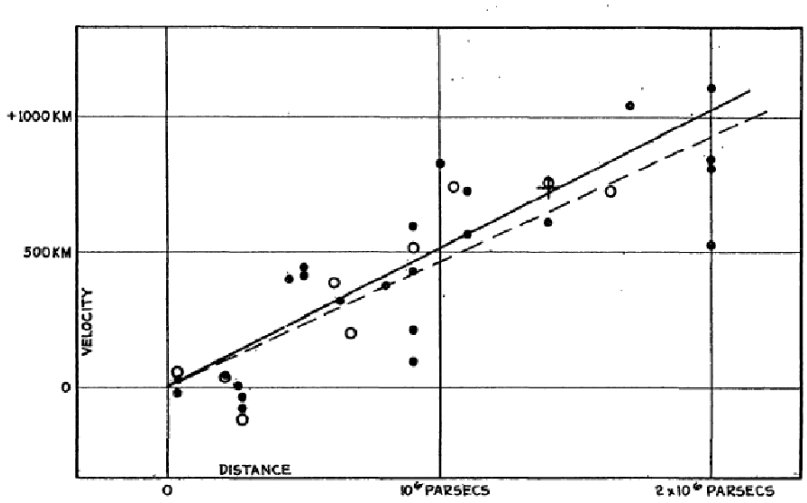

Figure 1.1: Distance relation among Extra-Galactic Nebulae. In this plot Hubble presented the radial velocities versus radius (Hubble, 1929).

Later on, with galaxy spectroscopy in the Coma cluster, Zwicky (1933) found that the galaxies have higher velocity dispersion than the escape velocity from the centre of the cluster. He concluded that there should be '*unseen*' matter in the cluster which later has been named as '*Dark Matter*'.

Independently, H.P. Robertson (1935) and A.G. Walker (1936) introduced a metric for a dynamic, homogeneous and isotropic universe.

F. Hoyle (1948), H. Bondi and T. Gold (1948) introduced the steady-state model for the universe which obeys an idealized cosmological principle (also known as '*the perfect cosmological principle*'). This principle implies that not only every point in the space is representative of the universe as a whole but each point in time represents the entire history of universe. In other words this model emphasizes that time should also be homogeneous.

From General Relativity, curved space-time around a galaxy by the action of its gravitational field can provide multiple images of a background source, arcs or distortion in the shapes

of the images of the background sources. According to the different paths light rays take, observer sees the images of the background source with time delays.

From the Einstein theory of deflection of light, in the presence of matter one can determine the mass distribution of the lens within the Einstein radius[1] (a typical maximum distance between the images produced by a gravitational lensing effect is in the order of the Einstein radius). With more detailed modelling and additional observables such as flux ratios, this can be done precisely in strong lensing. In the weak lensing regime the mass distribution can be estimated for larger radii.

Refsdal (1964a) showed that by having a strong gravitational lensing system with multiple images of a nearby source, one is able to determine the Hubble constant using positions of the images with respect to the position of the lens and the time delay between the images. Walsh et al. (1979) discovered the first multiple imaging systems in strong lensing. With improvements on the observational instruments, more systems have been discovered.

The strong lensing system JVAS CLASS B0218+357 was discovered by Patnaik et al. (1992). This double image system has the smallest image separation known among the strong lensing systems. Follow-up observations measured the redshift of the source, lens and the time delay between the images. These measurements and the isolation of this system give us an opportunity to determine the Hubble constant. Unlike other measurements of the Hubble constant on local scales, strong lensing obtains this parameter on cosmic scale.

Because of the small separation between the images, an accurate Point Spread Function (PSF) subtraction is required and to extract more information from data, images are combined. York et al. (2005) worked on B0218+357 data observed with the WFC camera of the ACS instrument of the Hubble Space Telescope (HST). They combined the images with the Drizzle algorithm (Multidrizzle software) and due to the imprecise PSF subtraction, the robustness of their result is limited.

The aim of this project is to determine the Hubble constant from B0218+357. To improve the results of previous work on the HST data, in this work an alternative method is developed to combine the data. The approach of this method is the direct fit of the data. The model of the brightness distribution of the sky is fitted to the observed data using the least-squares method and to have a unique solution, a smoothness constraint was added. Then, two PSF

[1]When a symmetric lens is located directly on the line of sight between source and observer, instead of images of the source, the observer sees a ring which is called Einstein ring and the radius of this ring is called Einstein radius.

and a spiral lens galaxy are fitted to the system to find the position of the images and the lens galaxy. With these positions and consideration of an isothermal profile for lensing potential, the Hubble constant is determined.

2

Modern Cosmology

In the late 19th and early 20th century, the fundamentals of modern cosmology were built. After Einstein introduced General Relativity in 1915, as mentioned in the introduction, there were several attempts on introducing a new model to explain the universe. Among all these which included the cosmological principle, the Friedmann-Robertson-Walker (FRW) model was the most successful one which could fit the observations[1]. In this chapter an introduction to the basic mathematical structure of our modern cosmological model based on Einstein's theory of gravity is provided. This chapter is based on some sections of the book written by Coles et al. (2002).

2.1 Cosmological Principle

In the 20th century (and still today), with limited knowledge about the distribution of matter in the universe, solving Einstein's field equations was too hard for arbitrary distribution of matter. Therefore, Einstein introduced the cosmological principle for simplicity of the solutions. This principle asserts that, on average, over sufficiently large scales, the universe is homogeneous and isotropic. The perfect principle as mentioned before is based on isotropy in all direction and times together with Copernican principle. This principle led to steady-state cosmology.

There are some models based on an inhomogeneity. Lemaitre (1933) and Tolman (1934)

[1]The observations can be named as CMB spectrum which was first measured by FIRAS instrument on the COBE satellite, Hubble expansion, the abundance of light elements (Big Bang nucleosynthesis), the large scale distribution and apparent evolution of galaxies which are predicted to occur due to gravitational growth of structure in the standard theory. These are sometimes called 'the four pillars of the Big Bang theory'.

and later Bondi (1947) investigated an inhomogeneous universe which is known as '*Lemaitre-Tolman-Bondi metric*' (LTB).

2.2 The Robertson-Walker Metric

To describe the universe, one needs a metric. The usual flat space-time Minkowski metric in Special Relativity

$$\mathrm{d}s^2 = c^2\,\mathrm{d}t^2 - \mathrm{d}x^2 - \mathrm{d}y^2 - \mathrm{d}z^2 \tag{2.1}$$

provides a trivial example of a static, homogeneous and isotropic metric. To generalize this to a dynamic metric, one can include an arbitrary time dependent scale function, like

$$\mathrm{d}s^2 = (c\,\mathrm{d}t)^2 - a(t)^2\left[f(r)\,\mathrm{d}r^2 + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right]. \tag{2.2}$$

And by including the cosmological principle, from homogeneity, the Ricci scalar[1] for the space part of the metric should be constant which gives us

$$\mathrm{d}s^2 = (c\,\mathrm{d}t)^2 - a(t)^2\left[\frac{\mathrm{d}r^2}{1-k\,r^2} + r^2\left(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2\right)\right] \tag{2.3}$$

where r, θ, ϕ are the co-moving coordinates; t is the proper time, $a(t)$ is a function named '*cosmic scale factor*' or '*expansion parameter*' and k is a constant which can be $1, 0, -1$. This metric is called Robertson-Walker metric. If $\mathrm{d}s^2 = 0$, the interval is light-like or null, which is important for strong lensing since we are interested in light-ray path and the time delay due to the different paths.

The value of k is important because it shows the curvature of the universe. $k = 1$ shows a closed universe which means that the universe has finite volume but without boundaries. If the value of k equals zero we end up with a flat universe and $k = -1$ gives us an open universe which behaves like hyperbolic space.

2.3 The Hubble Law

From the Robertson-Walker metric we know that the proper distance d_p for point p from the point which we assume to be the origin of the polar coordinate can be determined with $\mathrm{d}t = 0$

[1] In Riemannian geometry, the scalar curvature or Ricci scalar is the simplest curvature invariant of a Riemannian manifold.

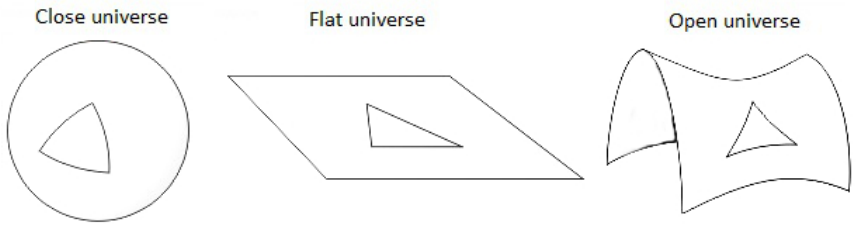

Figure 2.1: Examples of curved spaces in 2D

as

$$d_{\mathrm{p}}(t) = \int_0^r \frac{a(t)\,\mathrm{d}r'}{(1 - k\,r'^2)^{\frac{1}{2}}} = a(t)\,F(r), \tag{2.4}$$

where $F(r)$ is

$$F(r) = \begin{cases} \sin^{-1} r & k = 1 \\ r & k = 0 \\ \sinh^{-1} r & k = -1 \end{cases}$$

But the problem is that we can not measure all distance elements between the origin and p simultaneously. The proper distance at present is related to time t such that

$$d_{\mathrm{p}}(t_0) = a_0 F(r) = \frac{a_0}{a(t)} d_{\mathrm{p}}(t) \tag{2.5}$$

where a_0 is $a(t_0)$. By defining the co-moving distance as

$$d_{\mathrm{c}} = a_0 F(r) \tag{2.6}$$

we obtain the relation between co-moving coordinate and proper distance as

$$d_{\mathrm{c}}(t) = \frac{a_0}{a(t)}\, d_{\mathrm{p}}(t). \tag{2.7}$$

This proper distance d_{p} of a source varies with time due to $a(t)$. Therefore, the source has radial velocity at point p given by

$$v_{\mathrm{r}} = \dot{a}(t)\,F(r) = \frac{\dot{a}(t)}{a(t)}\, d_{\mathrm{p}}(t). \tag{2.8}$$

This equation is called the '*Hubble Law*' and the quantity

$$H(t) = \frac{\dot{a}}{a} \tag{2.9}$$

is called the '*Hubble parameter*'. The value of the Hubble parameter at present $H(t_0) = H_0$, is the Hubble constant. This equation explains what Hubble found in 1929. This means that any object (e.g. galaxy) that we observe is moving away from us and its velocity depends on the distance of that object from us. As an example, if we look at distant and nearby galaxies, we would see that the distant galaxies have higher velocity than the nearby ones.

2.4 Expansion of the Universe and Redshift

Since the scale factor $a(t)$ is not directly observable, astronomers use '*redshift*' z. We define the redshift as

$$z = \frac{\lambda_0 - \lambda_e}{\lambda_e} \tag{2.10}$$

where λ_0 is the observed wavelength (at time t_0) and λ_e is the wavelength of the light as emitted from the source (emitted at time t_e). The source as a part of the universe is moving with the expansion and is at a co-moving coordinate r. The radiation from the source is travelling on a light cone ($\mathrm{d}s^2 = 0$) and therefore we have

$$\int_{t_e}^{t_0} \frac{c\,\mathrm{d}t}{a(t)} = \int_0^r \frac{\mathrm{d}r'}{(1 - k\,r'^2)^{1/2}} = F(r). \tag{2.11}$$

If we consider the light emitted at $t'_e = t_e + \delta t_e$ from the same source, the observer receives it at $t'_0 = t_0 + \delta t_0$. Since $F(r)$ does not change with time and δt_e and δt_0 are assumed to be small, from above equation we have

$$\frac{\delta t_0}{a_0} = \frac{\delta t_e}{a(t_e)}. \tag{2.12}$$

If we have $\delta t_e = \lambda_e/c$ and $\delta t_0 = \lambda_0/c$, then we obtain

$$\frac{a}{\lambda_e} = \frac{a_0}{\lambda_0}, \tag{2.13}$$

which leads to

$$1 + z = \frac{a_0}{a}. \tag{2.14}$$

2.5 The Deceleration Parameter

As was denoted the previous sections, the Hubble parameter or in particular the Hubble constant describes the expansion of the universe. The behaviour of the Hubble parameter depends on the contents of the universe. This can be expressed by expanding the scale factor for a time t close to t_0 in a power series:

$$a(t) = a_0 \left[1 + H_0\,(t - t_0) - \frac{1}{2}\,q_0\,H_0^2\,(t - t_0)^2 + \ldots\right], \tag{2.15}$$

or equivalently

$$z = H_0\,(t_0 - t) + \left(1 + \frac{1}{2}\,q_0\right) H_0^2\,(t_0 - t)^2 + \ldots, \tag{2.16}$$

where

$$q_0 = -\frac{\ddot{a}(t_0)\, a_0}{\dot{a}(t_0)^2} \tag{2.17}$$

is the deceleration parameter. To find r as a function of redshift we have to imply (2.16) in (2.11) and by concerning the first three orders of the polynomial, we have

$$r = \frac{c}{a_0 H_0}\left[z - \frac{1}{2}(1 + q_0)\, z^2 + \ldots\right]. \tag{2.18}$$

2.6 Cosmological Distances

Since we can not make our measurements along a hyper surface of constant proper time, we ought to measure along our past light cone. Therefore, we should define other kinds of distances which can be measured directly.

One of these defined distances is the *luminosity distance*, d_L. Let L denote a power emitted by a source at point p with coordinate distance r at time t and l is the power received per unit area (flux) by the observer at time t_0. Then we define

$$d_L = \left(\frac{L}{4\pi l}\right)^{\frac{1}{2}}. \tag{2.19}$$

The area of a spherical surface in which the source is at centre is $4\pi a_0^2 r^2$. The photon which has been emitted from the source is redshifted in energy due to the expansion of the universe by factor of $\frac{a}{a_0}$. Also by including time-dilation effect (change in the arrival rate of photons) which we have seen in equation (2.12), we have

$$l = \frac{L}{4\pi\, a_0^2\, r^2}\left(\frac{a}{a_0}\right)^2 \tag{2.20}$$

which leads to

$$d_{\mathrm{L}} = a_0^2\, \frac{r}{a}. \tag{2.21}$$

By using the equation (2.18) we derive

$$d_L = \frac{c}{H_0}\left[z + \frac{1}{2}(1 - q_0)\, z^2 + \ldots\right]. \tag{2.22}$$

Another definition for distance is *angular diameter distance* d_A. If the proper diameter of a source is $D_p(t)$ at time t and distance r, one has

$$D_p = a(t)\, r\, \Delta\theta, \tag{2.23}$$

in which $\Delta\theta$ is the angle subtended by D_p. As a result we have

$$d_A = \frac{D_p}{\Delta\theta} = a(t)\, r. \tag{2.24}$$

In gravitational lensing the angular diameter distance is used because it relates physical transverse separation to angle.

2.7 Friedmann Equation

To go further in cosmology we need to use General Relativity to relate geometry to matter in the universe which can be seen in the Einstein's field equations

$$R_{ij} - \frac{1}{2}\, g_{ij} R = \frac{8\pi G}{c^4}\, T_{ij} \tag{2.25}$$

where R_{ij} and R are Ricci tensor and scalar (tensor and scalar for describing the curvature in the geometry) and T_{ij} is the energy-momentum tensor which for a perfect fluid[1] is defined as

$$T_{ij} = (p + \rho\, c^2)\, U_i\, U_j - p\, g_{ij}. \tag{2.26}$$

In this equation p denotes pressure and $\rho\, c^2$ is the energy density. U_k is the fluid velocity.

If we use the Robertson-Walker metric as metric which is describing the universe, we have

$$\ddot{a} = \frac{-4\pi}{3}\, G\left(\rho + \frac{3p}{c^2}\right) a \quad ; \quad \text{0-th component} \tag{2.27}$$
$$a\,\ddot{a} + 2\dot{a}^2 + 2\,k\,c^2 = 4\pi\, G\left(\rho - \frac{p}{c^2}\right) a^2 \quad ; \quad \text{space-space components.}$$

By eliminating $\ddot{a}$ from above equations, we obtain

$$\dot{a}^2 + k\, c^2 = \frac{8\pi}{3}\, G\rho\, a^2. \tag{2.28}$$

Another equation which can be derived from (2.27) is

$$\dot{\rho} + 3\left(\rho + \frac{p}{c^2}\right)\frac{\dot{a}}{a} = 0. \tag{2.29}$$

Equation (2.29) is the *fluid equation*[2]. There are two terms which can change the density. The first one corresponds to the dilution in the density because the volume has been increased and the second one $\frac{p}{c^2}$ is due to the loss of energy because the pressure of matter has done work due to the expansion of the volume of the universe.

[1] In physics, a perfect fluid refers to a fluid which can be characterized by its energy density and pressure.
[2] Equations (2.28) and (2.29) can also be achieved from a Newtonian approach.

2.8 Cosmological Constant

As mentioned in the previous chapter, Einstein formulated General Relativity without cosmological constant in 1916 and by the time he finished his calculation the majority of astronomers believed that the universe is static. From the 0-th component of equation (2.27) for a static universe we have $\rho = \frac{-3p}{c^2}$ which means that the pressure or energy density should be negative which is physically not reasonable. Therefore in 1917 Einstein added the cosmological constant term Λ

$$R_{ij} - \frac{1}{2}\, g_{ij} R - \Lambda\, g_{ij} = \frac{8\pi G}{c^4}\, T_{ij}. \tag{2.30}$$

We can redefine the energy-momentum tensor as

$$\tilde{T}_{ij} = T_{ij} + \Lambda\, \frac{c\, H}{8\pi G}\, g_{ij} = -\tilde{p}\, g_{ij} + \left(\tilde{p} + \tilde{\rho}\, c^2\right)\, U_i\, U_j, \tag{2.31}$$

where the effective pressure and effective energy density are

$$\tilde{p} = p - \frac{\Lambda\, c^4}{8\pi G},\ \tilde{\rho} = \rho + \frac{\Lambda\, c^2}{8\pi G}. \tag{2.32}$$

These equations indicate that $|\Lambda|^{-\frac{1}{2}}$ has the dimension of length. If we define $\rho_\Lambda = \frac{\Lambda\, c^2}{8\pi G}$ as the energy density of Λ, the positive value of Λ requires pressure $p_\Lambda = \omega\, \rho_\Lambda\, c^2$ to be negative. $\omega = -1$ represents the cosmological constant. This positive value of Λ or equivalently negative pressure implies an accelerated expansion of the universe.

For an expanding universe in Robertson-Walker metric, we have

$$\begin{aligned} \ddot{a} &= \frac{-4\pi}{3}\, G\, \left(\tilde{\rho} + 3\frac{\tilde{p}}{c^2}\right) a \\ \dot{a} + k\, c^2 &= \frac{8\pi G}{3}\tilde{\rho}\, a^2. \end{aligned} \tag{2.33}$$

These equations can give a static universe as with the condition below

$$\tilde{\rho} = -3\frac{\tilde{p}}{c^2} = \frac{3k\, c^2}{8\pi G\, a^2}. \tag{2.34}$$

A good approximation for our universe at present is the dust universe ($p = 0$) which in the static case leads us to

$$\Lambda = \frac{k}{a^2},\ \rho = \frac{k\, c^2}{4\pi G\, a^2}. \tag{2.35}$$

Since $\rho > 0$ then we should have $k = 1$ which leads to $\Lambda > 0$ and to have static universe, the value of Λ should be

$$\Lambda_E = \frac{4\pi G}{c^2}\rho. \tag{2.36}$$

This mode is called Einstein universe. After Hubble's discovery, this constant was omitted from Einstein field equation but in 1998 published observations of type Ia supernovae by the High-z Supernova Search Team and followed in 1999 by the Supernova Cosmology Project showed that the expansion of the universe is accelerating (Perlmutter et al., 1999).

2.9 Friedmann Model

A. Friedmann (1922) derived the properties of the Friedmann equations explained in section 2.7, although at early 1920s his model was not accepted because of the belief in a static universe.

By considering a perfect fluid with density ρ and pressure p from equation (2.25), (2.26) and (2.27) we are able to rewrite them in this way

$$\left(\frac{\dot{a}}{a_0}\right)^2 - \frac{8\pi}{3}G\rho\left(\frac{a}{a_0}\right)^2 = H_0^2\left(1 - \frac{\rho_0}{\rho_{0c}}\right) \tag{2.37}$$

where $H_0 = \frac{\dot{a_0}}{a_0}$ and $\rho_{0c} = \frac{3H_0^2}{8\pi G}$ is the critical density at present. If we define $\Omega_0 = \frac{\rho_0}{\rho_{0c}}$ we have

$$H_0^2\,(1 - \Omega_0) = \frac{-k\,c^2}{a_0^2}. \tag{2.38}$$

Now by including Λ (cosmological constant) in equation (2.30) and having Friedmann model for the universe we obtain

$$\begin{aligned} \ddot{a} &= \frac{-4\pi}{3}\,G\left(\rho + 3\frac{p}{c^2}\right)a + \frac{\Lambda\,c^2\,a}{3} \\ \dot{a}^2 + k\,c^2 &= \frac{8\pi\,G}{3}\rho\,a^2 + \frac{\Lambda\,c^2\,a^2}{3}. \end{aligned} \tag{2.39}$$

If we ignore the pressure and matter density we have

$$\left(\frac{\dot{a}}{a_0}\right)^2 - \frac{\Lambda\,c^2}{3} = H_0^2\left(1 - \frac{\Lambda}{\Lambda_c}\right) = H_0^2\,(1 - \Omega_{0\Lambda}) = \frac{-k\,c^2}{a_0^2} \tag{2.40}$$

where the critical value for Λ at present is

$$\Lambda_{0c} = \frac{3H_0^2}{c^2},\ \Omega_{0\Lambda} = \frac{\Lambda\,c^2}{3H_0^2}. \tag{2.41}$$

Now if we reinstate ordinary matter in above equations we see that the curvature is zero if $\Omega_{0\Lambda} + \Omega_0 = 1$.

Emphasis of this chapter was to review the basic knowledge of modern cosmology. In the next chapter, an introduction to the theory of gravitational lensing and how the Hubble constant can be determined from the strong gravitational lensing effect is provided.

3

Theory of Gravitational Lensing

In this chapter, the theory of gravitational lensing and in particular strong lensing which can provide us the determination of the Hubble constant is explained. The book by Schneider et al. (2006) was used in this chapter.

3.1 Light Deflection in General Relativity

After Einstein completed his calculation on General Relativity he realized that his calculation of the deflection angle of light in the presence of a gravitational field is a factor of two larger than what the Newtonian approach predicted[1]. In General Relativity this angle with direction radially towards the deflector is

$$\hat{\alpha} = \frac{4G\,M}{c^2\,\xi} = 1''.75\left(\frac{M}{M_\odot}\right)\left(\frac{\xi}{R_\odot}\right)^{-1} \tag{3.1}$$

where ξ is the impact parameter, M is the deflector's spherical mass, $M_\odot = 1.988 \times 10^{30}\,\mathrm{kg}$ and $R_\odot = 6.995 \times 10^5\,\mathrm{km}$ are mass and radius of sun. The term *lens* for the object causing the deflection was first used by Lodge (1919). Fritz Zwicky (1937) published two papers in which he considered lensing by galaxies (at that time it was '*extra galactic nebulae*').

First detection of multiple images

Walsh et al. (1979) discovered two quasars with separation of approximately 6 arcseconds (this system is known as QSO0957+561) with similar properties like same colour, spectra and

[1] Einstein showed that with the equivalence principle, the deflection angle is equal to the Newtonian approach. This shows that the factor of two in his later calculation equally corresponds to the equivalence principle and curvature of the space-time.

redshift of 1.41. With the VLA (Very Large Array) observations of these two quasars it was confirmed that both quasars are compact radio sources with same radio spectra (Roberts et al., 1979). Later observations with the VLBI (Very Long Baseline Interferometry), showed a similar core-jet structure in both quasars which was another proof of gravitational lensing in this system (Porcas et al., 1981).

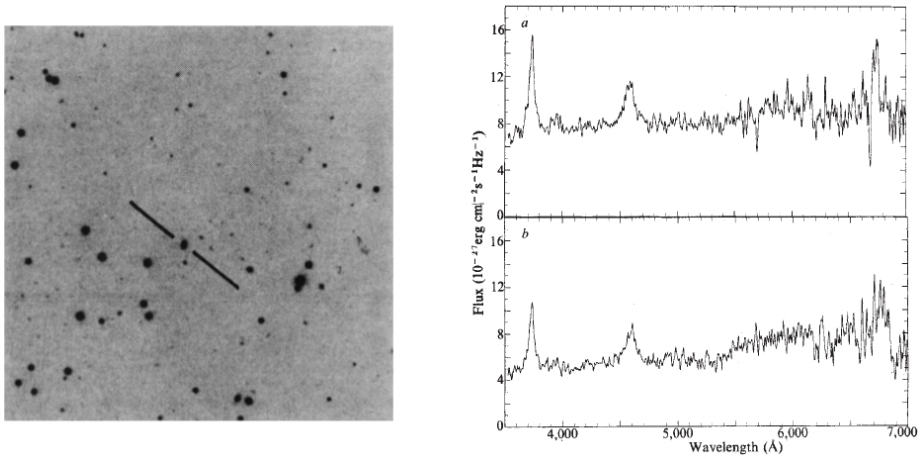

Figure 3.1: In the left panel the system QSO0957+561 is pointed with solid lines and in the right panel the spectra of the two images (Walsh et al., 1979).

Another lensing system with triple images of a quasar PG1115+080 was discovered by Weymann et al. (1980). Later observations on this system showed that the brightest image was a blend of two images with separation of $\sim 0.5''$.

3.2 Gravitational Lensing in Cosmology

Mass measurements

The gravitational potential can be related to the mass distribution of the object by Poisson's equation. If the light ray passes along this potential field, it would be deflected. Therefore, the deflection of the light ray depends on the mass distribution of the object (lens) i.e. the more massive the lens, the larger the deflection angle in the light ray path. From the equations which will be shown in this chapter one can estimate the mass distribution of the lens from observation of multiple images and rings, within the separation or ring diameter.

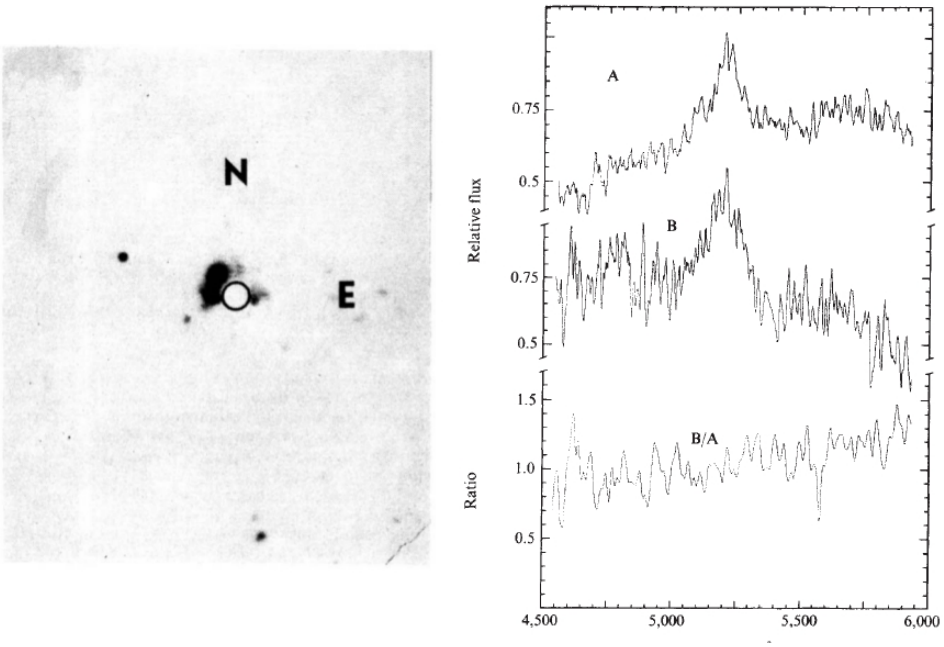

Figure 3.2: Left panel shows the system PG 1115+080 and right panel are the spectra of two of the images of this triple (later confirmed as quad) system (Weymann et al., 1980).

Estimates on cosmological parameters

Refsdal (1964) proposed that with the use of the time delays in multiple image systems, the Hubble constant can be obtained. Another cosmological parameter is the fraction of lensed high-redshift quasars which can be used to estimate the cosmological model.

3.3 Lens Equation

Let us assume that the source is at an angular size distance D_{S} and the lens at an angular size distance D_{L} from the observer. We presume that the extension of the mass distribution of the lens along the line of sight is negligible in comparison to D_{L} (*thin lens* approximation) and D_{LS} to be the distance between lens and source. If there is no other deflector along the light ray path, then we can approximate the curved path of the light ray with a straight line with kink in the lens plane as shown in Fig. 3.3.

If $\vec{\eta}$ is the position of the source with respect to the axis which is chosen to be the line of

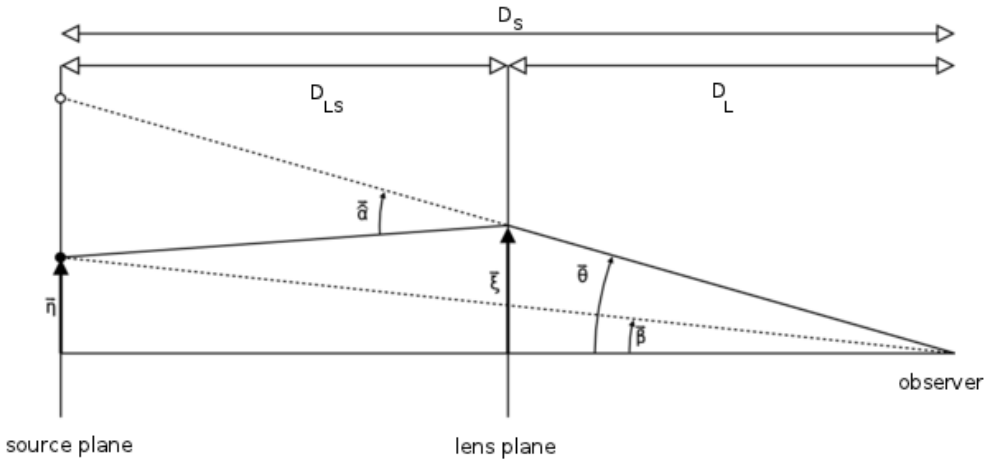

Figure 3.3: Sketch of a simple lens system.

sight and all angles are assumed to be small, then from geometry we have

$$\vec{\eta} = \frac{D_{\mathrm{S}}}{D_{\mathrm{S}}}\,\vec{\xi} - D_{\mathrm{LS}}\,\hat{\vec{\alpha}}(\vec{\xi}), \tag{3.2}$$

where $\vec{\xi}$ is distance in the lens plane and $\vec{\eta}$ in the source plane. $\vec{\eta}$ and $\vec{\xi}$ can be written as

$$\vec{\eta} = D_{\mathrm{S}}\,\vec{\beta},\ \vec{\xi} = D_{\mathrm{L}}\,\vec{\theta}. \tag{3.3}$$

By replacing (3.3) in (3.2) we have

$$\vec{\beta} = \vec{\theta} - \frac{D_{\mathrm{LS}}}{D_{\mathrm{S}}}\,\vec{\alpha}\left(D_{\mathrm{L}}\vec{\theta}\right) \equiv \vec{\theta} - \vec{\alpha}(\vec{\theta}), \tag{3.4}$$

where we defined the *scaled deflection angle* $\vec{\alpha}(\vec{\theta})$. If we have a source at position $\vec{\beta}$ then it can be seen by the observer at position $\vec{\theta}$ satisfying equation (3.4). If this equation has more than one solution for a fixed source at position $\vec{\beta}$ then the observer sees multiple images of the source in the sky.

3.4 Deflection Angle

Deflection of the light ray is due to the amount of mass distributed along its way. The field equation of General Relativity can be linearised if the gravitational field is weak. Let us assume $\rho(\vec{r})$ as the density distribution of the lens, then the mass distribution would be $\mathrm{d}m = \rho(\vec{r})\,\mathrm{d}V$. If we define the spatial trajectory $(\xi_1(\lambda), \xi_2(\lambda), r_3(\lambda))$ (these coordinates are chosen such that

the incoming light ray propagates along r_3) then with the approximation of a geometrically-thin lens and assumption of significant mass for the lens, we obtain $\vec{\xi}(\lambda) \approx \vec{\xi}$, $\vec{\xi} = (\xi_1, \xi_2)$. Therefore the deflection angle becomes

$$\begin{aligned} \hat{\vec{\alpha}}(\vec{\xi}) &= \frac{4G}{c^2} \sum \mathrm{d}m\,(\xi_1', \xi_2', r_3') \frac{\vec{\xi} - \vec{\xi}'}{\left|\vec{\xi} - \vec{\xi}'\right|^2} \\ &= \frac{4G}{c^2} \int \mathrm{d}^2\xi' \int \mathrm{d}r_3' \,\rho(\xi_1', \xi_2', r_3') \frac{\vec{\xi} - \vec{\xi}'}{\left|\vec{\xi} - \vec{\xi}'\right|^2}. \end{aligned} \tag{3.5}$$

With the definition of surface mass density

$$\Sigma(\vec{\xi}) \equiv \int \mathrm{d}r_3\, \rho(\xi_1, \xi_2, r_3), \tag{3.6}$$

the deflection angle for an arbitrary density distribution can be established as

$$\hat{\vec{\alpha}}(\vec{\xi}) = \frac{4G}{c^2} \int \mathrm{d}^2\xi'\, \Sigma(\vec{\xi}') \frac{\vec{\xi} - \vec{\xi}'}{\left|\vec{\xi} - \vec{\xi}'\right|^2}. \tag{3.7}$$

If we define a dimensionless surface mass density or *convergence* as below

$$\kappa(\vec{\theta}) := \frac{\Sigma(D_\mathrm{L}\vec{\theta})}{\Sigma_\mathrm{cr}};\ \Sigma_\mathrm{cr} = \frac{c^2}{4\pi G} \frac{D_\mathrm{S}}{D_\mathrm{L}\, D_\mathrm{LS}}, \tag{3.8}$$

the scaled deflection angle can be rewritten as

$$\vec{\alpha}(\vec{\theta}) = \frac{1}{\pi} \int \mathrm{d}^2\theta'\, \kappa(\vec{\theta}') \frac{\vec{\theta} - \vec{\theta}'}{\left|\vec{\theta} - \vec{\theta}'\right|^2}. \tag{3.9}$$

As can be seen, the *critical surface mass density*, Σ_cr depends on the distance of the source and lens. The case of $\kappa \geq 1$ or equivalently $\Sigma \geq \Sigma_\mathrm{cr}$ produces multiple images of the source in certain positions. Therefore this parameter is characteristic to divide strong and weak lensing regimes.

3.5 Deflection Potential

Like other parts of physics in which we define a potential, in gravitational lensing the fact that $\nabla \ln|\vec{\theta}| = \vec{\theta}/|\vec{\theta}|^2$ gives us the possibility of defining a *deflection potential* as follows

$$\psi(\vec{\theta}) = \frac{1}{\pi} \int_{\Re^2} \mathrm{d}^2\theta' \kappa(\vec{\theta}') \ln|\vec{\theta} - \vec{\theta}'|, \tag{3.10}$$

which has the property of

$$\vec{\alpha} = \nabla\psi, \tag{3.11}$$

and if we take the second derivative of this potential we obtain

$$\nabla^2\psi = 2\kappa, \tag{3.12}$$

which is the Poisson equation. Furthermore if we define a function

$$\tau\left(\vec{\theta};\vec{\beta}\right) = \frac{1}{2}\left(\vec{\theta}-\vec{\beta}\right)^2 - \psi(\vec{\theta}) \tag{3.13}$$

and take the derivative above equation with respect to θ we have

$$\nabla\tau\left(\vec{\theta};\vec{\beta}\right) = 0. \tag{3.14}$$

τ is called the *Fermat potential* and equation (3.14) is equal to equation (3.4). In optics, the Fermat's principle or the principle of least time is the principle that rays of light traverse the path of stationary optical length which can be expressed as Eq. (3.14).

This potential is important in the classification of multiple images in gravitational lensing systems. Apart from that, this potential can be used to understand the geometrical behaviour of the lens mapping (Blandford & Narayan, 1986).

3.6 Magnification, Shear and Distortion

In addition to multiple images of the source, gravitational lensing can cause distortion, magnification or demagnification of the image(s) of it. Since there is no emission and absorption in a gravitational lensing event, the surface brightness is conserved.

If $I_\nu^s\left(\vec{\beta}\right)$ is the surface brightness distribution of the source at frequency ν then the observed surface brightness would be

$$I_\nu(\vec{\theta}) = I_\nu^s\left(\vec{\beta}(\vec{\theta})\right). \tag{3.15}$$

Flux from a source in solid angle $\mathrm{d}\omega^*$ is

$$S_\nu^* = I_\nu \mathrm{d}\omega^*. \tag{3.16}$$

If the light bundle is deflected by a mass distribution, then the solid angle $\mathrm{d}\omega$ of the image is different from $\mathrm{d}\omega^*$. As deflection does not change the frequency nor intensity, therefore the flux changes to

$$S_\nu = I_\nu \mathrm{d}\omega. \tag{3.17}$$

Hence, the deflection changes the flux by the factor

$$|\mu| = \frac{S_\nu}{S_\nu^*} = \frac{\mathrm{d}\omega}{\mathrm{d}\omega^*}. \tag{3.18}$$

This magnification $|\mu|$ can be obtained from the Jacobian determinant which shows the distortion of lens mapping. If we define the Jacobian matrix as

$$A(\vec{\theta}) = \frac{\partial \vec{\beta}}{\partial \vec{\theta}},\ A_{ij} = \frac{\partial \beta_i}{\partial \theta_j}, \tag{3.19}$$

then the magnification factor is

$$\mu(\vec{\theta}) = \frac{1}{\det A(\vec{\theta})}. \tag{3.20}$$

This magnification factor can be both positive and negative which refers to positive or negative parity of the images. In the cases that $\det A(\vec{\theta})$ vanishes we have critical points/curves which are the points or curves in which the magnification mathematically goes to infinity. They are explained in more detail in the next section. We can rewrite the Jacobian matrix in the form of

$$\begin{aligned} A(\vec{\theta}) = \frac{\partial \vec{\beta}}{\partial \vec{\theta}} &= \left(\delta_{ij} - \frac{\partial^2 \psi(\theta)}{\partial\theta_i \partial\theta_j}\right) \\ &= \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}, \end{aligned} \tag{3.21}$$

where $\gamma \equiv \gamma_1 + \mathrm{i}\gamma_2 = |\gamma| e^{2\mathrm{i}\varphi}$ is the shear

$$\gamma_1 = \frac{1}{2}(\psi_{11} - \psi_{22}),\ \gamma_2 = \psi_{12}, \tag{3.22}$$

and κ is the convergence. Using the locally linearised lens equation at a certain point $\vec{\theta_0}$ and correspondingly $\vec{\beta_0}$, we have

$$I(\vec{\theta}) = I^s\left[\vec{\beta_0} + A(\vec{\theta_0})(\vec{\theta} - \vec{\theta_0})\right]. \tag{3.23}$$

From this equation for a circular source we obtain elliptical images. The ratios of semi-axes of this ellipse to the radius of circular source are $\lambda_{+/-} = 1 - \kappa \pm |\gamma|$. Therefore the magnification factor becomes

$$\mu = \frac{1}{\det A} = \frac{1}{(1-\kappa)^2 - |\gamma|^2}. \tag{3.24}$$

The total magnification of all the images of the source is

$$\mu_p(\vec{\beta}) = \sum_i \left|\mu(\vec{\theta_i})\right|. \tag{3.25}$$

For sources with finite extent (not point source) this equation changes to

$$\mu = \left[\int \mathrm{d}^2\beta\, I^s(\vec{\beta})\right]^{-1} \int \mathrm{d}^2\beta\, I^s(\vec{\beta})\,\mu_p(\vec{\beta}). \tag{3.26}$$

If we rewrite the equation (3.21) like

$$A(\vec{\theta}) = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \tag{3.27}$$

then we defined the reduced shear

$$g \equiv \frac{\gamma}{1-\kappa} = \frac{|\gamma|}{1-\kappa} e^{2\mathrm{i}\varphi} \equiv g_1 + \mathrm{i}g_2. \tag{3.28}$$

As can be seen from Eq. (3.27) the right side of equation is divided in two parts of convergence and reduced shear. The convergence is responsible for enlargement/reduction of the images of the source, but the second term consisting of reduced shear corresponds to distortion in the shape of the images. Therefore the observable quantity is reduced shear[1].

3.7 Critical Curves and Caustics

As mentioned before when the magnification of the Jacobian matrix diverges, we have closed, smooth curves called critical curves and by mapping them from the lens plane to the source plane we have *caustics*. The position of the source relative to these caustics can tell us about the number of images produced by lensing. These caustic curves are not smooth as critical curves. If the direction of the tangential vector to the critical curve is singular then the caustic is not smooth and has a *cusp*. Other parts of the caustic which are smooth are called *fold*.

3.8 Classification of Images

For a fixed $\vec{\beta}$, the Fermat potential $\tau(\vec{\theta};\vec{\beta})$ gives us a surface in which the images of the source are the extremum points or $\nabla\tau = 0$. These extremum points have one of the following three types below:

1. τ has a (local) minimum.

$$\det A > 0;\ \mathrm{tr}\, A > 0 \Rightarrow |\gamma| < (1-\kappa) \leq 1,\ \mu \geq \frac{1}{1-|\gamma|^2} \geq 1. \tag{3.29}$$

[1] more detail on the mathematical proof of reduced shear as an observable quantity can be found in Schneider et al. (1995).

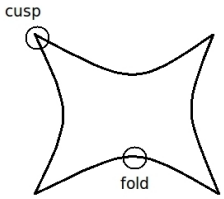

Figure 3.4: Different parts of a caustics

2. τ has a saddle point

$$\det A < 0 \Rightarrow (1-\kappa)^2 < |\gamma|^2. \tag{3.30}$$

3. τ has a (local) maximum.

$$\det A > 0;\ \mathrm{tr}\, A < 0 \Rightarrow (1-\kappa)^2 > |\gamma|^2,\ \kappa > 1. \tag{3.31}$$

If we consider the thin lens approximation with smooth surface mass distribution for the lens and n is the total number of images and n_1, n_2, n_3 refer to the number of images with types discussed above, we have $n = n_1 + n_2 + n_3$. If the source is not located on a caustic, then $n_1 \geq 1, n < \infty$ and $n_1 + n_3 = 1 + n_2$. This leads us to the fact that the number of images of the source must be odd which is called the *odd number theorem* in lensing. With a finite mass distribution, for large $|\vec{\theta}|$, the deflection potential increases as $\ln|\vec{\theta}|$. Therefore the Fermat potential $\tau(\vec{\theta};\vec{\beta})$ behaves as $|\vec{\theta}|^2/2$. With smooth mass distribution, τ behaves smoothly. Thus there is at least one minimum of τ corresponding to image of type 1. It can also be shown that the first image received by the observer is a type 1 image and brighter or with the same brightness of the source.

3.9 The Mass-Sheet Degeneracy

The main problem in lensing is that from observations we know the positions of the images and we want to reconstruct the mass distribution of the lens. Assume that we observe a multiple-imaged source in which the positions of images and fluxes can be measured. With this information one can reproduce the mass distribution or mass model of the lens. Falco et al. (1985)

and later Gorenstein et al. (1988) discussed the situation where one can have more than one possible solution for the mass distribution of the lens which is named 'mass-sheet degeneracy'.

In previous sections $\kappa(\vec{\theta})$ was introduced as the mass distribution of the lens. Now we assume that this $\kappa(\vec{\theta})$ fits to our observable quantities and we define $\kappa_\lambda(\vec{\theta})$ as below

$$\kappa_\lambda(\vec{\theta}) = (1-\lambda) + \lambda\,\kappa(\vec{\theta}) \tag{3.32}$$

and λ is an arbitrary value. If we include the $\kappa_\lambda(\vec{\theta})$ in Eq. (3.9) we obtain $\vec{\alpha}_\lambda = (1-\lambda)\vec{\theta} + \lambda\vec{\alpha}(\vec{\theta})$. With this new deflection angle the lens equation and deflection potential become

$$\vec{\beta} = \vec{\theta} - \vec{\alpha}_\lambda(\vec{\theta}), \tag{3.33}$$

$$\psi_\lambda(\vec{\theta}) = \frac{1-\lambda}{2}\,|\vec{\theta}|^2 + \lambda\,\psi(\vec{\theta}). \tag{3.34}$$

With these equations, the Poisson equation for deflection potential is still valid $\nabla^2\psi_\lambda(\theta) = 2\kappa_\lambda(\theta)$. If we expand α_λ in equation (3.33) we obtain

$$\frac{\vec{\beta}}{\lambda} = \vec{\theta} - \vec{\alpha}(\vec{\theta}). \tag{3.35}$$

This shows a rescaling in the source plane which can not be directly observed. As a consequence of this, the Jacobian matrix, magnification and shear are changed to

$$A_\lambda = \lambda A\,;\ \mu_\lambda = \frac{\mu}{\lambda^2}\,;\ \gamma_\lambda(\theta) = \lambda\,\gamma(\theta). \tag{3.36}$$

If we have other constraints, e.g. the Hubble constant from other observations or sources at different distances we are able to break this degeneracy. This degeneracy can also be broken using additional information such as magnification of extended structure of the source revealed in the images.

3.10 Strong Gravitational Lensing

The two largest gravitational lens surveys are the Cosmic Lens All-Sky Survey (CLASS) and the Sloan Lens ACS Survey (SLACS). Candidate selection in the CLASS was done in the radio with the VLA, MERLIN and VLBA (1990–2003), and resulted in the discovery of 22 lens systems[1]. The SLACS survey (2003–present) combined SDSS spectroscopic pre-selection with HST imaging follow-up and has yielded $\sim$ 100 new lens systems[2]. Currently there are

[1] http://www.jb.man.ac.uk/research/gravlens/class/class.html

[2] http://www.slacs.org/

$\sim$ 200 known strong lens systems in which with availability of time delay, redshift of the source or the lens provide a large number of solid scientific studies such as measurements of H_0, strong constraints on galaxy density profiles and their evolution.

Commonly we observe systems with double images (in some cases with Einstein ring) or four images. Quads can be classified in three different cases as shown in figure 3.5. There are cruciform quads, where the images form a cross pattern bracketing the lens when the source lies almost directly behind the non-spherical lens. If the source is near a fold caustic or fold-dominant case then we observe two magnified images close to each other. When we have cusp-dominant (source close to cusp caustic), we see a close triple of highly magnified images.

3.10.1 Circular lens model

As a simple and first step for modelling gravitational lensing systems one can start with a circular lens although in reality we cannot neglect angular structure of the gravitational potential.

In this model the lens potential is just a function of distance from the centre of the lens. Therefore one can calculate the deflection angle as below

$$\alpha(\theta) = \frac{4G\,M(<\xi)}{c^2\,\xi}\frac{D_{\rm LS}}{D_{\rm S}} = \frac{2}{\theta}\int_0^{\theta}\theta'\mathrm{d}\theta'\,\kappa(\theta'), \tag{3.37}$$

where $\xi = D_L\theta$. From this deflection angle one obtains

$$\vec{\beta} = \vec{\theta}\left(1 - \frac{\alpha(\theta)}{\theta}\right) = \vec{\theta}\,(1 - \langle\kappa(\theta)\rangle)\,, \tag{3.38}$$

where

$$\langle\kappa(\theta)\rangle = \frac{2}{\theta^2}\int_0^{\theta}\theta'\,\mathrm{d}\theta'\,\kappa(\theta') = \frac{\alpha(\theta)}{\theta}. \tag{3.39}$$

Since we have a circular symmetry, if we draw a line between all images it passes through the source and the centre of the lens. Also the Jacobian matrix (inverse magnification matrix[1]) can be written as follows

$$A^{-1} = \frac{\partial\vec{\beta}}{\partial\vec{\theta}} = (1-\kappa)\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} + \gamma\begin{pmatrix}\cos 2\phi & \sin 2\phi\\ \sin 2\phi & -\cos 2\phi\end{pmatrix}, \tag{3.40}$$

where $\vec{\theta} = \theta(\cos\phi, \sin\phi)$. As a result we have

$$\kappa = \frac{1}{2}\left(\frac{\alpha}{\theta} + \frac{\mathrm{d}\alpha}{\mathrm{d}\theta}\right),\ \gamma = \frac{1}{2}\left(\frac{\alpha}{\theta} - \frac{\mathrm{d}\alpha}{\mathrm{d}\theta}\right) = \langle\kappa\rangle - \kappa. \tag{3.41}$$

[1] magnification matrix is the inverse of Jacobian matrix

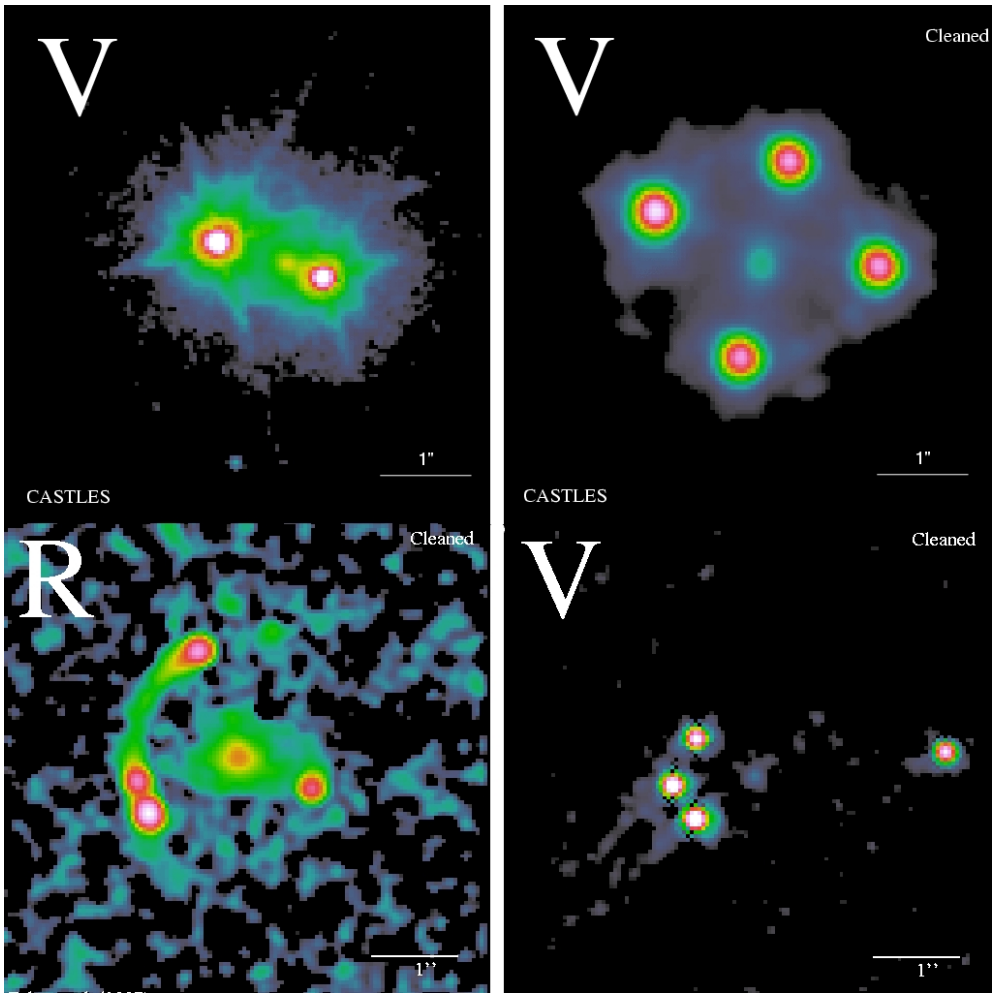

Figure 3.5: Image of double and three different configurations for quads. Upper left is the double image system QJ0158–4325 and upper right is a cross in the quad images system HE0435–1223. Bottom left is the fold-dominant system MG0414+0534 and bottom right is the cusp-dominant system RXJ0911+0551 (Kochanek et al.- CASTLES website)

The eigenvalues of A^{-1} are $\lambda_+ = 1 - \kappa + \gamma = 1 - \frac{d\alpha}{d\theta}$ in the radial direction and $\lambda_- = 1 - \kappa - \gamma = 1 - \langle\kappa\rangle$ in the tangential direction. If at least one of these eigenvalues goes to zero then we have infinite magnification and we are in radial or tangential critical curves.

If we assume that the lens galaxy has the power-law density profile $\rho \propto r^{-n}$, we have

$$\alpha(\theta) = b\left(\frac{\theta}{b}\right)^{2-n} \tag{3.42}$$

where b is a constant and it can be shown that

$$\kappa(\theta) = \frac{3-n}{2}\left(\frac{\theta}{b}\right)^{1-n}, \; \gamma(\theta) = \frac{n-1}{2}\left(\frac{\theta}{b}\right)^{1-n}. \tag{3.43}$$

Figure 3.6 presents the deflection angle for different values of n. The power-law lenses can cover most of simple and physical models. In this model the tangential magnification eigen-

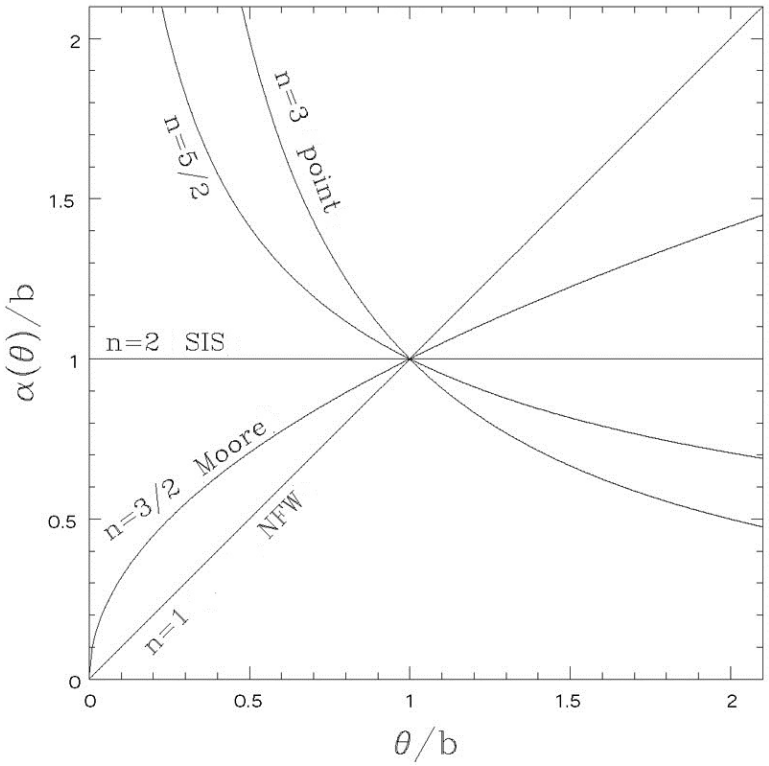

Figure 3.6: Deflection angle function with different values for n (Schneider et al., 2002).

value is

$$1 - \kappa - \gamma = 1 - \frac{\alpha}{\theta} = 1 - \langle\kappa\rangle = 1 - \left(\frac{\theta}{b}\right)^{1-n}, \tag{3.44}$$

which is always zero at $\theta = b = \theta_E$. This circle is what we call *Einstein ring* with Einstein radius θ_E. In this case the mean surface density inside this tangential critical radius is $\langle\kappa\rangle = 1$.

The definition of b depends on the model that we choose. The radial magnification eigenvalue of this model becomes

$$1 - \kappa + \gamma = 1 - (2 - n)\left(\frac{\theta}{b}\right)^{1-n} \tag{3.45}$$

which can be zero when $n < 2$.

Images are created or destroyed when the source crosses a caustics. For power-law models, the tangential critical line is $\theta = b$ which is the solution of $\alpha(b) = b$ and relates to $\beta = 0$ and a point source at origin is transformed into an Einstein ring with radius $\theta_E = b$.

Since in this thesis the emphasis is to determine the Hubble constant, not to do detailed modelling for the lens system B0218+357 (a singular isothermal model is considered for the mass profile of the lens galaxy), we continue this chapter with details on time delay theory.

3.11 Time Delay and Hubble Constant Determination

In the case of multiple images, light rays from a source take different paths and reach the observer at different times, an effect called 'time delay'. This delay can be measured from monitoring the image fluxes for variable sources.

Time delay is produced by bending of the light in the curved geometry in General Relativity (geometrical time delay) and the light traverse in a gravitational field of the lens (potential time delay). In the weak-field regime, coordinate travel time dt for a light ray to travel in Euclidean length dl is

$$c\,\mathrm{d}t \approx \left(1 - \frac{2\Phi_\mathrm{N}}{c^2}\right)\mathrm{d}l, \tag{3.46}$$

where Φ_N is the Newtonian potential. The classical way to express this effect is Refsdal's method (Refsdal 1964) which will be explained in the following subsection and later the general term of the current formulation of time delay.

3.11.1 Determining the Hubble parameter in Refsdal method

Refsdal (1964) suggested a lens system in which the source is a supernova and he showed that from these images of the source, the Hubble constant and the mass of the lens can be determined.

Refsdal considered the supernova S is far from the lens galaxy B and lies close to the line of sight. He also made two assumptions:

1. The lens galaxy is spherically symmetric.

2. The redshift of the source and lens are small.

Gravitational lensing introduces a deflection angle of light passing a lens object with mass M at distance r as

$$\alpha = 4G\,M\,c^{-2}\,r^{-1} = K\,r^{-1}. \tag{3.47}$$

In Fig. 3.7, the gravitational lensing system of two images with distances r_1 and r_2 with respect to lens position is shown. The distance between the observer O and lens galaxy B at redshift z_{L} is D_{L} and source S is D_{S} with redshift z_{S}. The distance from O to extension of SB is X. We assume that $r_1 > 0$, $r_2 < 0$ and $X > 0$ in right side and $X < 0$ in the left side of Fig. 3.7. By considering the approximation of small angles, from Fig. 3.7 we have

$$r^2 - X\,n^{-1}\,r - K\,D_{\mathrm{L}}\,n^{-1} = r^2 - X\,n^{-1}\,r - r_0^2 = 0, \tag{3.48}$$

where with the second assumption we have

$$n = \frac{D_{\mathrm{S}}}{D_{\mathrm{S}} - D_{\mathrm{L}}} = \frac{z_{\mathrm{S}}}{z_{\mathrm{S}} - z_{\mathrm{L}}} \tag{3.49}$$

and

$$r_0 = \sqrt{\frac{K\,D_{\mathrm{L}}}{n}} \tag{3.50}$$

which is the Einstein radius when $X = 0$. Therefore we have

$$r_1 = \frac{1}{2n}\left(X + \sqrt{X^2 + 4n^2\,r_0^2}\right) \tag{3.51}$$

$$r_2 = \frac{1}{2n}\left(X - \sqrt{X^2 + 4n^2\,r_0^2}\right) \tag{3.52}$$

Again from Fig. 3.7 we see

$$\theta_1 + \theta_2 = \theta \tag{3.53}$$

$$X = n\,D_{\mathrm{L}}\,\beta \tag{3.54}$$

By concerning equations (3.51), (3.52) and (3.53) we get

$$\theta_1 - \theta_2 = \frac{r_1 + r_2}{D_{\mathrm{L}}} = \frac{X}{n\,D_{\mathrm{L}}} = \beta \tag{3.55}$$

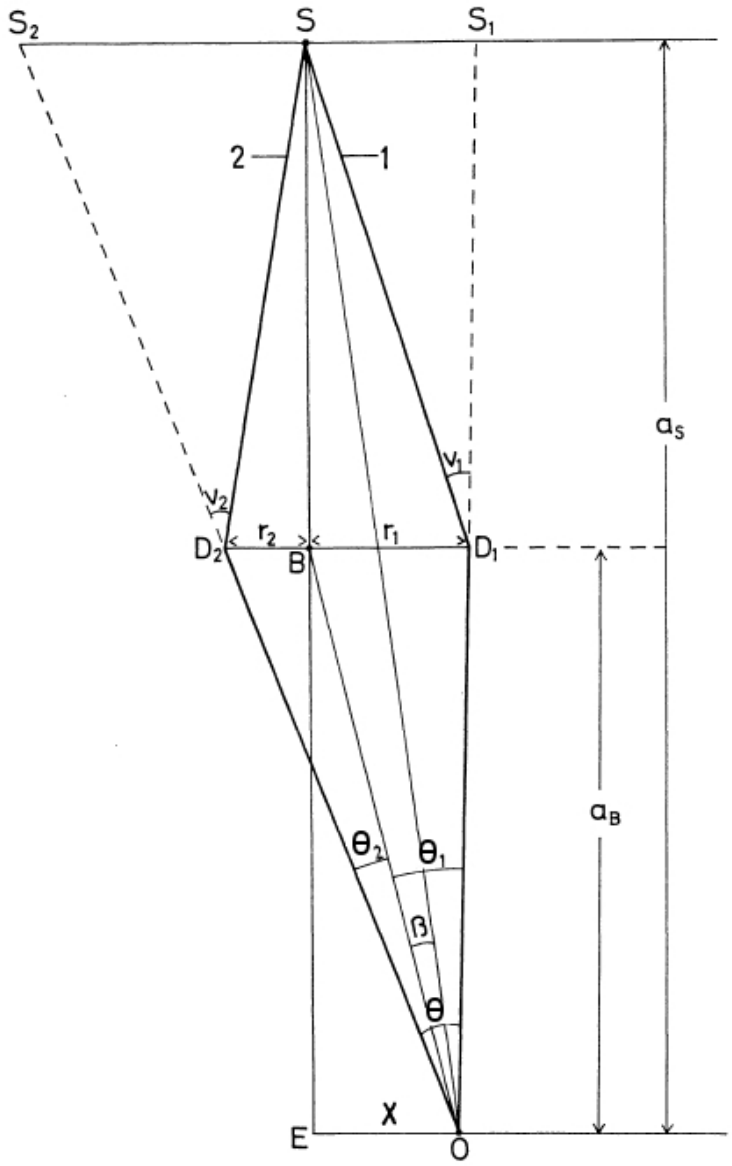

Figure 3.7: Two light ray path from source S to observer O (Refsdal, 1964).

From the assumptions at the beginning of the calculations we have $\theta_2 > 0$, thus $\theta_2 = -\frac{r_2}{D_\mathrm{L}}$. Then we obtain

$$\sqrt{X^2 + 4n^2\, r_0^2} = \sqrt{n^2\, D_\mathrm{L}^2 \beta^2 + n^2\, D_\mathrm{L}^2 \theta_0^2} = n D_\mathrm{L} \sqrt{\beta^2 + \theta_0^2}, \tag{3.56}$$

where θ_0 is the deflection angle for the case $X = 0$. The result is

$$\theta_0 = \frac{2r_0}{D_\mathrm{L}} = 2\sqrt{\frac{K}{n\, D_\mathrm{L}}}. \tag{3.57}$$

From equations (3.51) and (3.52) we see

$$r_1 - r_2 = n^{-1} \sqrt{X^2 + 4n^2\, r_0^2} = D_\mathrm{L}\theta. \tag{3.58}$$

Implying Eq. (3.56) in Eq. (3.58) leads us to

$$\theta = \sqrt{\theta_0^2 + \beta^2}. \tag{3.59}$$

Now if we assume that β is small we have

$$\theta \approx \theta_0 \left(1 + \frac{\beta^2}{2\theta_0^2}\right) \approx \theta_0. \tag{3.60}$$

The time delay between two light ray paths can be determined from

$$\Delta t = c^{-1} \int_0^X \theta \, \mathrm{d}X. \tag{3.61}$$

Applying equations (3.54) and (3.60) in Eq. (3.61) gives us

$$\Delta t = c^{-1} \int_0^\beta n\, D_\mathrm{L}\, \theta_0 \, \mathrm{d}\beta \approx n D_\mathrm{L} \theta_0 c^{-1} \beta. \tag{3.62}$$

Hence by inserting the definition of θ_0 we get

$$\Delta t = \frac{16G}{c^3} \frac{\beta}{\theta_0} M. \tag{3.63}$$

Our second assumption gives us

$$D_\mathrm{L} = z_\mathrm{L} \frac{c}{H_0},\; D_\mathrm{S} = z_\mathrm{S} \frac{c}{H_0}, \tag{3.64}$$

where H_0 is the Hubble constant. Now if we insert D_L from Eq. (3.62), β from Eq. (3.55) and definition of n in Eq. (3.64) the Hubble constant becomes

$$H_0 = \frac{z_\mathrm{S}\, z_\mathrm{L}\, \theta_0\, (\theta_1 - \theta_2)}{\Delta t\, (z_\mathrm{S} - z_\mathrm{L})} \tag{3.65}$$

and the lens mass is

$$M = \frac{\Delta t\, \theta_0\, c^3}{16G\, (\theta_1 - \theta_2)}. \tag{3.66}$$

3.11.2 General term for time delay

So far we assumed that the optical axis connects the observer to the centre of the lens. Now if we consider the observer to be at distance ζ from the optical axis O' as is shown in Fig. 3.8 and source is at distance η'. According to the definition of optical axis, the new optical axis has its origin at O and the distance of the source from this axis is η. As can be seen from figure 3.8 the lens equation (3.2) still holds and the relation between ζ and η' is

$$\vec{\phi} = \frac{\vec{\eta} - \vec{\eta'}}{D_{\mathrm{LS}}} = \frac{\vec{\zeta}}{D(z_{\mathrm{L}}, 0)} \tag{3.67}$$

where z_{L} is the redshift of the lens and $D(z_{\mathrm{L}}, 0) = (1 + z_{\mathrm{L}})\, D_{\mathrm{L}}$[1]. Therefore we have

$$\vec{\eta'} + \frac{D_{\mathrm{LS}}}{D_{\mathrm{L}}\,(1 + z_{\mathrm{L}})}\,\vec{\zeta} = \frac{D_{\mathrm{S}}}{D_{\mathrm{L}}}\,\vec{\xi} - D_{\mathrm{LS}}\,\vec{\theta}(\vec{\xi}). \tag{3.68}$$

With applying the definition of $\vec{\eta} = D_{\mathrm{S}}\,\vec{\beta}$ we obtain

$$\vec{\beta} = \frac{1}{D_{\mathrm{S}}}\left[\vec{\eta'} + \frac{D_{\mathrm{LS}}}{D_{\mathrm{L}}(1 + z_{\mathrm{L}})}\,\vec{\zeta}\right]. \tag{3.69}$$

If we consider the situation that the observer at position ζ observes two images of a fixed source at $\vec{\theta}^{(1)}$ and $\vec{\theta}^{(2)}$. Now if the observer moves along a curve $\vec{\zeta}(\lambda)$ and as long as the observer does not cross a caustic curve, there will be two image curves $\vec{\theta}^{(i)}(\lambda)$ which satisfy the lens equation for $i = 1, 2$. With the movement of the observer by the amount of $\mathrm{d}\vec{\zeta}$, the intersect of the wave-fronts of the two images are detected by the observer at different times. Therefore, this delay can be derived as

$$\mathrm{d}(c\,\Delta t) = \vec{\vartheta} \cdot \mathrm{d}\vec{\zeta}, \tag{3.70}$$

where $\vec{\vartheta} = \vec{\theta}^{(2)} - \vec{\theta}^{(1)}$ is the angular image separation of the images. With integration over $\vec{\zeta}(\lambda)$, the time delay becomes

$$c\,\Delta t(\vec{\zeta}) = \int_{\vec{\zeta}_0}^{\vec{\zeta}} \vec{\vartheta}(\vec{\zeta'}) \cdot \mathrm{d}\vec{\zeta'} + c\,\Delta t(\vec{\zeta_0}) \tag{3.71}$$

and $\vec{\zeta_0}$ is an arbitrary reference point. With use of the Fermat potential we are able to omit this reference point (Schneider, 1985). From Eq. (3.69) the dimensionless displacement of observer can be related to $\mathrm{d}\vec{\beta}$ as

$$\mathrm{d}\vec{\zeta} = (1 + z_{\mathrm{L}})\,\frac{D_{\mathrm{L}}\,D_{\mathrm{S}}}{D_{\mathrm{LS}}}\,\mathrm{d}\vec{\beta}. \tag{3.72}$$

[1] In this relation since the distance is considered from lens to observer term $(1 + z_{\mathrm{L}})$ should be multiplied to D_{L}.

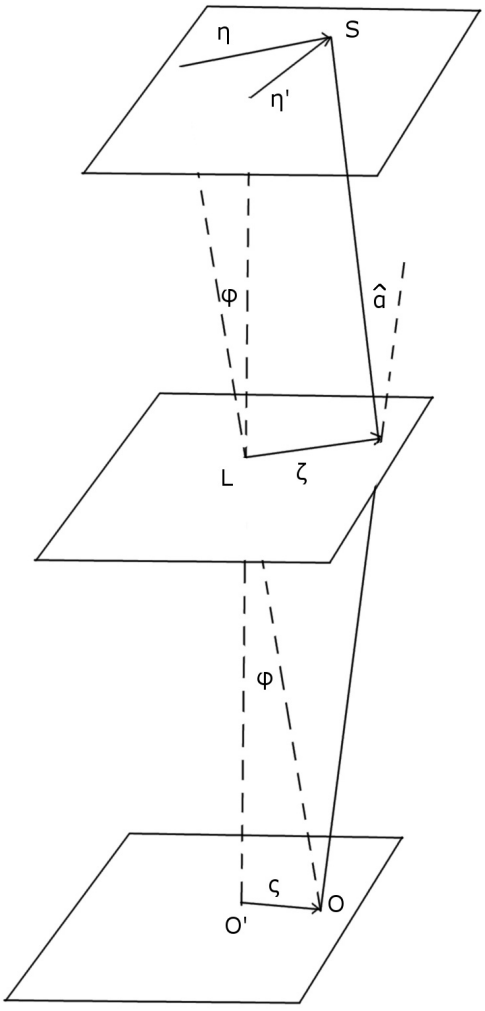

Figure 3.8: Schematic sketch of gravitational lensing.

Hence we obtain

$$c\,\Delta t(\vec{\beta}) = (1+z_{\rm L})\,\frac{D_{\rm L}\,D_{\rm S}}{D_{\rm LS}}\int_{\vec{\beta_0}}^{\vec{\beta}} {\rm d}\vec{\beta'}\,\left[\vec{\theta}^{(2)}(\vec{\beta'}) - \vec{\theta}^{(1)}(\vec{\beta'})\right] + c\,\Delta t(\vec{\beta_0}). \tag{3.73}$$

Including $\vec{\beta} = \vec{\theta} - \nabla\phi(\vec{\theta})$ and

$$\vec{\theta}\cdot{\rm d}\vec{\beta} = {\rm d}(\vec{\theta}\cdot\vec{\beta}) - \vec{\theta}\cdot{\rm d}\vec{\theta} + {\rm d}\psi(\vec{\theta}) = {\rm d}\left[\beta^2/2 - \tau(\vec{\theta};\vec{\beta})\right] \tag{3.74}$$

we derive

$$c\Delta t(\vec{\beta}) = (1+z_{\rm L})\,\frac{D_{\rm L}\,D_{\rm S}}{D_{\rm LS}}\,\left[\tau\left(\vec{\theta}^{(1)};\vec{\beta}\right) - \tau\left(\vec{\theta}^{(2)};\vec{\beta}\right)\right]. \tag{3.75}$$

Now if we consider a general form of an isothermal model[1][2] without external shear, Witt et al. (2000) showed that the time delay can be written only as a function of the image positions with respect to the lens galaxy (without including the ellipticity explicitly) as follow

$$c\,\Delta t_{i,j} = \frac{1+z_{\rm L}}{2H_0}\,\frac{d_{\rm L}d_{\rm S}}{d_{\rm LS}}\left(|\vec{\theta}_i - \vec{\theta}_0|^2 - |\vec{\theta}_j - \vec{\theta}_0|^2\right), \tag{3.76}$$

where $\vec{\theta}_0$ is the centre of the lens galaxy. As can be seen from equation (3.76) the Hubble constant (or time delay) has a linear relation with the position of the lens galaxy. Figure 3.9 presents the change of the Hubble constant with respect to the lens position for the lens system B0218+357.

[1]Isothermal models are referred to the density profile of $\rho \propto r^{-2}$ and a flat rotation curve for galaxies. Spiral galaxy rotation curves, stellar dynamics of elliptical galaxies are consistent with roughly isothermal profiles.

[2]Reader is encouraged to read the paper by Keeton (2001) for more information on different type of mass models in lens modelling.

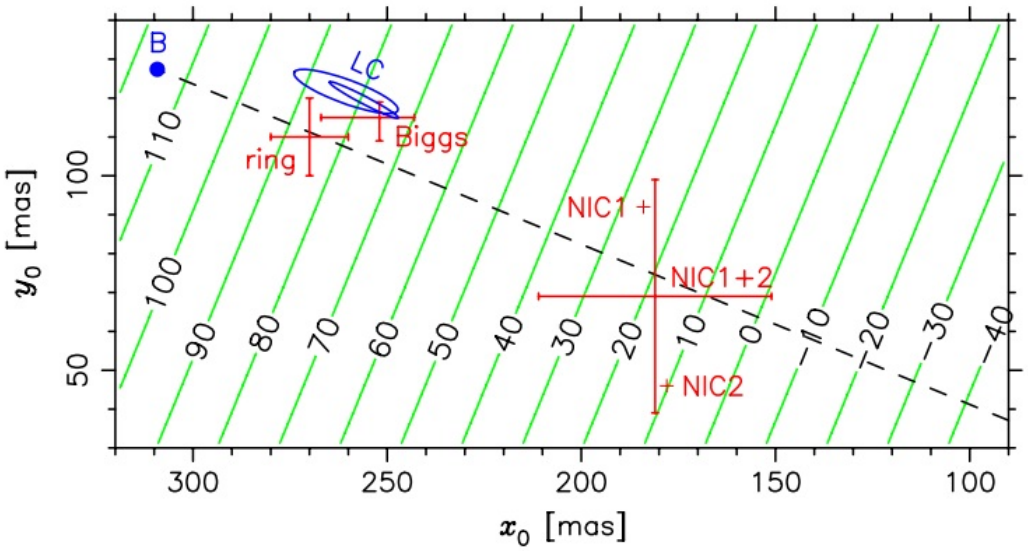

Figure 3.9: H_0 in units of km s^{-1} Mpc^{-1} as a function of the lens position for B0218+357. NIC1 and NIC2 are the positions of the lens galaxy from Lehar et al. (2000) and upper left are the derived positions for the lens galaxy by Biggs et al. (1999) and LS refers to the result from LensClean algorithm. The dashed line connects image B to image A (Wucknitz et al., 2004).

4

B0218+357

Golden Lenses are strong lensing systems (with the known time delays) in which the lens model can be constrained accurately from observations[1] and this leads to the determination of the Hubble constant. The lens system JVAS B0218+357 could be characterized as a Golden Lens systems which has double-images with a radio Einstein ring.

The separation of double images in B0218+357 is the smallest separation among the known systems in galactic scale lensing. Despite this, other properties of this system give us a chance to determine the cosmological parameter H_0. In this Chapter this system, its properties both in radio and optical and previous studies related to determination of the Hubble constant on this system are introduced.

4.1 Discovery and Properties of B0218+357

The lens system B0218+357 was discovered during the time when Jodrell-VLA Astronomic Survey (JVAS) was about to select interferometer phase calibration compact radio sources in the declination range $35^\circ \leq \delta \leq 75^\circ$, to be used as phase reference sources for MERLIN observations for the Cosmic Lens All-Sky Survey (CLASS) to search for lensed systems (Patnaik et al., 1992).

Observations with the VLA, MERLIN and VLBI[2] (Patnaik et al., 1993) showed that this

[1] the model for the lens should be explained with minimum parameters and no external perturbations (isolated lens galaxy) or well-known perturbations for these systems.

[2] VLA, MERLIN and VLBI are radio arrays and interferometers with frequency range of 74 MHz to 50 GHz, 151 MHz to 24 GHz and 300 MHz to 90 GHz (typically 1.4-22 GHz). Combining the data taken from them gives us higher resolution and sensitivity than previous data.

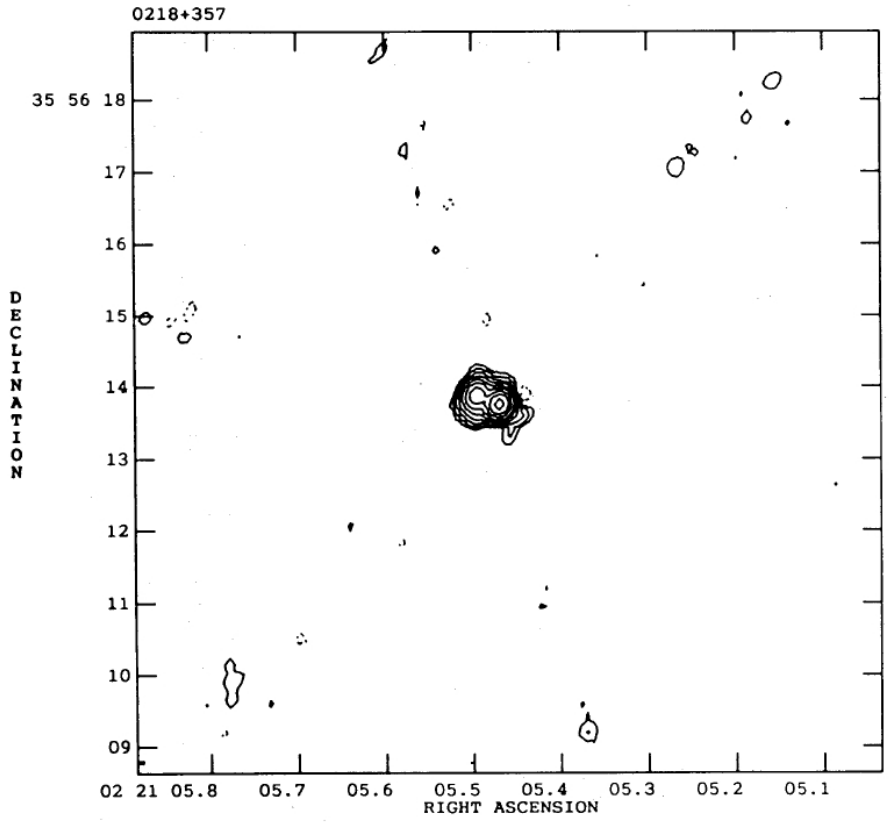

Figure 4.1: The first image of B0218+357 in radio wavelength (Patnaik et al., 1992).

diffuse structure around the double images of the blazar source is ring-like and the two compact components of the system have the same radio spectra. The ring in this system could not be confirmed as a complete ring and was believed to be an image of parts of the extended structure of kpc scale radio jet. Therefore, it varies on time scales longer than the two images of the compact cores (Biggs et al., 1999).

B0218+357 was observed with the William Herschel Telescope (WHT) in 1992 for spectroscopy and the results showed that the redshift of the lens galaxy is $z = 0.6847$ (Browne et al., 1993; Stickel & Kuehr, 1993). From evidences such as Faraday rotation, absorption lines in radio spectra of the system, they claimed that the lens galaxy should be spiral.

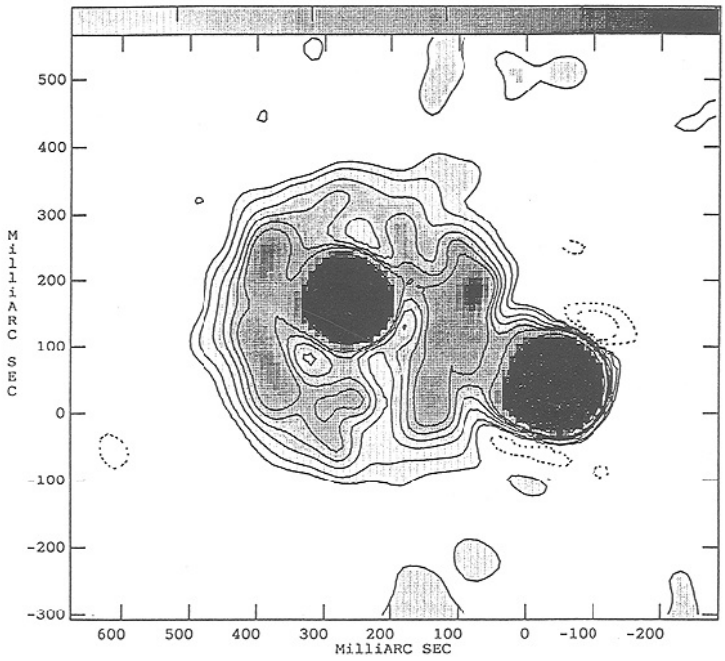

Figure 4.2: The image of B0218+357 in radio wavelength with Einstein ring (Patnaik et al., 1993)

Grundahl et al. (1995) published a paper on the optical appearance of B0218+357. This image was taken with the Nordic Optical Telescope (NOT) and they used aperture and point spread function photometry to resolve the two images in this system with separation of 335 mas. As presented in figure 4.2, image A (west) is brighter than image B (east) in radio but they found that in optical image B is brighter by a factor of 3. Therefore, they concluded that in the line of sight of image A there is a giant molecular cloud in the lens galaxy which absorbs the light coming from image A. This indicates that the lens galaxy is a late type galaxy.

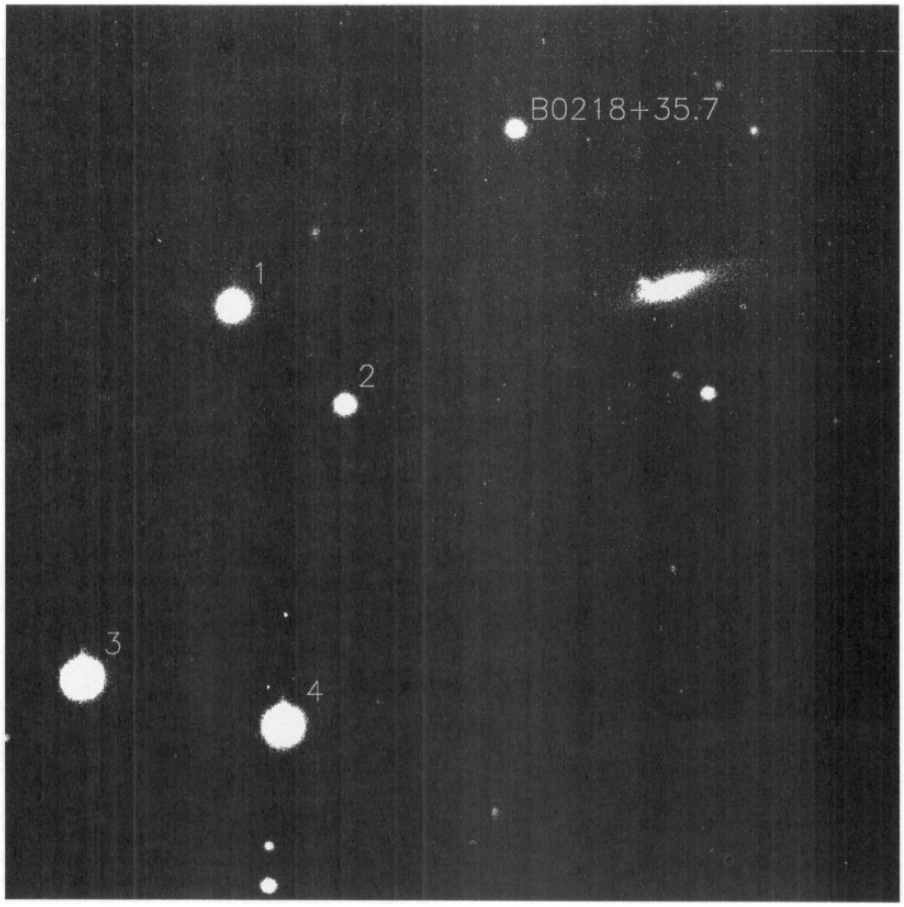

Figure 4.3: The image of B0218+357 in optical (Grundahl et al., 1995).

Later, Biggs et al. (1999) determined the time delay of this system using radio monitoring observations made with the VLA and consistency of variation in polarization and polarization angle. The value that they found as time delay is 10.5 ± 0.4 days. In their work they also derived the Hubble constant for this system. They created 5000 data-sets with Gaussian distribution error on the observed positions, subcomponent sizes and flux density ratio. The solution for minimum χ^2 of the mass model parameters gave the time delay of $7.2^{+1.3}_{-2.0}/h$ days between the images. Implying the time delay they obtained from observation and singular isothermal ellipse (SIE) model for the mass, they found the Hubble constant to be $69^{+13}_{-19}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. The error bars on their result were mainly due to the inaccurate position fitting for the centre of the mass and not well constrained model for the lens galaxy (they assumed cosmological model with $\Omega_0 = 1$ and $\Lambda = 0$).

Optical observation of this system with the Near infra red Camera and Multi-Object Spectrometer (NICMOS) camera of Hubble Space telescope (HST) showed that as predicted the lens galaxy is spiral. Jackson et al. (2000) in their work with NICMOS data were not able to determine the position of the lens galaxy due to its overlap with image B. Cohen et al. (2003) found the redshift of the source to be 0.944 (Lawrence (1996) also determined the source redshift to be 0.96). Apart from lensing, other studies were done on this system like research on gas and dust in proto-galaxies or HI absorption etc. (Wiklind et al., 1995; Combes et al., 1997; Gerin et al., 1997; Kochanek et al., 2000 and Henkel et al., 2005).

Wucknitz et al. (2004) analysed the VLA data (with combination of the constraints from the VLBI data) with the improved version of LensClean method (readers are encouraged to read the paper by Wucknitz (2003) for more details on this method) to determine the position of the lens galaxy for isothermal models. They also studied the effect of deviation from isothermal models with analysis of the structure of the images. This approach let them to constrain the power-law mass profile for this system. At the end they concluded that the deviation from isothermal model has no significant effect on this system. Although their study gave a value of the Hubble constant with smaller errors ($H_0 = 78 \pm 6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) in comparison to Biggs et al. work, the uncertainty on the position of the lens galaxy was a major problem. Therefore, optical observation with high resolution and signal to noise was required to enable us to improve the Hubble constant.

4.2 B0218+357 with the HST

Servicing Mission 3B flown by Columbia (STS-109) in March 2002 saw the installation of a new instrument, ACS. In the ACS, three independent, high-resolution channels cover the spectrum region of ultraviolet to the near-infra red. As mentioned before, because of the uncertainties on obtaining the positions of the spiral lens galaxy in B0218+357, Jackson et al. proposed an observation of this system with the Wide Field Channel (WFC) instrument of the Advanced Camera for Survey (ACS) detector of the HST. This observation was made at the end of 2002 and beginning of 2003. This data is used in this work. In Fig. 4.4 one of the images taken by the ACS/WFC is shown.

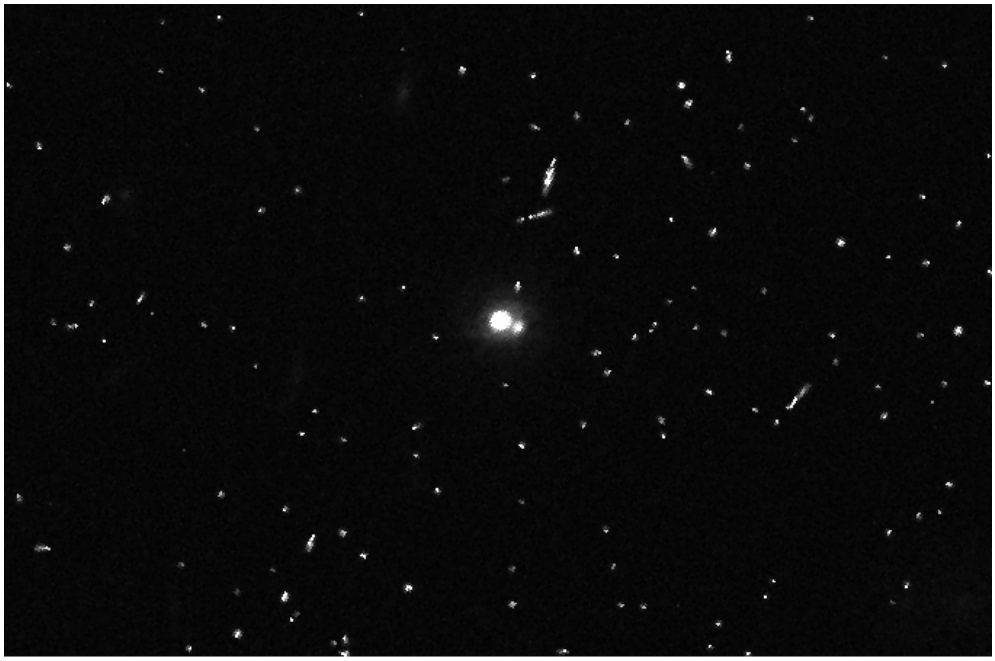

Figure 4.4: One of the raw images of B0218+357 taken with the ACS/WFC.

Table 4.1 presents the visits in this dataset and their exposure times.

Table 4.1: Observed data of B0218+357 by HST

visit nr.	observation date	exposure time(s)	file name root
10	28.02.2003	20×360	j8d410
11	01.03.2003	20×360	j8d411
12	17-18.01.2003	20×360	j8d412
13	06.03.2003	20×360	j8d413
15	11.03.2003	20×360	j8d415
14	26-27.10.2002	20×360	j8d414
16	11.09.2002	12×360	j8d416

This system was observed with the filter F814W. As can be seen in fig 4.5 this filter is a broad band near-infrared filter.

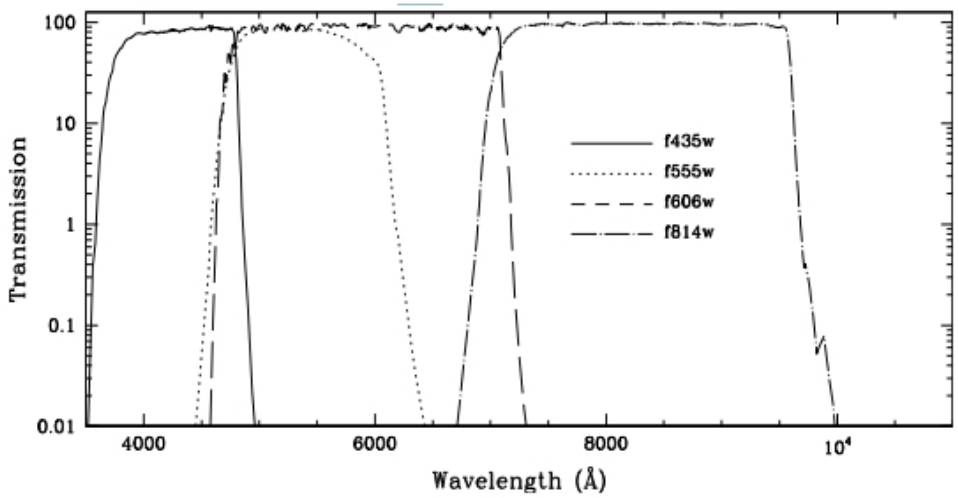

Figure 4.5: The broad band filters of the WFC/ACS instrument (Gonzaga et al., 2011)

Work by York et al. (2005) on this data was not fully satisfactory either. They reduced the data with MultiDrizzle (Drizzle algorithm which is explained in more detail in the next chapter) with additional shifts. They shrank the pixel size of the output image to 25 mas and chose the Gaussian kernel to reduce the blurring and slightly improve the resolution. To subtract the PSF from the two images, one needs an accurate PSF model. Therefore, they produced different PSFs. They used the Landolt PG0213+051B stars but their resulting PSF

suffered from artefacts like extended wings with length of 80 mas. Thus they created a PSF per visit by averaging field stars. To forbid any bias, they also used parametric models (Airy and Gaussian function) with field stars.

From PSF subtraction, they found the separation of the images to be 317 ± 2 mas when field stars are used, 315 ± 4 mas with Gaussian and 311 ± 10 mas with Airy function for modelling the PSF. They used the residuals to find the galaxy position. They applied a maximum symmetry criterion as a goodness of fit parameter. With this approach they obtained the position of the centre of the lens. At the end they determined $H_0 = 79 \pm 7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the case of fitting only the galaxy and $H_0 = 66 \pm 9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when arms are masked out with 95 per cent confidence.

As described before, because of the small separation of the images, the PSF plays an important role in this system and precise PSF subtraction is needed. In York et al. work, they used the Drizzle algorithm which produces additional PSF to the images. This affects the PSF subtraction process which York et al. work suffered from. One solution is to develop an alternative method to combine the images to avoid this additional PSF which is the main goal of this work. In chapter 6, the aspects of this alternative method and in chapter 7, the results of our study on this system are discussed in more detail.

5

HST Data Reduction Pipeline

A condensed version of parts of the following chapters about the direct model fitting method has been submitted for publication in A&A. To write the section 5.2 of this chapter, the ACS instrument handbook (Ubeda et al., 2011) is used.

5.1 Introduction

A distribution of light rays coming from a source towards the CCD is affected by the properties of the optical instrument. The most important effect is the Point-Spread-Function (PSF). If we neglect the small variation of the PSF with position, the true light distribution is convolved with the PSF. The provenances of this effect is the diffraction in the instrument. For ground-based telescopes, the atmosphere also adds additional spreading to the apparent light distribution of the source.

The design of the instrument can cause a geometric distortion. The off-axis location of the detectors (tilted focal surface with respect to the light rays) and the Optical Telescope Assembly (OTA) causes variations in the plate scale (in size and orientation of the pixels). The Wide Field Channel, WFC, a detector of the Advance Camera for Surveys, ACS, of the Hubble Space Telescope, HST, suffers from this optical distortion. In the detector, pixels have response function and integrate the light within their area. In other words, this function describes the sensitivity of the pixel. The last effect that should be considered is the regular sampling effect which can be explained as multiplication with a 'bed-of-nails' function.

Therefore, if we denote the true light distribution as I^*, P the PSF profile, D as the effect of

optical distortion, R for response function and S as sampling, the observed light distribution is

$$I^{\mathrm{obs}} = S \cdot R \star D \otimes P \star I^{*}. \tag{5.1}$$

Here $\star$ denotes a convolution, while $\otimes$ indicates an arbitrary mapping.

In order to reveal information on scales smaller than the pixel size, observations are 'dithered', i.e. split into several exposures that are shifted relative to each other by sub-pixel displacements. Inverting the convolution with P requires the conservation of structures on scales smaller than the PSF width; the sampling must be sufficiently fine to resolve the PSF. For ground-based telescopes, P is dominated by the atmospheric seeing, and modern CCD instruments generally have sufficient resolution. For the WFC detector of the ACS, the pixel size of 0.05 arcseconds (50 mas) is comparable to the overall PSF width, and dithering is essential to recover any of the fine-scale structure of the PSF.

A basic method to combine these dithered images is a linear technique, 'shift-and-add'. In this method, pixels of each image are transferred to a finer grid, shifted to the same position and added to the output image. In the formalism introduced above this corresponds to a convolution with a function R' that represents the size and shape of the pixels. Neglecting the distortion for the moment, and assuming a uniform and complete dithering pattern (so that the sampling S can be neglected), the result would correspond to $R' \star R \star P \star I^{*}$. This additional convolution with the pixel size reduces the resolution of the result. Another disadvantage of this method is that no correction on the geometric distortion is implemented in it.

Another method in linear reconstruction was introduced by Fruchter and Hook, called *Drizzle* (Fruchter et al., 1996). 'Drizzle', the standard method to combine dithered HST images, can handle images with arbitrary shifts and rotations. It allows for different R', correct for the distortion, and identifies and flags pixels affected by cosmic rays. Drizzle inverts the measurement process explained in Eq. (5.1) from left to right. Since there is no precise inversion, this method adds artefacts to the combined image.

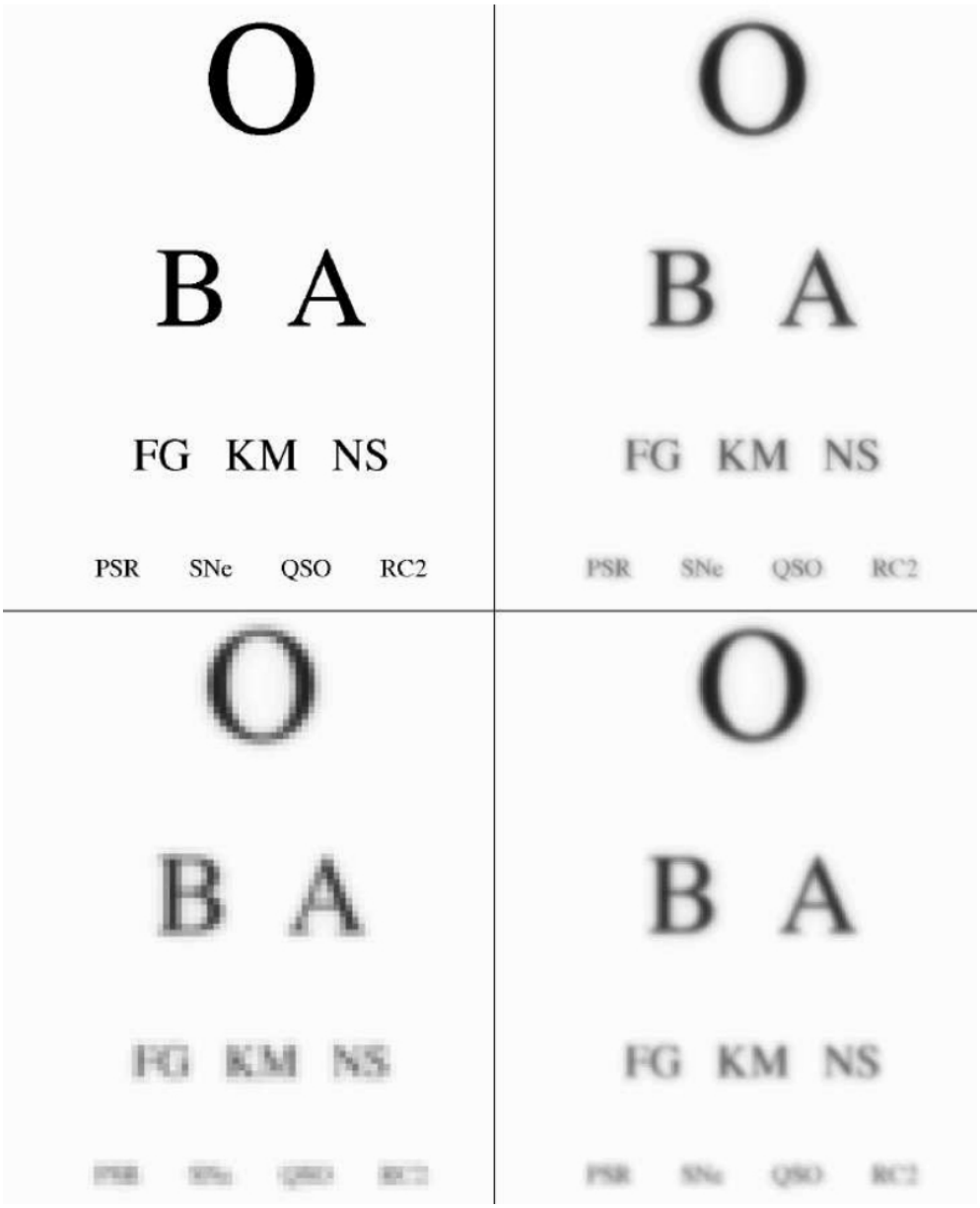

Figure 5.1: The upper left image is the image of the source, upper right shows the light distribution convolved with PSF, lower left image shows the image that we see from the CCD and lower right is the image reconstructed with shift-and-add method- from Fruchter et al. (1996)

5.2 ACS Camera Designs and Properties

ACS provides three different channels, each for a specific purpose:

1. Wide-Field Channel (WFC): This channel has a field of view of $\sim 202 \times 202$ arcseconds from 3700-11000 Å. The plate scale for this instrument is $\sim 0.049 \frac{\text{arcsec}}{\text{pixel}}$. The pixel size of this CCD is $15 \times 15 \mu$m and the image format is $2 \times 2048 \times 4096$ pixels.

2. High-Resolution Channel (HRC): 29×25 arcsecond field of view from 2000-11000 Å. In this instrument the plate scale is $\sim 0.028 \times 0.025 \frac{\text{arcsec}}{\text{pixel}}$ with the pixel size of $21 \times 21 \mu$m. The total size of the CCD is 1024×1024.

3. Solar Blind Channel (SBC): Field of view of 35×31 arcsecond in range of 1150-1700 Å. Plate scale of $\sim 0.034 \times 0.03 \frac{\text{arcsec}}{\text{pixel}}$ is a good compromise between resolution and FOV. In this CCD the pixel size is $25 \times 25 \mu$m and the total number of pixels are 1024×1024.

In our case, B0218+357 was observed with the WFC in the ACS. Figure 5.2 shows the structure of the ACS detector. As can be seen from this figure, the light ray passes the filter wheel and reaches the CCD after it has been reflected by 3 mirrors on its way. This structure of the detector causes the geometric distortion which is explained in more detail in following subsection.

The ACS CCDs are thinned, backside-illuminated devices which are cooled by TEC (Thermo-Electric Cooler). The spectral response of the WFC CCDs are optimized for observing in the range of visible to near-IR wavelength. The CCD camera of the WFC produces a time-integrated image in ACCUM data-taking mode. In this mode, the WFC CCD accumulates the signal during the exposure in response to photons. The charge is read out at the end of the exposure.

The minimum exposure time for the WFC is 0.5 seconds and the minimum time between successive identical exposure is 135 seconds. CCD observations should be broken into multiple exposures whenever possible to allow removal of cosmic rays in post-observation data processing.

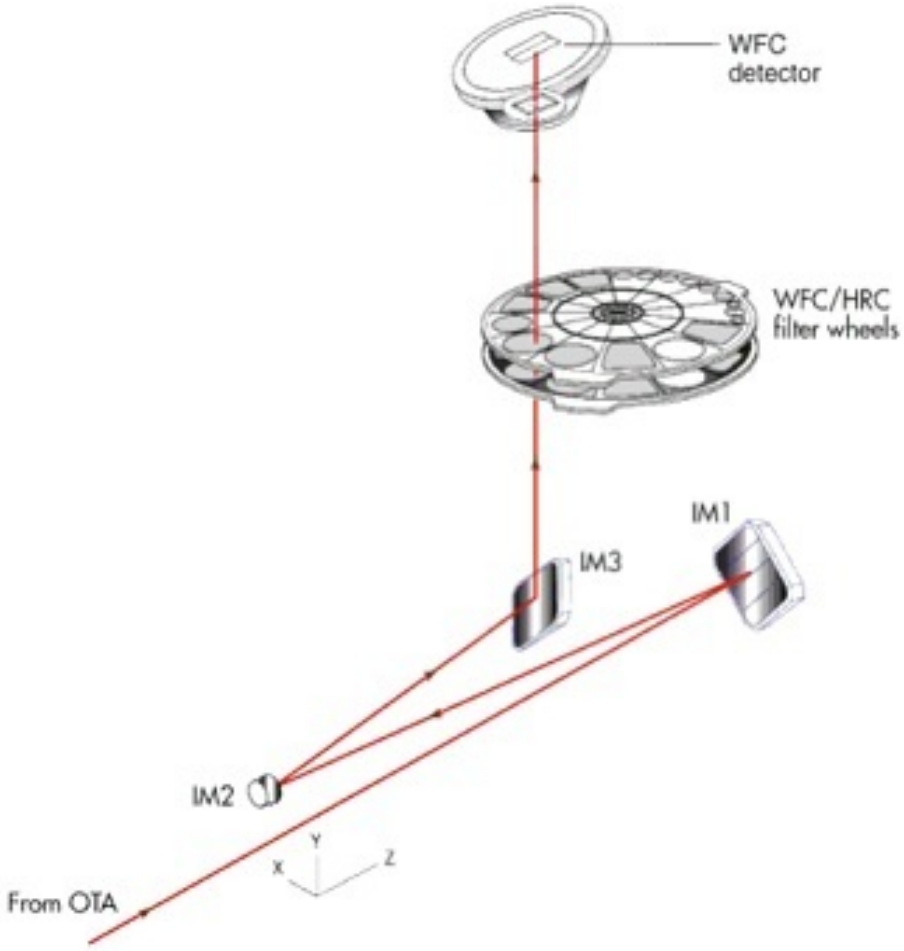

Figure 5.2: Schematic view of ACS detector (Gonzaga et al., 2011)

5.2.1 ACS Distortion

ACS optics were designed to be consistent with the correction of spherical aberration induced by the Optical Telescope Assembly (OTA), without introducing coma. This causes an offset between focal surface and rays at OTA which introduces a distortion in the sky image. This distortion reveals in two ways, first is that the pixel scale are smaller along the radial direction of the OTA field of view than along tangential direction (due to the tilted elongation of the ACS apertures). The second is the variation of pixel area over the CCD. The area on the sky covered by a WFC pixel varies by about 18% from corner to corner, corrections for which must be made in photometry of extended objects. This variation of the scale creates a problematic effect in combining the ACS images by the fact that an integral pixel shift near the centre of the detector changes to a non-integral displacement for the pixels near the edges. This implies some computational complexity in accurate alignment in order to combine the images and will depend on the accuracy of the geometric distortion model.

Distortion Model

The distortion correction model connects the distorted pixel position on CCD to sky coordinates. This transformation is made by polynomial functions given as

$$\begin{aligned} x_c &= \sum_{i=0}^{k} \sum_{j=0}^{i} a_{i,j}(x - x_r)^j (y - y_r)^{i-j} \\ y_c &= \sum_{i=0}^{k} \sum_{j=0}^{i} b_{i,j}(x - x_r)^j (y - y_r)^{i-j}, \end{aligned} \tag{5.2}$$

where (x, y) are the positions of the pixel on the CCD, k is the polynomial order of this fit and (x_r, y_r) are the positions of the reference pixel which is decided to be the centre of each detector chip. In above relation (x_c, y_c) are the undistorted position in arcseconds. The coefficients $a_{i,j}$ and $b_{i,j}$ are free parameters (these coefficients are fitted from observations of star fields). For the WFC, with two CCD chips, we should add a value for the gap between two chips such as

$$\begin{aligned} X' &= x_c + \Delta x(\text{chip}) \\ Y' &= y_c + \Delta y(\text{chip}). \end{aligned} \tag{5.3}$$

The offsets Δx(chip) and Δy(chip) are zero for the first chip and for the second chip it is the gap between two chips which is around 50 pixels. (X', Y') are the tangential plane position in

arcseconds. Finally the transformation will be complete when we apply these relations to these tangential coordinate points

$$\begin{aligned} X_{\rm sky} &= \Delta X_i + \cos(\Delta\theta_i)X' - \sin(\Delta\theta_i)Y' \\ Y_{\rm sky} &= \Delta Y_i + \sin(\Delta\theta_i)X' + \cos(\Delta\theta_i)Y', \end{aligned} \tag{5.4}$$

where ΔX_i, ΔY_i and $\Delta\theta_i$ are the positions and rotation offsets of frame i which are determined from the pointing of the telescope and its roll angle (orientation). To solve for the free parameters ($a_{i,j}$, $b_{i,j}$, ΔX_i, ΔY_i and $\Delta\theta_i$) in the distortion model, the positions of stars (stars in the main sequence turn-off) observed several times in the dithered star fields are used. The first ACS geometric distortion campaign observed the core of 47 Tucanae with the WFC and HRC. A fit for $k = 4$ corrects the distortion to an accuracy of 0.1–0.2 pixels over the entire FOV. In figures 5.3 and 5.4 the non-linear component of distortion fitting and correction are shown.

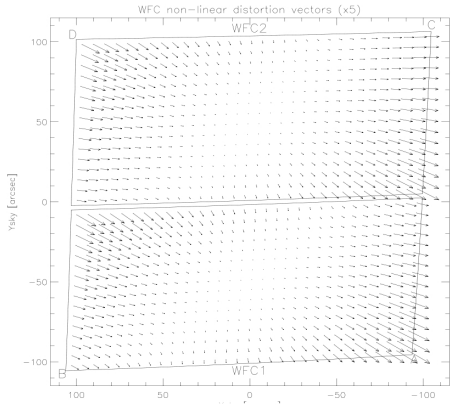

Figure 5.3: Non-linear component of the ACS distortion for the WFC detector using a F475W quadratic fit. The vectors show the degree of distortion (Gonzaga et al., 2011).

5.2.2 Point Spread Function of ACS

As mentioned before, the design of the telescope can cause spreading of the true brightness distribution on the sky. The CCDs in the WFC suffer from a halo due to the red photons that pass through the instrument and are scattered back into the substrates of the detector in wavelength higher than 9000Å. Figure 5.5 shows a simulated PSF image for WFC. The TinyTim software produces a PSF image for the HST (Krist, J., 1993). Although this software is taking into account the HST optical aberrations and obscuration as well as the CCD pixel response function,

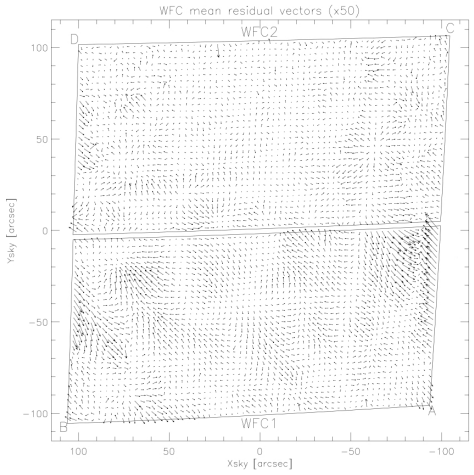

Figure 5.4: Binned residuals to quadratic distortion correction model fits for the ACS WFC detector, for F475W. These residuals are dependent to the filter used (Gonzaga et al., 2011).

but the real PSF will differ from the model because of the jitter in the HST pointing, HST focus variation (focus breathing), and other instrumental effects (Krist, J., 1993). Table 5.1 gives the ACS model PSFs in the central 5×5 pixel region. The listed numbers are the fraction of the total energy received in each pixel. The models have been generated using TinyTim.

Table 5.1: WFC model PSF for filter F814W in the central 5×5 pixel region

0.01	0.01	0.02	0.01	0.01
0.01	0.03	0.07	0.03	0.02
0.02	0.07	0.18	0.07	0.02
0.01	0.03	0.07	0.03	0.01
0.01	0.02	0.02	0.01	0.00

5.2.3 CCD Pixel Response Function

The sharpness of the CCD PSF is degraded by charge diffusion into adjacent pixels. This effect is described as Pixel Response Function (PRF) which shows the distribution of the flux in the neighbouring pixels. This causes a loss of ~ 0.2 magnitude for long wavelengths and ~ 0.5 for short wavelengths in the WFC. Since the thickness is not the same all over the detector, this charge diffusion is not a constant in the field of view. In different filters this PRF is represented

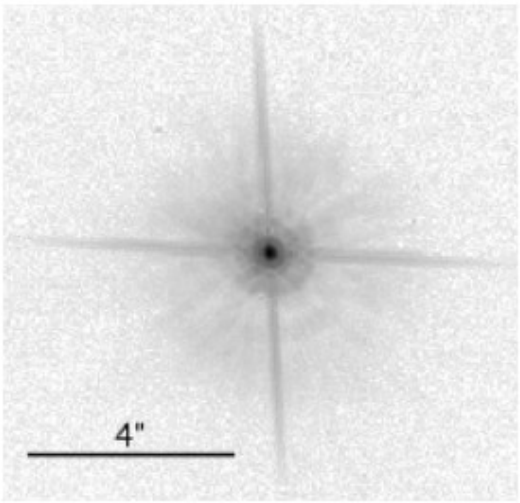

Figure 5.5: This image is a simulated PSF of the WFC in filter F625W with TinyTim (Gonzaga et al., 2011).

by a kernel for the center of the field which for the wavelength of 8000Å close to F814W (which B0218+357 was observed) is

$$K_{\mathrm{WFC}} = \begin{bmatrix} 0.02 & 0.07 & 0.02 \\ 0.07 & 0.64 & 0.07 \\ 0.02 & 0.07 & 0.02 \end{bmatrix}. \tag{5.5}$$

This shows that 64% of the photons receiving a pixel are recorded in the central pixel and the rest will be distributed over the surrounding pixels.

5.3 Drizzle Algorithm

Drizzle was originally developed to combine the dithered images of the Hubble Deep Field North (HDF-N) project taken with the WFPC2 camera and it is the common method used for combining the images of other HST cameras. Drizzle has the ability to work with the images with the arbitrary shifts, rotations and geometric distortion.

In this method, pixels of the original image are mapped into the pixels in the sub-sampled output grid. In this mapping, shifts and rotations between the images and correction of optical distortion of the camera are included. Also there is an option for the user to shrink the pixel before it is averaged into the output image, as shown in figure 5.6.

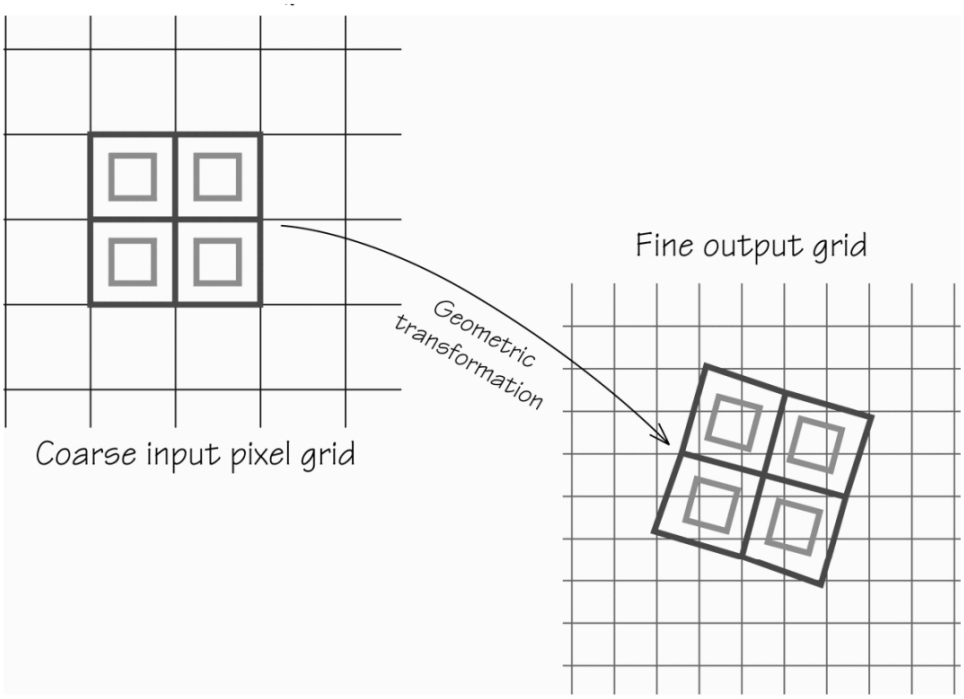

Figure 5.6: Schematic view of how Drizzle algorithm works (Fruchter et al., 1996).

Thereafter these new shrunken pixels or 'drops' are distributed over the sub-sampled output image. The values of input pixels are averaged into the output pixel by concerning a weight proportional to the overlapping area between the drop and the output pixel. In the last step, the sub-sampled image is mapped back to the CCD plane. To evaluate the value of each pixel in the final output image, this method uses the overlapping area $a_{x_i y_i x_o y_o}$ between the final output image pixel and the corresponding pixel with the value of $d_{x_i y_i}$ in the image i such that the value of the pixel in the final output grid is proportional to $a_{x_i y_i x_o y_o}$. Therefore the value of the pixel $I_{x_o y_o}$ and its weight $W_{x_o y_o}$ in the output image becomes

$$\begin{aligned} W_{x_o y_o} &= \sum_i a_{x_i y_i x_o y_o} w_{x_i y_i} \\ I_{x_o y_o} &= \sum_i \frac{d_{x_i y_i} a_{x_i y_i x_o y_o} w_{x_i y_i} s^2}{W_{x_o y_o}}, \end{aligned} \tag{5.6}$$

where s^2 is a factor to conserve surface intensity. The rain down of the drops produces an additional convolution to the output image. In general this method can be explained as

$$I^{\text{Drizzle}} = L_D \star D^{-1} \otimes I^{\text{obs}} \tag{5.7}$$

where L_D represents the additional convolution.

The advantage of this method is that it preserves the absolute surface and point source photometry. Also due to cosmic rays or bad pixels some pixels of the output image might receive no information but as long as there are enough input dithered images these gaps will be filled by them[1]. In figure 5.7 the same procedure provided in figure 5.1 is shown but the combined image is made by Drizzle.

[1]For more detail the reader is encouraged to read the MultiDrizzle handbook (Fruchter et al., 2009)

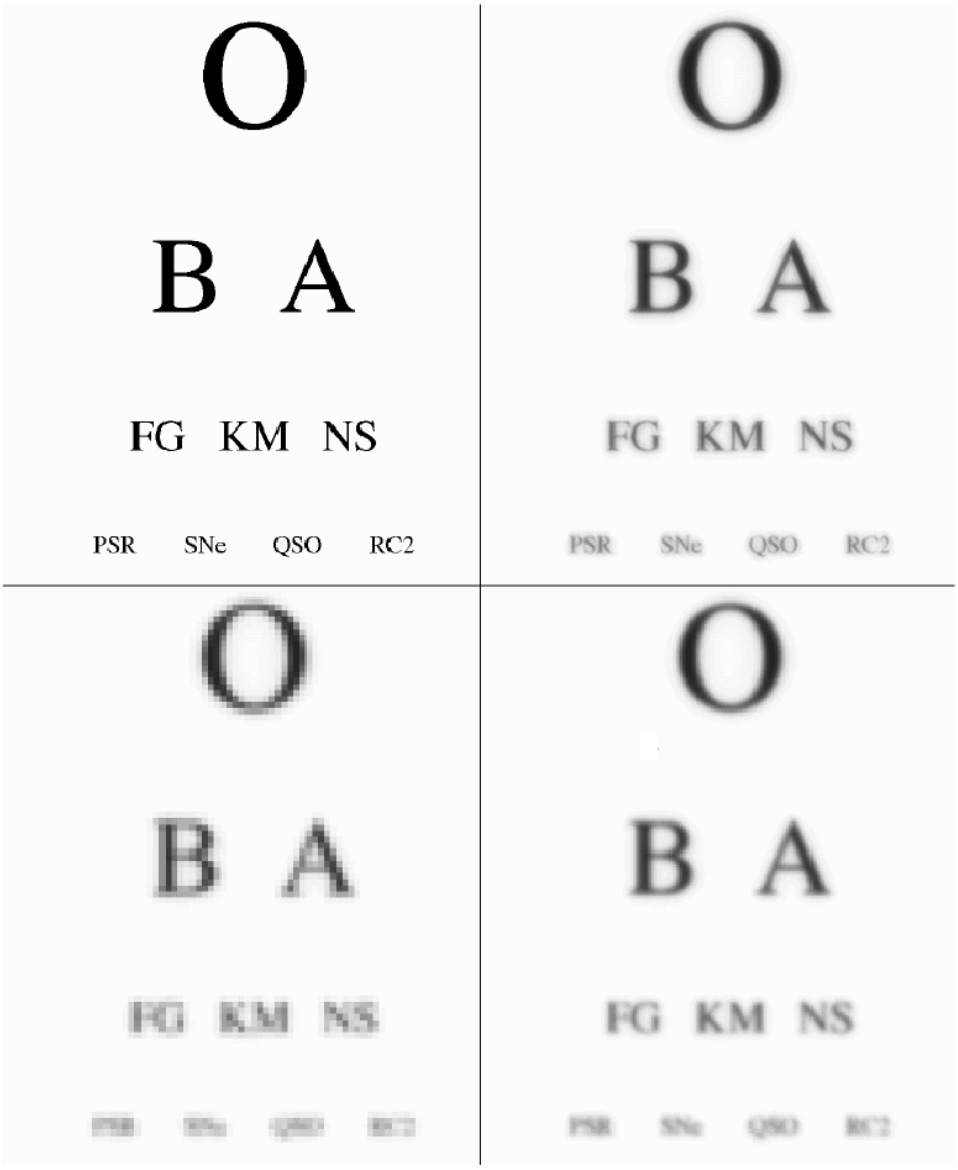

Figure 5.7: Upper left image is the image of the source, upper right shows the light distribution convolved with PSF, lower left image shows the image that we see from the CCD and lower right is the image reconstructed with Drizzle method (Frucher et al., 1996).

The Drizzle algorithm was developed to work with small, faint and partially resolved sources and it can fail to reconstruct a proper result if we want to work on high signal-to-noise unresolved objects. Also Drizzle adds small high-frequency artefacts to the image (figure 5.8). For the case where preservation of the true PSF is essential, one might prefer an algorithm producing highest frequency features more accurately in the image.

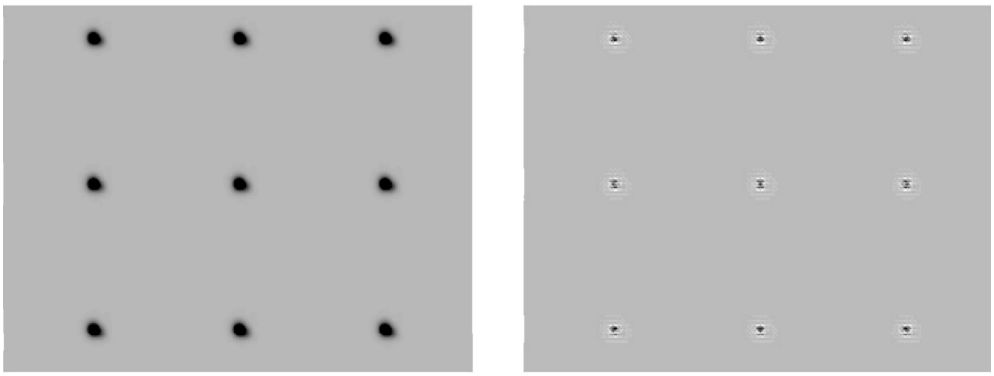

Figure 5.8: On the left a series of synthetic PSFs for the HST/ACS. On the right, a drizzled approximation of the image subtracted from the original (Fruchter, 2011).

6

Direct Model Fitting to Combine Dithered ACS Images

As discussed in the previous chapter, Drizzle produces an additional PSF on the combined image during the image reconstruction which is inconvenient for some systems like B0218+357 in which the precise PSF subtraction to find the positions is essential. In this chapter, the direct fitting method that is developed to combine the ACS/WFC images is introduced. For this reason in the following sections some mathematical aspects of this method are reviewed. Later, the method and its effects in a mathematical and practical form are described and in the last part the results of this method on simulated and real data are presented.

6.1 Least-Squares Method as Maximum Likelihood Estimator

A standard method to find a solution of the systems is the least-squares which minimizes the sum of the errors (squared residuals or in other words the square of the difference between the value provided by the model and the observed value) made in solving every single equation. The main application of this method is 'data fitting'.

There are two types of categories in least-squares method, linear and non-linear. The approach of this method in the linear regime is when the parameters of the model are linearly related to the measured parameters. This type has a closed-form solution (expression) which can be reached in a finite number of steps. The non-linear problem has no closed-form solution and it can be solved by iterative refinement which means at each iteration the method is approximated by a linear system.

Lets assume that we have a data set of n points (x_i, y_i) where x_i is independent variable and y_i is dependent one which comes from observation. The model has the form of $f(x,\beta)$ where β is a vector of free parameters in the model.

To find the best fit of the model f on our data set, we employ least-squares. In this method, the function S as defined below must achieve its minimum.

$$S = \sum_{i=1}^{n} \frac{(f(x_i,\beta) - y_i)^2}{\sigma_i^2} \tag{6.1}$$

where σ_i is the error in each observed data point. In subsection 6.1.1 and section 6.2, we follow the book by Press et al. (1992).

6.1.1 BFGS minimization method

In the first step, we assume an arbitrary function $f(\mathrm{x})$[1] which can be locally approximated by

$$f(\mathrm{x}) \approx c - \mathrm{b}\cdot\mathrm{x} + \frac{1}{2}\mathrm{x}\cdot\mathrm{A}\cdot\mathrm{x}. \tag{6.2}$$

where Roman letters represent vectors and matrices. The values of the parameters b and A are unknown. The basic idea in quasi-Newtonian methods like Broyden-Fletcher-Goldfarb-Shanno (BFGS) is to have a good approximation on the inverse Hessian matrix A^{-1}, that is, to build a sequence of matrices H_i with the property of

$$\lim_{i\to\infty} \mathrm{H}_i = \mathrm{A}^{-1}. \tag{6.3}$$

We want to achieve the limit in N iterations instead of ∞. Near the current point x_i (value of variable x at the iteration i) we have

$$f(\mathrm{x}) = f(\mathrm{x}_i) + (\mathrm{x}-\mathrm{x}_i)\cdot\nabla f(\mathrm{x}_i) + \frac{1}{2}(\mathrm{x}-\mathrm{x}_i)\cdot\mathrm{A}\cdot(\mathrm{x}-\mathrm{x}_i), \tag{6.4}$$

thus

$$\nabla f(\mathrm{x}) = \nabla f(\mathrm{x}_i) + \mathrm{A}\cdot(\mathrm{x}-\mathrm{x}_i). \tag{6.5}$$

In the Newtonian approach $\nabla f(\mathrm{x})$ is set to zero to enable us to determine the next iteration points

$$\mathrm{x} - \mathrm{x}_i = -\mathrm{A}^{-1}\cdot\nabla f(\mathrm{x}_i). \tag{6.6}$$

[1]This function is an arbitrary function for explaining the minimization process in this section and it is not related to the f we define later in the direct fitting method.

The left-hand side of this equation is the steps we take and the right-hand side is known when we accumulate an accurate $\mathrm{H} \approx \mathrm{A}^{-1}$. In quasi-Newtonian the actual Hessian is not needed but its current approximation. If we assume a descent directions of f at x_i with directions P, then $\nabla f(\mathrm{x}_i) \cdot \mathrm{P} < 0$. Therefore, for the Newtonian method we have

$$\nabla f(\mathrm{x}_i) \cdot (\mathrm{x} - \mathrm{x}_i) = -(\mathrm{x} - \mathrm{x}_i) \cdot \mathrm{A} \cdot (\mathrm{x} - \mathrm{x}_i) < 0. \tag{6.7}$$

This requires A to be positive-definite. In the quasi-Newtonian approach the basis is to start with the points with positive-definite and symmetric approximation to A to be sure to have the correct direction for minimization (although the real Hessian A may not be positive-definite). Now if we subtract the Eq. (6.6) for two points of x_{i+1} and x_i we have

$$\mathrm{x}_{i+1} - \mathrm{x}_i = \mathrm{A}^{-1} \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i)) \,. \tag{6.8}$$

Since we went from x_i to x_{i+1}, we want that the new approximation H_{i+1} satisfies Eq. (6.6)

$$\mathrm{x}_{i+1} - \mathrm{x}_i = \mathrm{H}_{i+1} \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i)) \,. \tag{6.9}$$

If we assume that H_{i+1} is H_i plus 'corrections' and equation (6.9) still holds for H_{i+1}, the formula for H_{i+1} becomes

$$\begin{aligned} \mathrm{H}_{i+1} &= \mathrm{H}_i + \frac{(\mathrm{x}_{i+1} - \mathrm{x}_i) \otimes (\mathrm{x}_{i+1} - \mathrm{x}_i)}{(\mathrm{x}_{i+1} - \mathrm{x}_i) \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))} \\ &- \frac{[\mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))] \otimes [\mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))]}{(\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i)) \cdot \mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))} \\ &+ [(\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i)) \cdot \mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))] \, \mathrm{u} \otimes \mathrm{u}, \end{aligned} \tag{6.10}$$

where $\otimes$ denotes outer or direct product of two vectors and u is defined as

$$\mathrm{u} = \frac{\mathrm{x}_{i+1} - \mathrm{x}_i}{(\mathrm{x}_{i+1} - \mathrm{x}_i) \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))} - \frac{\mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))}{(\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i)) \cdot \mathrm{H}_i \cdot (\nabla f(\mathrm{x}_{i+1}) - \nabla f(\mathrm{x}_i))}. \tag{6.11}$$

In this approach after N iterations the algorithm converges to A^{-1}.

6.2 Interpolation

Interpolation methods are used in numerical analysis in mathematics for constructing new data points within a set of known data points. In science it is frequent to interpolate the value of the function obtained by sampling for an intermediate value of independent variables.

In n-dimensional interpolation we look for an estimate of $y(x_1, x_2, ..., x_n)$ from a table of values of y with respect to n independent variables $x_1, x_2, ..., x_n$. In the case of two-dimensional interpolation we have a matrix of functional values $f[1..m][1..n]$. We have the lists of independent variables $x_1[1..m]$ and $x_2[1..n]$. The relation of these input quantities to an underlying function $y(x_1, x_2)$ is

$$f[j][k] = y(x_1[j], x_2[k]) \tag{6.12}$$

We seek to estimate the function y at some untabulated points (x_1, x_2). An important concept is that of the grid in which the point (x_1, x_2) falls, that is, the four tabulated points that surround the desired interior point as shown in figure 6.1.

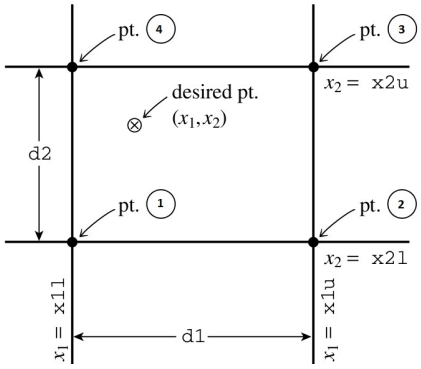

Figure 6.1: Schematic view of a two-dimensional interpolation (Press et al., 1992).

Therefore we define, if

$$\begin{aligned} x_1[j] &\leq x_1 \leq x_1[j+1] \\ x_2[k] &\leq x_2 \leq x_2[k+1], \end{aligned} \tag{6.13}$$

then

$$\begin{aligned} y_1 &\equiv f[j][k] \\ y_2 &\equiv f[j+1][k] \\ y_3 &\equiv f[j+1][k+1] \\ y_4 &\equiv f[j][k+1]. \end{aligned} \tag{6.14}$$

The simplest way of interpolation is bilinear interpolation on the grid square. Its formulas are

$$\begin{aligned} t &\equiv \frac{x_1 - x_1[j]}{x_1[j+1] - x_1[j]} \\ u &\equiv \frac{x_2 - x_2[k]}{x_2[k+1] - x_2[k]} \end{aligned} \tag{6.15}$$

and at the end the interpolated value would be

$$y(x_1, x_2) = (1-t)(1-u)\, y_1 + t\,(1-u)\, y_2 + t\, u\, y_3 + (1-t)\, u\, y_4. \tag{6.16}$$

6.3 Theoretical Aspects of the Method

As mentioned in the previous chapter, the convolution of the true brightness distribution with effects such as PSF and sampling are not accurately known. Therefore, the inverse procedures to deconvolve them from the data can cause additional errors thus may not be a proper solution. As explained before, the Drizzle algorithm may also not be the best option in some specific cases like B0218+357. Here, we introduce an alternative method to combine the observed images. This method is based on model fitting. We use the least-squares method to find the optimum fit for the sky brightness distribution model with observed data. Since we want to overcome the undersampling problem, we use a finer grid for our model. A smoothness constraint is also employed to have a unique solution for the final result. As the size of the output pixels is shrunk in comparison to the input pixels, there would be some unconstrained pixels which are avoided with dithered images and the smoothness constraints.

The model that we define for our sky brightness distribution is a pixelized model in sky coordinates. In this method, the model is compared to all the observed images pixel by pixel. Because the model has a different coordinate system than the observed images we use a mapping function which relates each pixel of the model to the pixel in each observed image. In general, this method minimizes the function below

$$f = \sum_{j=1}^{n_{\rm im}} \sum_{i=1}^{n_{\rm pix}} \frac{\left(I^{\rm mod}_{{\rm int}(j)}[i] - I^{\rm obs}_{(j)}[i]\right)^2}{\sigma^2_{ij}} w_{ij} + \lambda R(n) \tag{6.17}$$

where $n_{\rm im}$ is the total number of images that we want to combine, $n_{\rm pix}$ is the number of CCD pixels per image, $I^{\rm mod}$ is the pixelized sky brightness distribution model in sky coordinates, and $I^{\rm obs}_{(j)}$ the observed image j. The subscript 'int(j)' denotes interpolation to the same grid as the observed image j, taking into account the distortion, shift and rotation in the conversion between CCD coordinates and sky coordinates. $I^{\rm mod}_{{\rm int}(j)}[i]$ is the interpolated value of the model pixel in sky coordinates corresponding to pixel i of image j in Cartesian coordinates. σ_{ij} is the uncertainty of pixel i in image j. The additional weight function w_{ij} is set to zero for flagged or masked bad or unwanted pixels, and to one otherwise. Most important is cosmic ray flagging using the data quality layer of the calibrated and flat fielded, 'flt' data (flt images are provided

by the CALACS pipeline of the HST). The strength of the smoothness constraint is given by the coefficient λ.

$R(n)$ is a quadratic operator that measures non-smoothness, where n denotes the order of derivatives included. Most commonly used are gradient minimization ($n = 1$) and curvature minimization ($n = 2$):

$$R(1) = \sum_{i=1}^{n_{\mathrm{modpix}}} \left[\left(\frac{\partial I^{\mathrm{mod}}}{\partial x}[i] \right)^2 + \left(\frac{\partial I^{\mathrm{mod}}}{\partial y}[i] \right)^2 \right] \tag{6.18}$$

$$R(2) = \sum_{i=1}^{n_{\mathrm{modpix}}} \left(\frac{\partial^2 I^{\mathrm{mod}}}{\partial x^2}[i] + \frac{\partial^2 I^{\mathrm{mod}}}{\partial y^2}[i] \right)^2 . \tag{6.19}$$

Here we are summing over all n_{modpix} model pixels which is generally different from n_{pix}. The derivatives of the discrete model brightness distributions are determined using finite differences[1].

In our work we used $n = 1$ but for the study of the effect of smoothing below we also considered the case of $n = 2$. For minimization of this function, we used the BFGS minimization method.

In the previous parts the general form of this method is described. As mentioned in the previous chapter, the light distribution that one sees from an image from a telescope is the convolution of different functions

$$I^{\mathrm{obs}} = S \cdot R \star D \otimes P \star I^{*}. \tag{6.20}$$

Now the question arises here is that what would be the effect of this method, particularly the regularization part (smoothing constraint), on the combined image. If we assume small pixels (large n_{pix} and n_{modpix}) and the error $\sigma = 1$, we can write Eq. (6.17) with continuous integrals as

$$f = \iint \left[I^{\mathrm{mod}}(x, y) - I^{\mathrm{obs}}(x, y) \right]^2 \mathrm{d}x\,\mathrm{d}y + \lambda R(n) \tag{6.21}$$

with

$$R(1) = \iint \left[\left(\frac{\partial I^{\mathrm{mod}}}{\partial x} \right)^2 + \left(\frac{\partial I^{\mathrm{mod}}}{\partial y} \right)^2 \right] \mathrm{d}x\,\mathrm{d}y, \tag{6.22}$$

$$R(2) = \iint \left(\frac{\partial^2 I^{\mathrm{mod}}}{\partial x^2} + \frac{\partial^2 I^{\mathrm{mod}}}{\partial y^2} \right)^2 \mathrm{d}x\,\mathrm{d}y. \tag{6.23}$$

[1] difference of each pixel with its neighbours

This approach is plausible for the ideal case of fine grids for I^{obs} or proper dither patterns that fill all the gaps. From Fourier transform[1][2] and Parseval's theorem[3] we can rewrite Eq.(6.21) as

$$f = \iint \left|\widehat{I^{\text{mod}}}(u,v) - \widehat{I^{\text{obs}}}(u,v)\right|^2 \,\mathrm{d}u\,\mathrm{d}v + \lambda' \iint k^{2n} \left|\widehat{I^{\text{mod}}}(u,v)\right|^2 \,\mathrm{d}u\,\mathrm{d}v, \tag{6.24}$$

where $k^2 = u^2 + v^2$, $\lambda' = (2\pi)^{2n}\lambda$ and the symbols with hat are the Fourier transforms of those functions. Since we want to minimize this function, the integrand should be minimized. Therefore

$$\delta\left(\left|\widehat{I^{\text{mod}}}(u,v) - \widehat{I^{\text{obs}}}(u,v)\right|^2 + \lambda' k^{2n} \left|\widehat{I^{\text{mod}}}(u,v)\right|^2\right) = 0 \Rightarrow \widehat{I^{\text{mod}}}(u,v) = \frac{\widehat{I^{\text{obs}}}(u,v)}{1+\lambda' k^{2n}}. \tag{6.25}$$

Also from Fourier transform theory we know that the convolution of two functions in Fourier space is

$$\mathcal{F}\{g \star h\} = \hat{g}\hat{h}. \tag{6.26}$$

With this approach our method can act as a convolution function on the data and this function in Fourier space can be shown as

$$\widehat{C}(u,v) = \frac{1}{1+\lambda' k^{2n}}. \tag{6.27}$$

The remaining issue is to inverse this convolution function to the real space

$$C(x,y) = \int_{-\infty}^{\infty} \widehat{C}(u,v) \exp(2\pi \mathrm{i}\, \vec{k}\cdot\vec{x})\,\mathrm{d}u\,\mathrm{d}v\,,\ \vec{k} = (u,v). \tag{6.28}$$

In the following the inverse transform for $n = 1$ which represents the gradient approach and $n = 2$ for curvature is studied.

n=1

For working with gradient constraint in the smoothing function n should be equal to 1. In this case Eq.(6.24) becomes

$$f = \int \left|\widehat{I^{\text{mod}}}(u,v) - \widehat{I^{\text{obs}}}(u,v)\right|^2 \,\mathrm{d}u\,\mathrm{d}v + \lambda' \int k^2 \left|\widehat{I^{\text{mod}}}(u,v)\right|^2 \,\mathrm{d}u\,\mathrm{d}v. \tag{6.29}$$

[1] Definition used here for Fourier transform is: $\mathcal{F}\{f\}(x) = \int_{-\infty}^{\infty} \widehat{f}(k)\, \exp(2\mathrm{i}\pi\, k\, x)\,\mathrm{d}k$

[2] $\mathcal{F}\left\{\frac{d^n f}{dx^n}\right\}(x) = (2\pi\, k)^n\, \widehat{f}(k)$

[3] Parseval's theorem is often written as: $\int_{-\infty}^{\infty} |f(x)|^2\, dx = \int_{-\infty}^{\infty} \left|\widehat{f}(k)\right|^2 dk$

Thus by minimizing the integrand we have

$$\widehat{I^{\mathrm{mod}}}(u,v) = \frac{\widehat{I^{\mathrm{obs}}}(u,v)}{1+\lambda' k^2} \rightarrow \widehat{C}(u,v) = \frac{1}{1+\lambda' k^2}. \tag{6.30}$$

Now the inverse Fourier transform in spherical coordinate gives us

$$C(x,y) = \frac{1}{\lambda'} \int_0^{\infty} \int_0^{2\pi} \frac{k}{\frac{1}{\lambda'}+k^2} \mathrm{e}^{2\pi \mathrm{i}\, k\, r \cos(\theta-\phi)}\, \mathrm{d}k\, \mathrm{d}\theta, \tag{6.31}$$

where $x = r\cos\phi$ and $y = r\sin\phi$. The integral over θ is

$$C(x,y) = \frac{2\pi}{\lambda'} \int_0^{\infty} \frac{k}{\frac{1}{\lambda'}+k^2} \mathrm{J}_0(2\pi\, k\, r)\, \mathrm{d}k. \tag{6.32}$$

This integral is the definition of the *Hankel* transform[1] of

$$\mathrm{H}\left\{\frac{1}{\frac{1}{\lambda'}+k^2}\right\} \tag{6.33}$$

which leads us to

$$C(x,y) = \frac{2\pi}{\lambda'} \mathrm{H}\left\{\frac{1}{\frac{1}{\lambda'}+k^2}\right\}(x,y). \tag{6.34}$$

The Hankel transform of this function from Piessens, R. (2000) is the modified Bessel function of the second kind[2]. Therefore the convolution function becomes

$$C(x,y) = \frac{2\pi}{\lambda} \mathrm{K}_0\left(\frac{2\pi\, r}{\sqrt{\lambda'}}\right). \tag{6.36}$$

The behaviour of this convolution function is shown in figure 6.2

[1] More information of this transform can be found in the Appendix.

[2] The zero order of the modified Bessel function the second kind (for $\Re\{x\} > 0$) is defined as

$$\mathrm{K}_0(x) = \int_0^{\infty} \cos(x \sinh t)\mathrm{d}t. \tag{6.35}$$

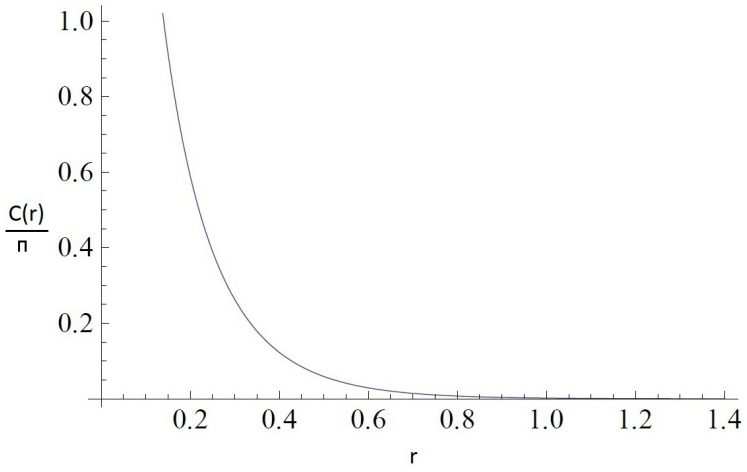

Figure 6.2: In this figure the graph of the convolution function of first order smoothing for $\lambda' = 1$ is plotted.

n=2

For the case of $n = 2$ which represents the curvature approach on giving a smoothing constraint, the Eq.(6.24) becomes

$$f = \int \left|\widehat{I^{\text{mod}}}(u,v) - \widehat{I^{\text{obs}}}(u,v)\right|^2 \mathrm{d}u\,\mathrm{d}v + \lambda' \int k^4 \left|\widehat{I^{\text{mod}}}(u,v)\right|^2 \mathrm{d}u\,\mathrm{d}v. \tag{6.37}$$

As previously mentioned for the case $n = 1$ we should minimize the integrand which leads us to

$$\widehat{I^{\text{mod}}}(u,v) = \frac{\widehat{I^{\text{obs}}}(u,v)}{1+\lambda' k^4} \rightarrow \widehat{C}(u,v) = \frac{1}{1+\lambda' k^4}. \tag{6.38}$$

From inverse Fourier transform in spherical coordinate we have

$$C(x,y) = \frac{1}{\lambda'} \iint \frac{k}{\frac{1}{\lambda'}+k^4} \mathrm{e}^{2\pi \mathrm{i}\, k\, r \cos(\theta-\phi)}\,\mathrm{d}k\,\mathrm{d}\theta \tag{6.39}$$

where $x = r\cos\phi$ and $y = r\sin\phi$. Like the case of $n = 1$ equation (6.39) contains the definition of *Hankel* transform of the function $C(u,v)$ or

$$\mathrm{H}\left\{\frac{1}{\frac{1}{\lambda'}+k^4}\right\}. \tag{6.40}$$

Thus the convolution function can be written as

$$C(x,y) = \frac{2\pi}{\lambda'}\mathrm{H}\left\{\frac{1}{\frac{1}{\lambda'}+k^4}\right\}. \tag{6.41}$$

Again from Piessens, R. (2000) the Hankel transform of this function is *Kelvin* function Kei_0[1]. Thus the convolution function becomes

$$C(x,y) = \frac{-2\pi}{\sqrt{\lambda'}}\mathrm{Kei}_0\left(\frac{2\pi\, r}{\sqrt[4]{\lambda'}}\right). \tag{6.43}$$

The behaviour of this convolution function is shown in figure 6.3. As can be seen between the two figures 6.2 and 6.3 the approach of $n = 1$ has a singularity at the centre of the distribution.

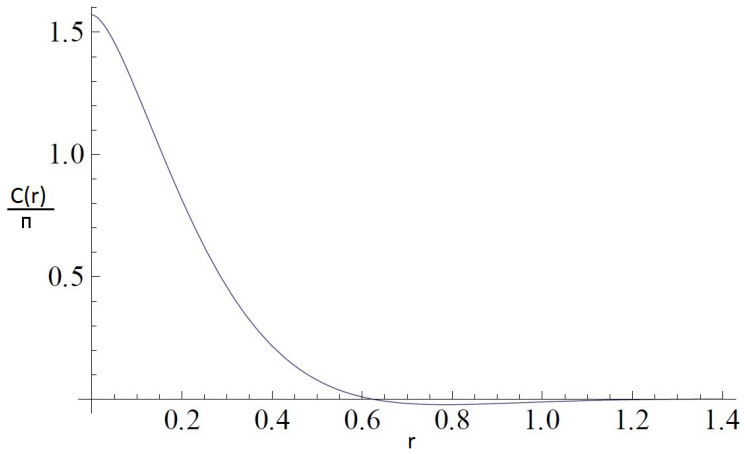

Figure 6.3: In this figure the convolution function produced by the smoothing for $\lambda' = 1$ is plotted for the case $n = 2$.

6.4 Practical view of the method

In the first step, a grid with shrunken pixel size in the sky coordinates is created. After this, each pixel of all the input images are mapped to a (or maybe more) pixel(s) in this grid (geometric distortion correction is included in this mapping). This gridding is done in such that the first pixel in the grid represents the lowest right-ascension (RA) and declination (Dec) and the last pixel of it corresponds to the highest values of right-ascension and declination. This lowest and

[1]The function $\mathrm{Kei}_0(x)$ is the imaginary part of the zero order modified Bessel function of the second kind. This functions is named after 'William Thomson', first Baron Kelvin. The argument x is taken to be real. $\mathrm{Kei}_0(x)$ is defined as

$$\mathrm{Kei}_0(x) = \Im\{e^{\frac{-\alpha\pi i}{2}}\mathrm{K}_0(x\, e^{\frac{\pi i}{4}})\}. \tag{6.42}$$

highest values of right-ascension and declination are determined by comparing the lowest and highest RA and DEC in each observed image in the data set. The gridding is done as

$$RA_i = RA_{\min} + \left(\frac{\text{pixel size} \times m}{\cos(DEC_i)} \right) \tag{6.44}$$

$$DEC_i = DEC_{\min} + \text{pixel size} \times n, \tag{6.45}$$

where RA_i, DEC_i are the right-ascension and declination of the pixel i and $RA_{\min}$, $DEC_{\min}$ are the minimum values of right-ascension and declination. pixel size in these formulas is the pixel size of the model image which is preferred to be smaller than pixel size of observed image. n and m are the positions of the pixel in the model image corresponding to pixel i. Therefore the total number of pixels in model image is

$$\text{nmodpix} = n_{\max} \times m_{\max} \tag{6.46}$$

where $n_{\max}$ and $m_{\max}$ the maximum number of pixels in each row and column of the model image. Due to lack of memory for running this code, right-ascension and declination, mapping of the pixel of the observed image to the pixel in the model image and interpolation coefficient for each pixel are stored as arrays and added as additional layers to the original observed images. This was done in python using the 'pyfits' library which is made to work with fits files in python environment.

As mentioned in the description of f, to combine the observed data, one has to correct for cosmic-ray effects which is an extension in the observed images of the ACS with different values. If the pixel has the value of 4096, it means that this pixel is affected by a cosmic-ray and its weight (w_{ij}) is set to zero. Apart from this extension, there is another extension in the observed image corresponding to uncertainties of each pixel in the image. The error values in this extension are used as σ_{ij} in χ^2. These extensions are provided in 'flt' images of the ACS (these data are calibrated with the CALACS pipeline of the HST).

The minimization process is written in the C programming language using the 'gsl' library (Galassi et al., 2009) for the optimization of a function and the BFGS method which is described in previous subsections is chosen for minimization procedure. Despite the fact that the least-squares in this method is in the linear regime, due to the higher speed of the algorithm, we decided to use BFGS which is for non-linear regime. This minimization library uses the function and its first derivative to compute the minimum of the function. The convergence limit of this method and pixel size for the model image and smoothing coefficient λ is given by the user.

6.5 The results of applying the direct fitting method on data

In this section the results of applying the fitting method on simulated and real data are presented. First the results on the simulated data produced with the use of header information of the ACS images is discussed. Then in the following the results of applying this method on observed data are explained. The observed data is two strong lensing systems B1608+656 and B0218+357 (the result of B0218+357 is provided with more detail in the next chapter).

Simulated data

To test the method, 20 simulated images based on the observational information of B0218+357 (pixel size, dither pattern, reference point, orientation and geometric distortion correction coefficients) are produced. These images are created with the 'pyfits' library in Python. For sources in the images, two sources with Gaussian brightness distribution are simulated. The fainter image has FWHM equal to 2.354 with the amplitude of 100 and the brighter image has 3.531 with the amplitude of 150. The background has the fix value of 35. For the noise in these data we used the Poisson distribution in which σ^2 of each pixel is equal to the value of that pixel. The pixel size of the combined image is 30 mas.

To choose the right smoothing coefficient λ for the simulated data, the method is applied on simulated data with different values of λ. We used the Gaussian distribution function of 'galfit' software (Peng et al., 2010) to model the two images in the combined image (for more detail see Appendix B). As can be seen from figure 6.4 the right image has a better residual image than the left hand side of the figure. This shows that smaller values of λ provide us the closer model for brightness distribution of the images to the Gaussian distribution.

Another test on this method is to see if it preserves the separation between the images. We fix the separation between two simulated objects to 0.7235 arcseconds. Thus the results of galfit (positions of the centre of two modelled sources) are used and compared with the 'true' separation of the sources. In table 6.1 the differences between the separations are given. The lower the value of λ, the more precise value we get from the direct fitting method.

As can be seen from the table 6.1 and figure 6.4, λ equal to 10^{-5} gives us better results. The pixel size of the input images are the chosen to be 50 mas (the same as the WFC detector). Table 6.1 shows that our method can preserve the astrometry of the objects with accuracy of far less than a pixel size.

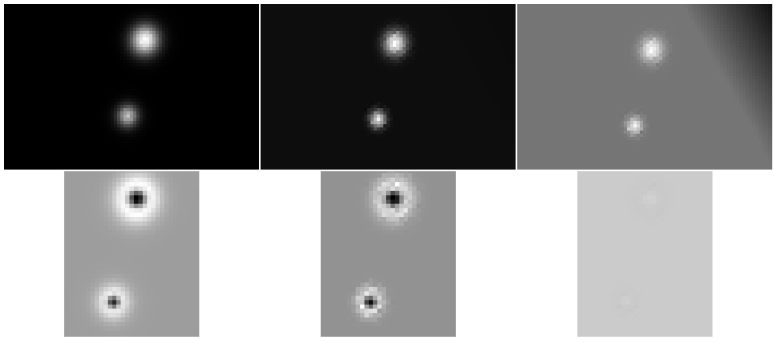

Figure 6.4: The upper images are the reconstructed images of the fitting method with the smoothing coefficient from left equal to $10^{-1}, 10^{-3}$ and 10^{-5}. The bottom images are the residual images from upper images above and the model produced with galfit with the fixed parameters on Gaussian distribution for sources.

Table 6.1: Comparison between the true separation of 0.7235 and the separation determined by galfit

Smoothing coefficient λ	separation	Difference in separation (in arcsec)
10^{-1}	0.7268	0.0032
10^{-2}	0.7260	0.0025
10^{-3}	0.7257	0.0022
10^{-4}	0.7251	0.0015
10^{-5}	0.7238	0.0002

B1608+656

The quadruple-image gravitational lens B1608+656 was discovered in the Cosmic Lens All-Sky Survey (CLASS) (Myers et al., 1995; Browne et al., 2003; Myers et al., 2003). Figure 6.5 is an image of B1608+656, showing the spatially extended source surface brightness distribution (with lensed images labeled by A, B, C, and D) with two lens galaxies G1 and G2.

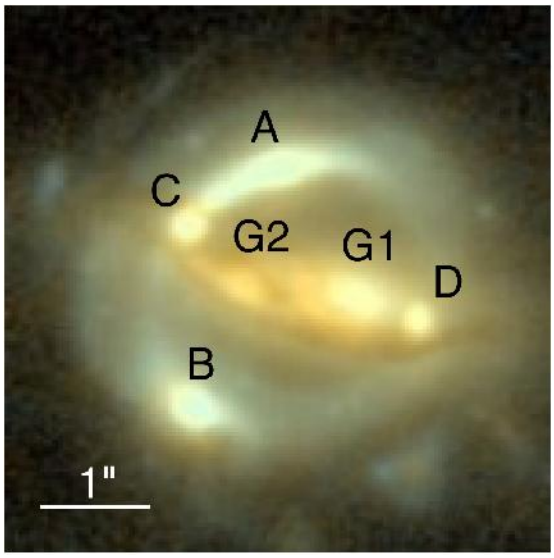

Figure 6.5: HST ACS image of B1608+656. The lensed images of the source galaxy are labelled by A, B, C, and D and the two lens galaxies are G1 and G2. 1 arcsec corresponds to approximately 7 kpc at the redshift of the lens- from Suyu et al. (2010)

The redshift of the lens galaxies is $z = 0.6304$ (Myers et al., 1995) and the redshift of the source is $z = 1.394$ (Fassnacht et al., 1996). The lens galaxies are in a group with all galaxy members lying within $\pm 300\,\mathrm{km\,s^{-1}}$ of the mean redshift (Fassnacht et al., 2006a). The separation between the images in table 6.2 are provided by CASTLES website (Kochanek et al.) from radio observations.

Table 6.2: Separation between images in mas

components	A	B	C	D
A	-	2095.0	879.8	1677.6
B	2095.0	-	1500.0	2003.8
C	879.8	1500.0	-	2035.4
D	1677.6	2003.8	2035.4	-

In this work, 44 images of this system observed with the ACS/WFC camera in the filter F814W are chosen. To have an estimate about an optimum value of λ, we plot λ versus χ^2 for three different pixel sizes of the output image.

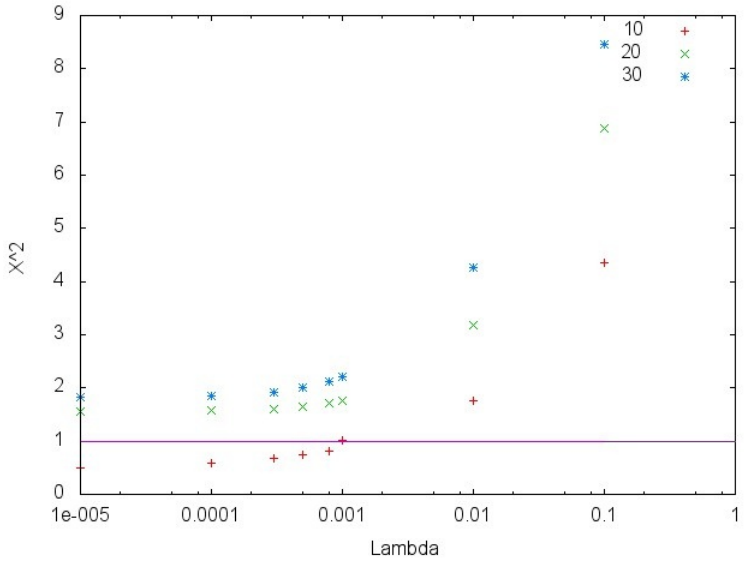

Figure 6.6: The changes of reduced χ^2 as a function of λ is presented in this plot. Different colours and symbols represent pixel sizes from 10 to 30 mas.

As can be seen from the plot in Fig. 6.6, for a pixel size of 10 mas we find that $\lambda = 10^{-3}$ results in a reduced χ^2 of 1.029, very close to the expected 1. Generally smaller values for λ produce a better fit, as expected. However, for pixel sizes of 20 and 30 mas, χ^2 never drops below unity. In these cases the large pixels themselves serve as additional regularization that is too strong to achieve a reduced χ^2 of unity.

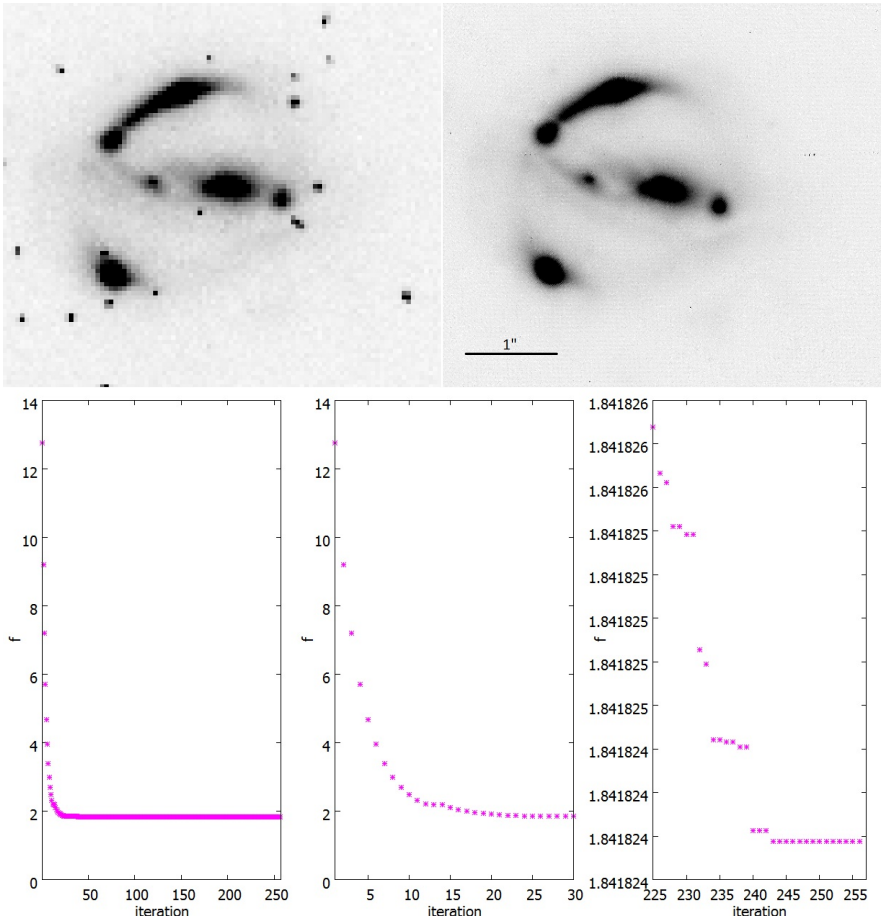

Figure 6.7: In the upper image, the left is one of the raw images from ACS and on the right is the combined image from our method with the pixel size of 10 mas. The bottom left plot shows the minimization process of f, the middle plot is the zoom for first 30 iteration the right plot is the shows the iteration 225 till the end of minimization. The reduced χ^2 for the combined image is 1.029.

The separation of the quad images in the combined image from direct fitting method is provided in table 6.3. To estimate the positions of these images we chose the position of the image to be the pixel with the largest value among the pixels in each image and no PSF is subtracted which can also cause uncertainty to the true position of the images. The optical images are not point like but extended and distorted images of a galaxy so the positions cannot be determined more accurately. With this uncertainty in the determination of the positions, the separation of the images in the output image is still around 1 pixel.

Table 6.3: Separation between images in mas- in parenthesis the difference of each separation from table 6.2 is given.

components	A	B	C	D
A	-	2103.30 (8.3)	882.47 (2.67)	1689.12 (11.52)
B	2103.30 (8.3)	-	1500.0 (0)	2003.21 (−0.59)
C	882.47 (2.67)	1500.0 (0)	-	2036.56 (1.16)
D	1689.12 (11.52)	2003.21 (−0.59)	2036.56 (1.16)	-

7

The Analysis of B0218+357 for the Determination of H_0

In the previous parts, we presented the aim of this project and the scientific knowledge needed to understand the path we should take. In this chapter the first results of the direct fitting method on the ACS/WFC data of the strong lensing system B0218+357, PSF subtraction from these results using 'galfit' (a brief introduction on galfit is provided in Appendix B) and at the end the determination of the Hubble constant are provided.

7.1 The Direct Fitting Method on B0218+357

As mentioned in chapter 4, B0218+357 was observed in 7 visits. To combine the images, a pixel size of 30 mas was chosen for the image model with smoothing coefficient $\lambda = 10^{-5}$. The first attempt to combine the data showed multiple images of each object in the field (the pointing information in the exposures were not reliable for different visits). Therefore, to stack the same objects in different exposures on top of each other, during the combination of all different visits, additional shifts were added to the positions of the reference points of the images. These additional shifts were determined by applying the fitting method on each visit separately (no shifts were required in the combination process in each visits) and using the 'sextractor' software (Bertin et al., 1996) to create a catalogue of the objects in each visit. Then the positions of the common objects in all seven catalogues were used to estimate the relative additional shifts for each visit (visit 12 was used as a reference catalogue). Thereupon, the related shifts for each visit were added to all the exposures within the visit. These shifts are

given in table 7.1.

Table 7.1: Added Shifts to the B0218+357 data in arcseconds.

visit nr.	shift in RA	shifts in Dec
10	−1.565	+0.237
11	−1.547	−0.243
12	0.0	0.0
13	−1.580	−0.283
15	−1.601	−0.377
14	+0.437	−2.302
16	+0.485	−2.238

After applying these shifts and combining the images, the minimization was still not satisfactory. Figure 7.1 presents the combined image of the 112 exposures using direct fitting method and the evolution of f with respect to the iterations.

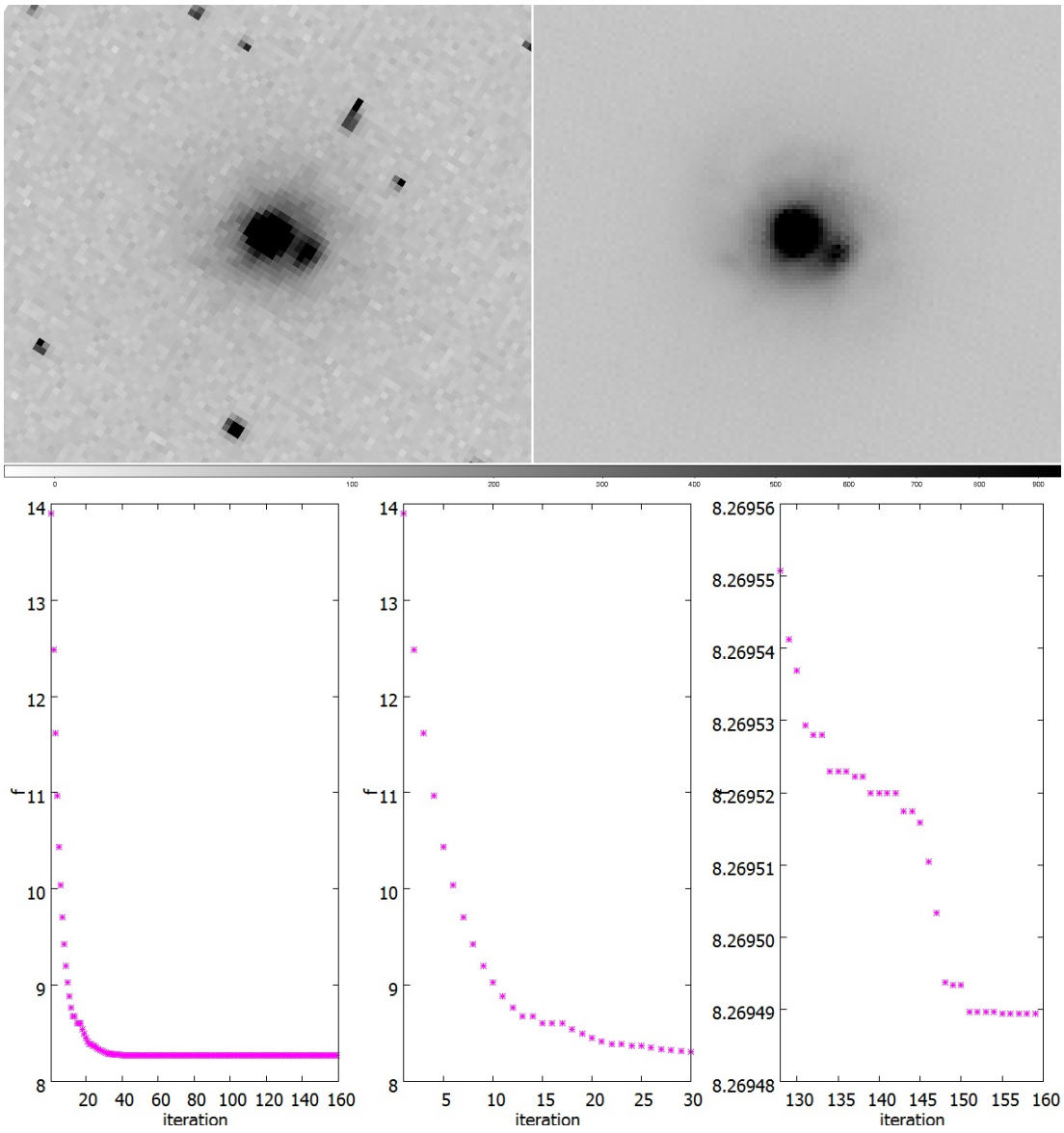

Figure 7.1: Top: the left image is one of the raw images and the right image is the result of the direct fitting method with pixel size of 30 mas and $\lambda = 10^{-5}$. Bottom: the left plot represents the minimization process of f. The middle plot is the zoom of the left plot for the first 30 iterations and the right plot shows the changes of f in the last 30 iterations.

As can be seen from the plots in figure 7.1, f reaches the value of ~ 8.269 with $\chi^2 \sim 8.253$ which indicates that the fitting is not good (for the accurate result the χ^2 should be of the order of 1). Changes in the PSF, geometric distortion of the telescope, sensitivity of the instrument and brightness variation of the lensed images with time are the main reasons that we decided to work on each visit separately (another possible reason is that the pixel size of 30 mas was too large in the case of combining 112 exposures for the χ^2 to converge to the value close to one). Thus we exerted the method on each visit separately with the same parameters. Figures 7.2 and 7.3 present the minimization process of the visit 10 and 15.

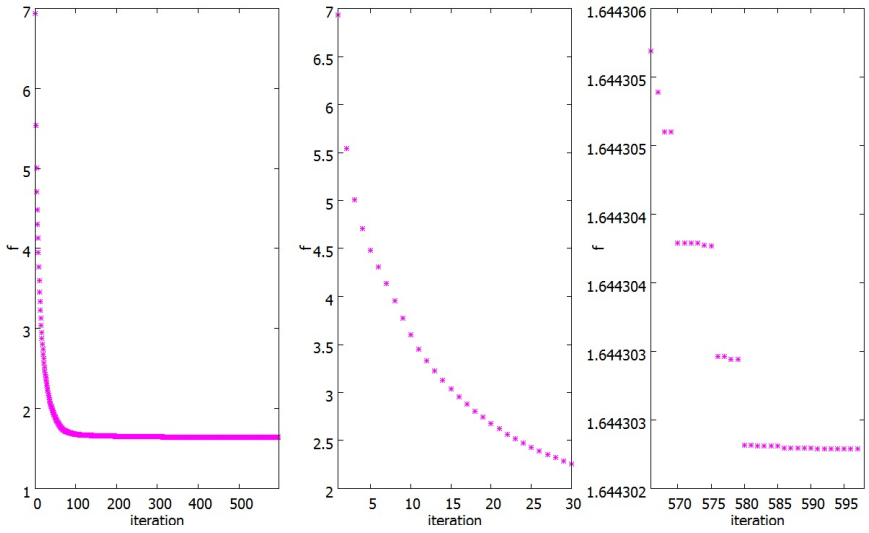

Figure 7.2: visit 10- The left plot represents the minimization process of f. The middle plot is the zoom of the left plot in the first 30 iterations and the right plot shows the changes of f in the last 30 iterations.

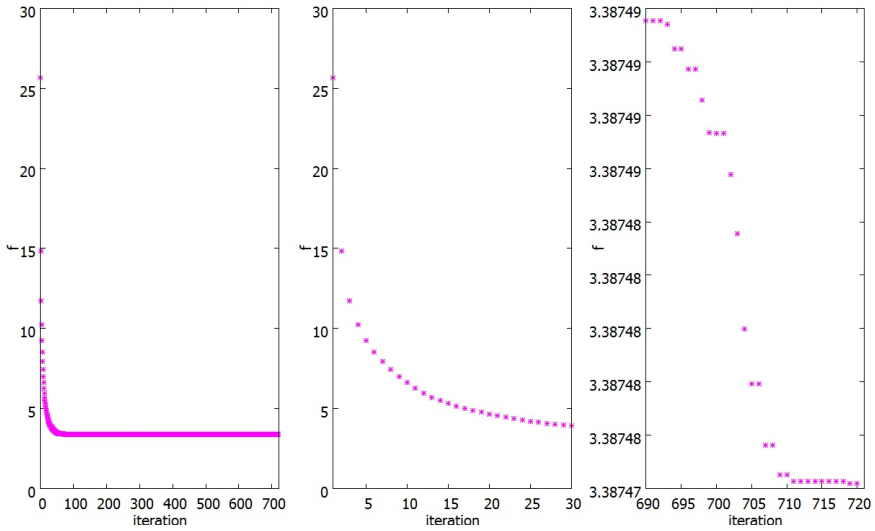

Figure 7.3: visit 15- The left plot represents the minimization process of f. The middle plot is the zoom of the left plot in the first 30 iterations and the right plot shows the changes of f in the last 30 iterations.

In table 7.2, the value of χ^2 and f is provided for each visit. Plots in Fig. 7.3 and values

Table 7.2: The resulting value of the minimized function f and χ^2 for pixel size of 30 mas

visit	f	χ^2
10	1.64	1.21
11	1.33	1.19
12	1.38	1.28
13	1.45	1.24
14	1.29	1.11
15	3.38	3.37
16	1.38	1.31

in table 7.2 reveal that there is a problem in visit 15 which causes the high value of f (which is 3.387 with χ^2 value of 3.376). In the previous work on these data, York et al. (2005) also excluded visit 15 (they claimed that there is contamination of the stray light from the WFPC2 detector calibration lamp with the images in this visit).

To have a higher resolution and reliable results we use smaller pixel size for the output image. Another effect that should be taken into account is the smoothing coefficient λ and how it should be chosen. Plot 7.4 shows the χ^2 values for different pixel sizes and λ. As can be

seen, smaller pixels and proper λ, gives us a χ^2 value close to 1. As an example, choosing pixel size of 10 mas for the output image requires $\lambda = 10^{-3}$ to have χ^2 close to 1.

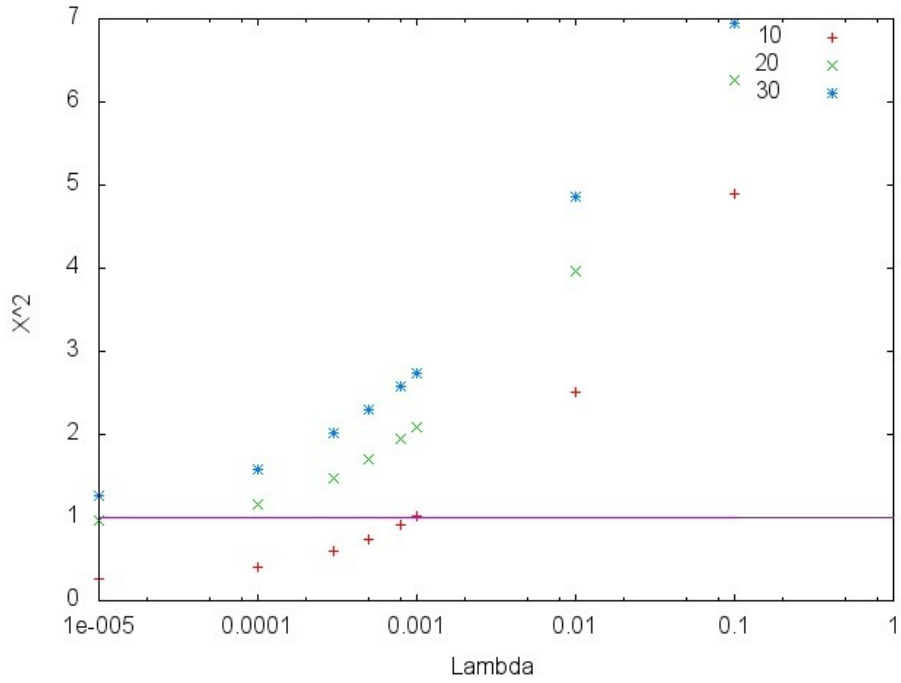

Figure 7.4: Different values of λ versus resulted value of the reduced χ^2 after the minimization process for pixel sizes of 10, 20 and 30 mas for visit 13.

Figure 7.5 presents the combined image of visit 13 using pixel size of 10 mas and $\lambda = 8 \times 10^{-4}$. For this visit the reduced χ^2 is 1.01.

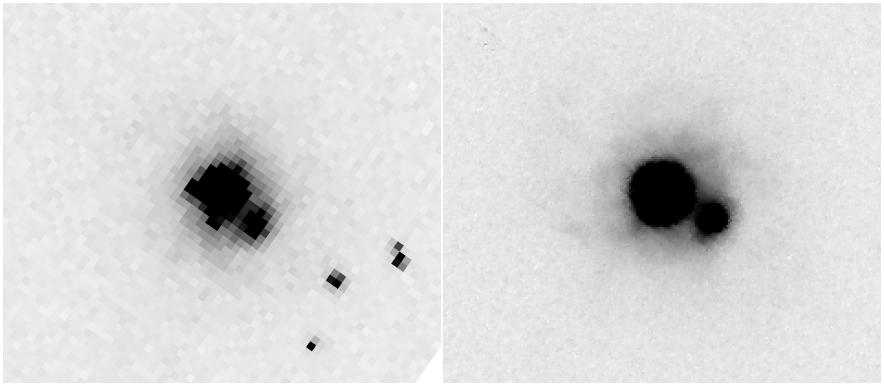

Figure 7.5: visit 13- The left image shows one of the raw images of the visit and right image is the model image produced by direct fitting method with the pixel size of 10 mas and $\lambda = 8 \times 10^{-4}$.

Although the pixel size of 10 mas with $\lambda = 10^{-3}$ gives us a χ^2 closer to 1, primary attempts on the PSF subtraction of the two lensed images of the source were not satisfactory. In the following, to compare our results with previous work of York et al. (2005) on these data, the pixel size of 25 mas is chosen. The value of $\lambda = 10^{-5}$ was chosen for each visit. The values of resulting reduced χ^2 is given in table 7.3

Table 7.3: Chosen value of λ for each visit and the resulting χ^2 for the pixel size of 25 mas

visit	λ	χ^2
10	10^{-5}	1.16
11	10^{-5}	1.12
12	10^{-5}	1.23
13	10^{-5}	1.04
14	10^{-5}	1.03
16	10^{-5}	1.25

7.2 PSF Subtraction and Position Analysis

In the next step, the combined images for each visit are used for the PSF and galaxy fitting with galfit. For the PSF subtraction, one of the stars in the field was chosen as a PSF model. Figure

7.6 shows an example of the PSF subtraction of the two images of B0218+357 with galfit on visit 13.

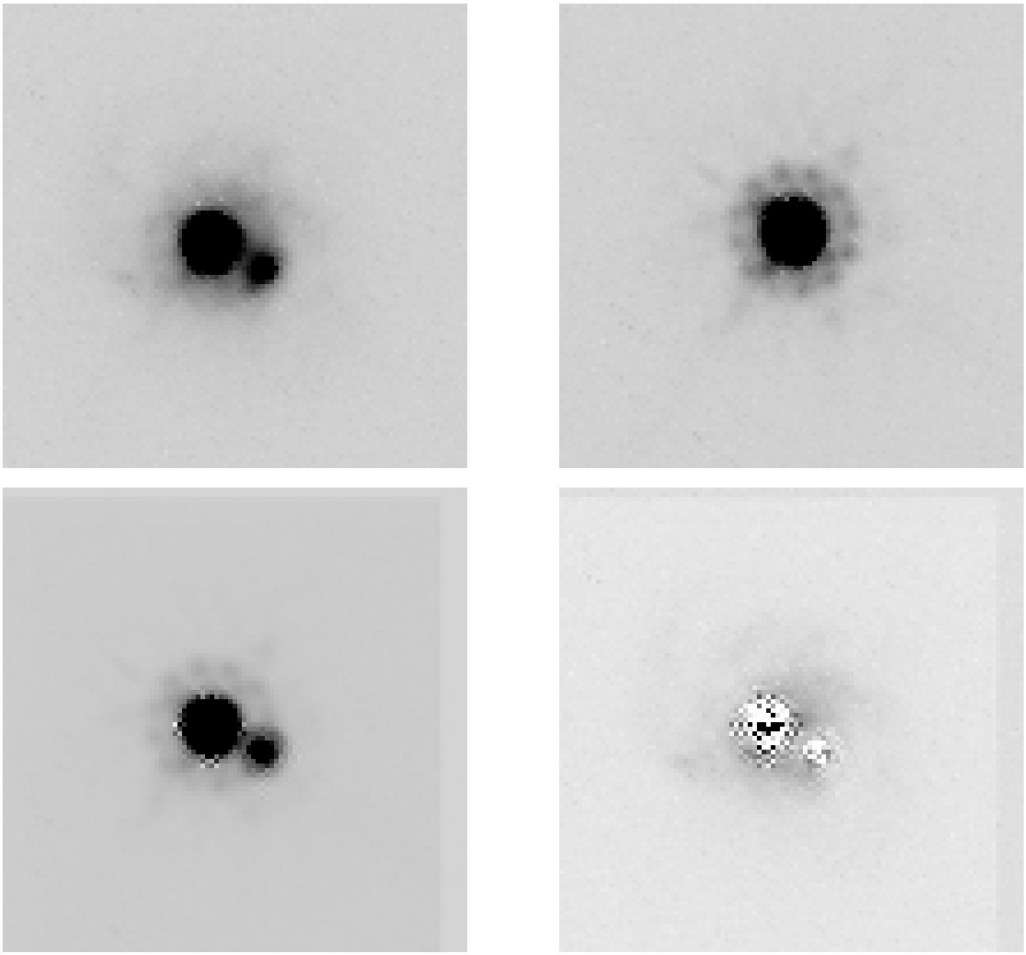

Figure 7.6: The upper left image represents the image model from direct fitting method for visit 13. The upper right image is the star chosen as a PSF model. Bottom left is the model image produced by galfit and the bottom right image is the residual image.

York et al. (2005) determined the Hubble constant in two ways of finding the position of the lens galaxy with and without masking arms of the spiral lens galaxy. In order to compare our results with them, in the first part a Sersic profile in galfit is fitted to B0218+357 to model the lens galaxy[1]. The Sersic profile is used for modelling the brightness distribution of the galaxies and bulges. Sersic index $n = \frac{1}{2}$ gives us a Gaussian distribution and $n = 1$ presents

[1]More information of definition and parameters of Sersic profile is provided in Appendix B.

an exponential profile which is normally used to model spiral galaxies and dwarf elliptical galaxies. In our fitting the index between 0.7 and 0.8 (depending on the visit) gave us the best results.

In the second part we modelled the arms of the lens galaxy with galfit's option to model the arms of the spiral galaxy. Following figures 7.7–7.12 present the optimum results galfit provided for modelling B0218+357.

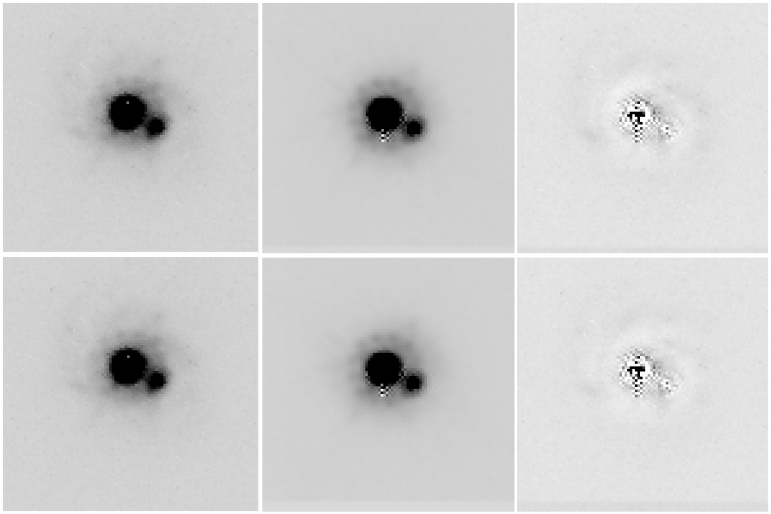

Figure 7.7: visit 10- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

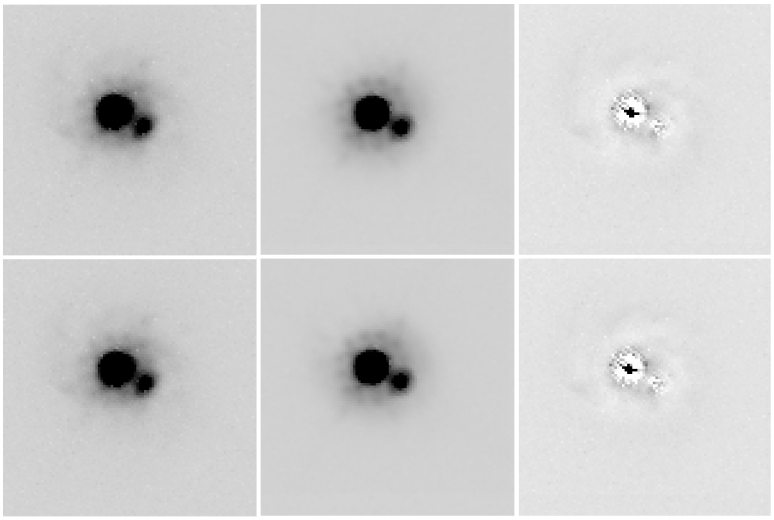

Figure 7.8: visit 11- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

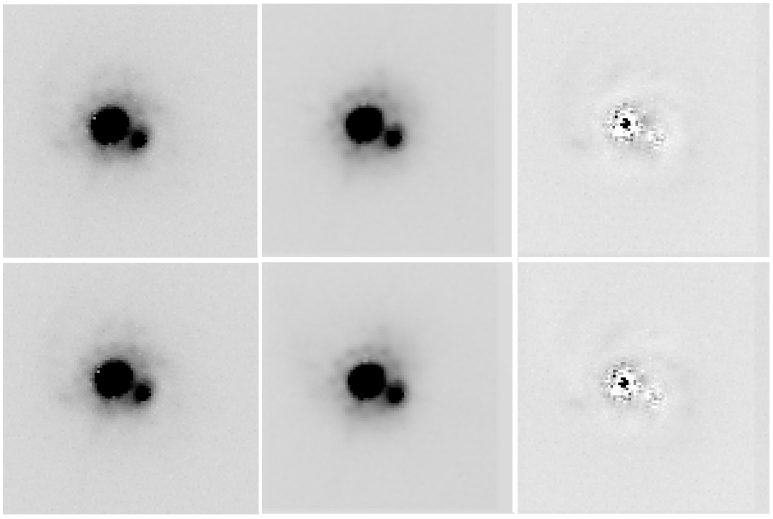

Figure 7.9: visit 12- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

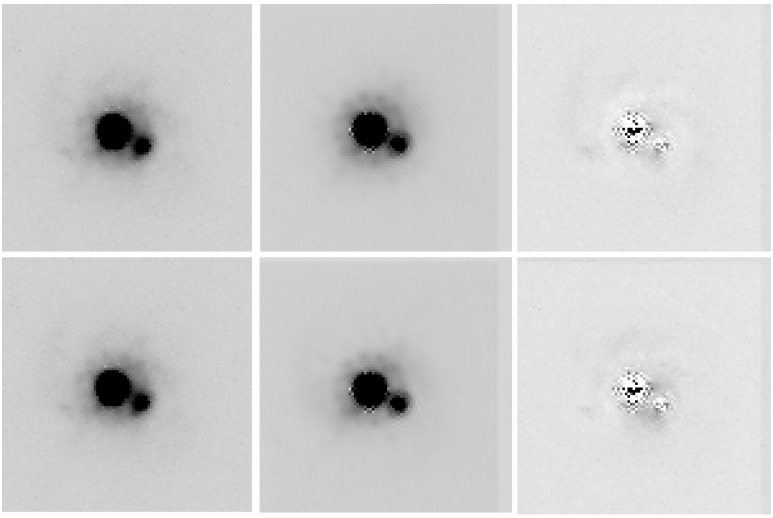

Figure 7.10: visit 13- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

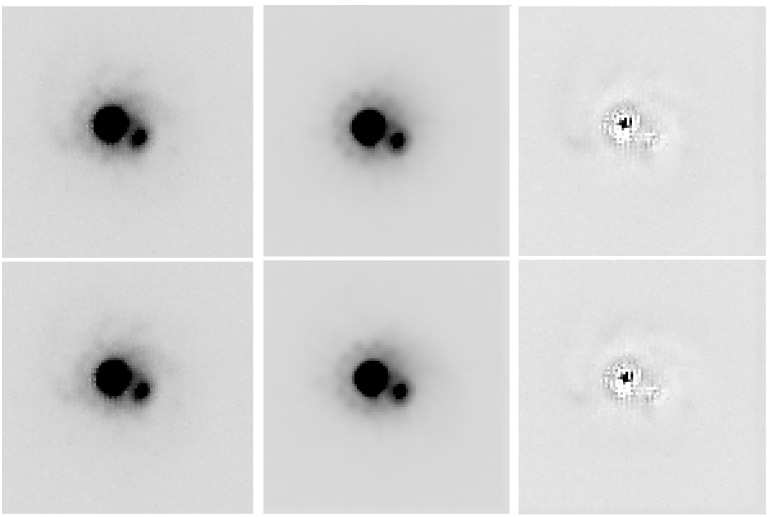

Figure 7.11: visit 14- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

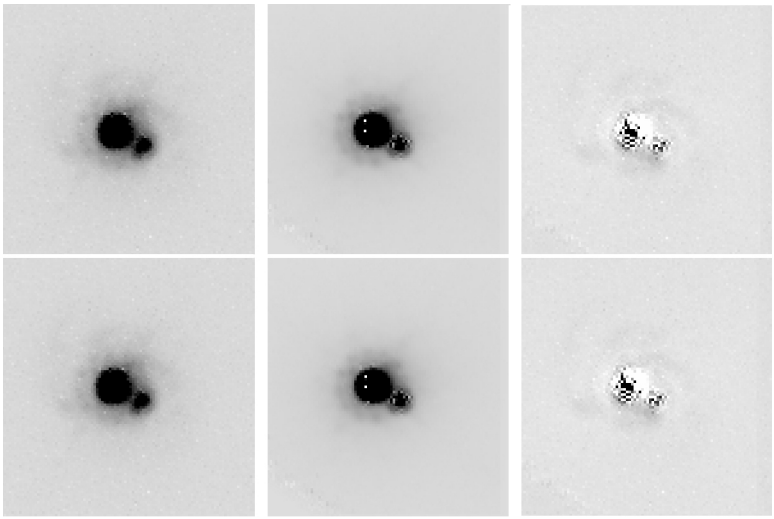

Figure 7.12: visit 16- The upper row presents from left to right the input image of galfit, model image produced by galfit and the residual image for the case when the arms of the spiral galaxy are not fitted. The bottom row images have the same arrangement but the arms are fitted.

In the presented images, the highest values of the black and white areas in the residual images are around 15 percent of the original value of those pixels. In this system, image B and the galaxy are overlapping which causes uncertainties in the fit of the lens galaxy and the two images but as can be seen, most parts of the arms close to the images are fitted.

Table 7.4 presents the positions of the two images with respect to the centre of the spiral lens galaxy from the galfit models of the B0218+357 and the separation of the double image.

Table 7.4: Relative position of images with respect to the lens galaxy centre and image separation. Values in parentheses represent the result of not fitting the arms

visit	image A-G (mas)	image B-G (mas)	image separation (mas)
10	261.64 (264.14)	53.83 (52.70)	314.76 (316.24)
11	276.39 (260.34)	62.49 (66.88)	311.92 (312.00)
12	264.54 (251.35)	47.18 (60.72)	311.65 (311.44)
13	271.79 (252.88)	50.09 (64.98)	313.15 (313.30)
14	257.79 (262.20)	59.62 (57.70)	315.78 (314.67)
16	271.75 (260.57)	51.49 (57.19)	317.66 (323.32)
mean	267.27 (258.58)	54.12 (60.03)	314.15 (315.16)
rms	6.49 (4.76)	5.35 (4.81)	2.14 (3.98)

York et al. (2005) found the mean separation to be 317 ± 2 mas (1σ) when they used the star field. In their work they presented the separation of image B with respect to the lens centre in both right-ascension and declination. Table 7.5 compares the separation of image B from the centre of the lens with York et al. (2005) work. Wucknitz et al. (2004) obtained the separation of the image A and the lens galaxy to be (255 ± 6, 119 ± 4). In our work the separation is (224 ± 10, 123 ± 6). As mentioned in chapter 4, the separation between the images in optical is smaller than radio which has more precise measurements on the position of the images. York et al. (2005) indicated the hypothesis of the existence of a giant molecular cloud in the line of sight of image A which absorb the lights coming through that region. This extinction in image A causes the shift in the position of image A in optical.

Table 7.5: The comparison our result on the separation of image B from the centre of the lens galaxy with York et al (2005). The values in parenthesis represent the result of not fitting/masking the arms (both with and without parenthesis represent $\Delta\alpha$ and $\Delta\delta$ in mas)

visit	our results	York et al. (2005)
10	+79, +14 (+62, +16)	+70, +12 (+50, +6)
11	+70, −26 (+62, +15)	+69, −18 (+60, −4)
12	+54, +18 (+71, +19)	+84, +8 (+59, +9)
13	+68, +1 (+80, +7)	+72, −5 (+54, −2)
14	+71, +14 (+71, +3)	+76, −16 (+79, −14)
16	+61, +3 (+66, +20)	+79, −14 (+61, −6)
mean±rms	+67±8, +4±12 (+69 ±6, 13 ± 6)	+75±6, −6±13 (+57±4, +1±6)

The general form of time delay between images in gravitational lensing as mentioned in chapter 3 is

$$c\Delta t_{i,j} = \frac{1+z_{\rm L}}{H_0}\frac{d_{\rm L}d_{\rm S}}{d_{\rm LS}}(\Phi_i - \Phi_j), \tag{7.1}$$

where $d_{\rm L}$, $d_{\rm S}$ and $d_{\rm LS}$ are the angular size distances normalized to Hubble constant $d_{\rm k} = D_{\rm k}\frac{H_0}{c}$ with k = L, S, LS and

$$\Phi_i = \frac{1}{2}|\nabla\psi(\vec{\theta_i})|^2 - \psi(\vec{\theta_i}), \tag{7.2}$$

where ψ is the potential profile for the lens galaxy. In a general isothermal model for the potential (since B0218+357 is an isolated system, no external shear is needed to be considered (Lehar et al., 2000)), equation (7.1) changes to a simple form of

$$c\Delta t_{i,j} = \frac{1+z_{\rm L}}{2H_0}\frac{d_{\rm L}d_{\rm S}}{d_{\rm LS}}(|\vec{\theta_i} - \vec{\theta_0}|^2 - |\vec{\theta_j} - \vec{\theta_0}|^2), \tag{7.3}$$

where $\vec{\theta}_0$ is the centre of the lens galaxy.

Table 7.6 presents the obtained Hubble constant for each visit assuming an isothermal model for the lens galaxy.

Table 7.6: Hubble constant value for each visit using isothermal

visit	Hubble constant with fitting	Hubble constant without fitting
10	68	69
11	75	67
12	70	61
13	73	61
14	65	67
16	73	70
rms	3	4

By averaging over these values, the mean value for the Hubble constant for an isothermal model when the arms of the spiral lens galaxy are left unfitted is $66 \pm 4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. For the case of fitting the arms, our mean value for the Hubble constant is $70 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Due to the difficulties in the error estimation, in this study the normal rms is used as an error for these results.

As mentioned earlier, because the optical position of image A maybe affected by extinction, measurements of the position of the images in radio are more accurate. Therefore, in the next step we combine our results on the positions of the lens galaxy and image B with the positions measured for image A relative to image B in radio. In order to combine the data, we assumed the difference between the separation of the images are due to the shift in the position of image A in optical data and use the new position of image A to obtain the Hubble constant. Table 7.7 presents the results of combining the positions derived in this work and the VLBI positions provided in Wucknitz et al. (2004) paper.

Table 7.7: The additional shifts in the position of image A from radio and corresponding value of the Hubble constant. The values in parenthesis represent the result of leaving the spiral arms unfitted (in the second column, both with and without parenthesis are represented $\Delta\alpha$ and $\Delta\delta$ in mas)

visit	additional shifts of image A from radio	H_0
10	+25, −6 (+26, −6)	76 (77)
11	+29, −5 (+29, −6)	84 (77)
12	+30, −5 (+30, −5)	80 (72)
13	+29, −5 (+28, −6)	78 (80)
14	+25, −5 (+26, −6)	73 (77)
16	+27, −7 (+22, −18)	81 (74)

The mean value for the Hubble constant when we do not fit the arms of the lens galaxy is $76 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and with fitting the arms we obtained $79 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. With combination of optical and radio data, York et al. (2005) determined the Hubble constant to be $79 \pm 7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when they did not mask the arms and $66 \pm 9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when they masked the arms. The determined values for the Hubble constant in this work rule out the lower value of the Hubble constant ($66 \pm 9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$) obtained by York et al. (2005) work when they masked the spiral arms of the lens galaxy. The result of this work on the value of the Hubble constant is consistent with Wucknitz et al. (2004) $H_0 = 78 \pm 6\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

8

Conclusion and Outlook

One of the essential parameters in the standard model of cosmology is the Hubble constant. With a precise observed value of the Hubble constant and with the use of other information, one can determine the physical distance to objects and the age of the universe more accurately and constrain the dark energy equation of state. This parameter can be obtained from different approaches such as supernovae (SNe) and Cepheids by measuring the luminosity distances at different scales in the local and distant universe. Cepheids provide us a luminosity distances through period-luminosity relation in the local universe. The characteristic peak of the SNe allow us to measure cosmic distances but they have to be calibrated with Cepheids. The possibility of redshift dependency evolution of SNe can affect the results on these distance measurements. Combination of the effect of Sunyaev-Zel'dovich (SZ), which explains the distortion of Cosmic Microwave Background (CMB) due to the high energy electrons in the galaxy cluster via inverse Compton scattering, with X-ray flux of the cluster give us an estimate on the angular distances. The assumption of spherical symmetry for the cluster (which can not be true for all the clusters of galaxies) can change the measured distances in this method (Paraficz et al., 2010)).

Theory of gravitational lensing tells us that an object with certain condition for its mass can produce multiple images of a background source reaching the observer with different light travel time. Refsdal showed that with known time delay and redshift of the source and lens one is able to determine the Hubble constant. This method has the advantages of measuring the Hubble constant on cosmic scales with the known physical aspects of the theory of gravitational lensing and well understood problems in the observation area (e.g. mass-sheet degeneracies and flux-ratio anomalies).

8. CONCLUSION AND OUTLOOK

The lens JVAS B0218+357 with double image, an Einstein ring in radio, and an accurate measurement of the time delay between the images gives us a good opportunity to determine the Hubble constant. The challenge in this system is the small separation of the double image of the background blazar in which the removal of the true PSF is essential (which requires careful analysis on the images). Since in radio wavelengths the lens galaxy cannot be observed, previous attempts on this system to determine the Hubble constant were model-dependent. Therefore, this system was observed with the ACS/WFC detector of the HST. The results showed that the separation of the images and relative fluxes of the images in radio and optical are different. The hypothesis is the high extinction of image A due to a giant molecular cloud in the region and some of the optical emission on the image A arises from the host galaxy not the active galactic nucleus (AGN) which dominates the image B and therefore the centre of the image A is not in the agreement with the AGN image (York et al., 2005).

In optical observations, exposures with sub-pixel shifts for large pixel sizes (dithering) are required to obtain information on scales smaller than a pixel. Drizzle is the standard method to reconstruct this information. In the case of gravitational lensing where the mass profile of the lens and the brightness distribution of the source are needed to be fitted, the inversion approach (which is adverse in the cases where preservation of the true PSF is needed) used in Drizzle is not the well suited method. The previous attempt on B0218+357 with the HST data to obtain the Hubble constant suffered from the high-frequency artefacts added by Drizzle and corresponding difficulties in the PSF subtraction (York et al., 2005).

In this work, instead of combining the images by shift-and-add or the Drizzle method, the true brightness distribution of the sky is fitted to the observed exposures. In this direct fitting method, the geometric distortion correction and flagging of cosmic-rays are taken into account. To be able to have a unique solution and avoid unrealistic high fluctuations on small scales, a smoothness constraint is added to this method. This method can handle any arbitrary shifts, rotations and dithering patterns. To have a higher resolution, the pixels of the output result of this method are chosen to have smaller size in comparison to the original exposures taken from Hubble Space Telescopes. This method is tested on both simulated and real data and produces accurate results.

The images of the different visits of B0218+357 are used and combined with the direct fitting method with the model pixel size of 25 mas (half of the original pixel size of the WFC data). We were able to reconstruct faint spiral arms of the lens galaxy which can not be distinguished in the raw images. With the use of a star in the field to model the PSF and a fit for the

lens galaxy and its spiral arms, the positions of the images relative to the centre of the spiral lens galaxy are derived. With an isothermal profile for the lens, we determined the Hubble constant to be $70 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ with fitting and $66 \pm 4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the case of leaving the arms unfitted.

As mentioned earlier, in the optical data of B0218+357, image A suffers from extinction which results in a significant systematic shift on the measured position of the image A. To have more precise results, we combined the positions of image B and galaxy from our results with the positions of image A from radio observations. In this approach, the obtained value for the Hubble constant for the case not fitting the spiral arms changes to $76 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and when we fit the spiral arms we derived the value of $79 \pm 3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ for the Hubble constant. When York et al. (2005) used the combination of both VLBI and optical data and not masking the arms, they determine $H_0 = 79 \pm 7\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $66 \pm 9\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when armed were masked. With our new results we obtained more consistent results and we can rule out the low value for the Hubble constant obtained by York et al. (2005) and their high value of the Hubble constant without masking the arms are fully consistent with our new results.

The direct fitting approach can be used in other applications where physical models have to be fitted to the observed images. This method avoids the intermediate steps of image combination which can reduce possible systematic errors. The outlook for this work is to implement the direct fit of the unlensed brightness distribution of the source. For this implementation, the convolution with PSF and the pixel response function, and the linear effect of gravitational lensing on the brightness distribution will be included in the method. To find the optimum mass distribution of the lens, a minimization on the remaining residuals will be applied.

Appendix A

Hankel transform

The Hankel transform of order zero is an integral transform equivalent to a two-dimensional Fourier transform of a radially symmetric function and also called the Fourier-Bessel transform. The two-dimensional Fourier transform is

$$\begin{aligned} g(u,v) &= \mathcal{F}\{f\}(u,v) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(r)\, e^{-2\pi \mathrm{i}\,(ux+vy)}\,\mathrm{d}x\,\mathrm{d}y. \end{aligned} \tag{A.1}$$

Now we define

$$(x,\ y) = r\,(\cos\theta,\ \sin\theta) \tag{A.2}$$

and

$$(u,\ v) = k\,(\cos\phi,\ \sin\phi). \tag{A.3}$$

By applying Eq. (A.2) and (A.3) in equation (A.1) we obtain

$$\begin{aligned} g(k) &= \int_{0}^{\infty}\int_{0}^{2\pi} f(r)\, e^{-2\pi \mathrm{i}\, r\, k\,(\cos\theta\cos\phi+\sin\theta\sin\phi)}\, r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \int_{0}^{\infty}\int_{0}^{2\pi} f(r) e^{-2\pi \mathrm{i}\, r\, k\,(\cos(\theta-\phi))}\, r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \int_{0}^{\infty}\int_{-\phi}^{2\pi-\phi} f(r)\, e^{-2\pi \mathrm{i}\, r\, k\,(\cos\theta)}\, r\,\mathrm{d}r\,\mathrm{d}\theta \end{aligned} \tag{A.4}$$

$$= \int_0^\infty \int_0^{2\pi} f(r)\, e^{-2\pi \mathrm{i}\, r\, k\, (\cos\theta)}\, r\, \mathrm{d}r\, \mathrm{d}\theta$$
$$= \int_0^\infty f(r) \left[\int_0^{2\pi} e^{-2\pi \mathrm{i}\, r\, k\, (\cos\theta)}\, \mathrm{d}\theta \right] r\, \mathrm{d}r.$$

The term in the bracket in equation (A.4) is the definition of *zeroth order Bessel function of the first kind* $2\pi\, J_0$. Therefore we get

$$g(k) = 2\pi \int_0^\infty f(r)\, J_0(2\pi\, k\, r)\, r\, \mathrm{d}r. \tag{A.5}$$

Hence the Fourier and inverse Fourier transform changes to

$$\begin{aligned} g(k) &= 2\pi \int_0^\infty f(r)\, J_0(2\pi\, k\, r)\, r\, \mathrm{d}r \\ f(r) &= 2\pi \int_0^\infty g(k)\, J_0(2\pi\, k\, r)\, k\, \mathrm{d}k. \end{aligned} \tag{A.6}$$

The function

$$\mathcal{H}\{p\}(x) = \int_0^\infty p(x)\, J_0(a\, x)\, x\, \mathrm{d}x \tag{A.7}$$

is called the *Hankel transform* of the zero order. Thus the equation (A.6) becomes

$$\begin{aligned} g(k) &= 2\pi\, \mathcal{H}\{f(r)\}(2\pi k) \\ f(r) &= 2\pi\, \mathcal{H}\{g(k)\}(2\pi\, r). \end{aligned} \tag{A.8}$$

Appendix B

Galfit

Galfit was developed to extract information on galaxies, stars etc. by using parametric functions to model objects as they appear in the observed image. This tool provides functions to fit more complicated details of shapes that are curved, truncated or have spiral arms.

In this study, the PSF option to fit the PSF of the two images with the use of a star in the field (since the source is blazar, it can be assumed to be point source) and a 'Sersic' profile for the lens galaxy plus additional function to fit the arms are used. Below a brief explanation of each component in the fitting process is provided. This Appendix is based on the paper by Peng et al. (2010).

PSF profile

The common way to fit the PSF image is to convolve a point source (δ-function) with the input PSF (from stars in the field). In galfit the δ-function is approximated with Gaussian function with small width (FWHM < 0.3 pixel size). The position and magnitude of the image are the free parameters of this option.

Sersic profile

This profile is the most commonly used profile in galaxy morphology studies. This profile is useful for modelling flat disks, bulges of galaxies and has the form

$$\Sigma(r) = \Sigma_e \exp\left[-\kappa\left(\left(\frac{r}{r_e}\right)^{1/n} - 1\right)\right] \tag{B.1}$$

where Σ_e is the pixel surface brightness at effective radius r_e. n is the concentration (power-law index) parameter and κ is connected to n such that half of the total flux stands within r_e. In this function, positions of the centre of the galaxy, magnitude, n, r_e, ellipticity and position angle are free parameters.

Gaussian profile

A Gaussian profile is a special case of a Sersic profile with $n = \frac{1}{2}$ and FWHM represents the r_e. This profile has the form of

$$\Sigma(r) = \Sigma_0 \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{B.2}$$

where FWHM = $2.354\,\sigma$. The 6 free parameters in this profile are positions of the centre, the total magnitude, FWHM, ellipticity[1] and position angle.

Logarithmic-Hyperbolic Tangent

To fit the winding rate of spiral arms in late-type galaxies, usually the logarithmic relation with radius is assumed. In order to characterize this function an empirical parameter CDEF is defined such that at bar radius $r_{\rm in}$, the rotation angle reaches 20°. Then with defining

$$A = \frac{2 \times \mathrm{CDEF}}{|\theta_{\rm out}| + \mathrm{CDEF}} - 1.00001 \tag{B.3}$$

and

$$B = \left(2 - \tanh^{-1}(A)\right)\left(\frac{r_{\rm out}}{r_{\rm out} - r_{\rm in}}\right) \tag{B.4}$$

one can define the hyperbolic tangent as

$$\tanh(r_{\rm in}, r_{\rm out}, \theta_{\rm incl}, \theta_{\rm PA}^{\rm sky}; r) \equiv 0.5 \times \left(\tanh\left[B\left(\frac{r}{r_{\rm out}} - 1\right) + 2\right] + 1\right). \tag{B.5}$$

where $r_{\rm out}$ is the radius where 96 percent of the hyperbolic tangent stands inside it. $\theta_{\rm incl}$ is the inclination of the galaxy along the line of sight and $\theta_{\rm PA}^{\rm sky}$ is the position angle of the galaxy in the sky. With these parameters the logarithmic winding rate can be define as

$$\theta(r) = \theta_{\rm out} \tanh(r_{\rm in}, r_{\rm out}, \theta_{\rm incl}, \theta_{\rm PA}^{\rm sky}; r) \times \left[\log\left(\frac{r}{r_{\rm ws}} + 1\right) / \log\left(\frac{r_{\rm out}}{r_{\rm ws}} + 1\right)\right]. \tag{B.6}$$

[1] Equation (B.2) shows the circular Gaussian function but this equation can be modified to elliptical Gaussian as well.

In this formulation, the hyperbolic tangent regulates the bar-length and speed of rotation within $r_{\rm out}$. Outside this radius, logarithmic part is responsible for the rotation. This function has 6 free parameters $r_{\rm in}$, $r_{\rm out}$, $\theta_{\rm incl}$, $\theta_{\rm PA}^{\rm sky}$, $\theta_{\rm out}$ and $r_{\rm ws}$.

List of Figures

List of Tables

References

The Cosmic Lens All-Sky Survey (CLASS). Available from: `http://www.jb.man.ac.uk/research/gravlens/class/class.html`.

SLACS: The Sloan Lens ACS Survey. Available from: `http://www.slacs.org/`.

M. Bartelmann and P. Schneider. **Weak gravitational lensing**. *Phys. Rep.*, **340**:291–472, January 2001.

E. Bertin and S. Arnouts. **SExtractor: Software for source extraction**. *Astron. Astrophys. Supp. Ser.*, **117**(2):393–404, 1996.

A. D. Biggs, I. W. A. Browne, P. Helbig, L. V. E. Koopmans, P. N. Wilkinson, and R. A. Perley. **Time delay for the gravitational lens system B0218+357**. *MNRAS*, **304**:349–358, 1999.

A. D Biggs, O. Wucknitz, R.W. Porcas, I.W.A. Browne, N.J. Jackson, S. Mao, and P.N. Wilkinson. **Global 8.4-GHz VLBI observations of JVAS B0218+357**. *MNRAS*, **338**(3):599–608, 2003.

R. Blandford and R. Narayan. **Fermat's principle, caustics, and the classification of gravitational lens images.** *ApJ*, **310**:568, 1986.

H. Bondi. **Spherically symmetrical models in general relativity.** *MNRAS*, **107**:410, 1947.

H. Bondi and T. Gold. **The Steady-State Theory of the Expanding Universe**. *MNRAS*, **108**:252, 1948.

I.W.A Browne, P.N Wilkinson, N.J.F Jackson, S.T Myers, C.D Fassnacht, L.V.E Koopmans, D.R Marlow, M. Norbury, D. Rusin, C.M Sykes, et al. **The Cosmic Lens All-Sky Survey–II. Gravitational lens candidate selection and follow-up**. *MNRAS*, **341**(1):13–32, 2003.

J.G. Cohen, C.R. Lawrence, and R.D. Blandford. **The Redshift of the Lensed Object in the Einstein Ring B0218+357**. *ApJ*, **583**:67–69, 2003.

P.P. Coles and F. Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure*. Wiley, 2002.

W. de Sitter. **Planetary motion and the motion of the moon according to Einstein"s theory.** *Koninklijke Nederlandse Akademie van Weteschappen Proceedings Series B Physical Sciences*, **19**:367–381, 1917.

A. Einstein. *Näherungsweise Integration der Feldgleichungen der Gravitation*. Wiley Online Library, 1916.

E. Einstein and W. de Sitter. **On the Relation between the Expansion and the Mean Density of the Universe**. *Proc Natl Acad Sci USA.*, **18**:213, 1932.

E.E. Falco, M.V. Gorenstein, and I.I. Shapiro. *ApJ*, **289**:L1, 1985.

C.D. Fassnacht, R.R. Gal, L.M. Lubin, J.P. McKean, G.K. Squires, and A.C.S Readhead. **Galaxy Groups Associated with B1608+ 656**. *ApJ*, **642**:30–38, 2006.

C.D. Fassnacht, D.S. Womble, G. Neugebauer, I.W.A. Browne, A.C.S. Readhead, K. Matthews, and T.J. Pearson. **B1608+656: A Gravitationally Lensed Poststarburst Radio Galaxy**. *ApJ*, **460**:L103, 1996.

A. Friedmann. **Über die Krümmung des Raumes**. *Zeitschrift für Physik*, **10**:377, 1922.

A. Fruchter, M. Sosey, and et al. *The MultiDrizzle Handbook, version 3.0*. Baltimore, STScI, 2009.

A. S. Fruchter and R. N. Hook. **Drizzle: A Method for the Linear Reconstruction of Undersampled Images**. *Publ. Astron. Soc. Pac.*, **114**(792):pp. 144–152, 2002.

Andrew S. Fruchter. **A New Method for Band-limited Imaging with Undersampled Detectors**. *Publ. Astron. Soc. Pac.*, **123**(902):pp. 497–502, 2011.

M. Galassi and et al. *GNU Scientific Library Reference Manual (3rd Ed.)*. 2009.

M. Gerin, T.G. Phillips, D.J. Benford, K.H. Young, K.M. Menten, and B. Frye. *ApJ*, **488**:L31, 1997.

S. Gonzaga et al. *ACS Data Handbook, Version 6.0*. Baltimore, STScI, 2011.

M.V. Gorenstein, N.L. Cohen, I.I. Shapiro, A.E.E. Rogers, R.J. Bonometti, E.E. Falco, N. Bartel, and J.M. Marcaide. **VLBI observations of the gravitational lens system 0957+ 561-Structure and relative magnification of the A and B images**. *ApJ*, **334**:42–58, 1988.

F. Grundahl and J. Hjorth. **The optical appearance of the gravitational lens system B0218+35.7**. *MNRAS*, **275**:L67–L71, August 1995.

C. Henkel, N. Jethava, A. Kraus, K.M. Menten, C.L. Carilli, M. Grasshoff, D. Lubowich, and M.J. Reid. **The kinetic temperature of a molecular cloud at redshift 0.7: ammonia in the gravitational lens B0218+ 357**. *A&A*, **440**(3):893–899, 2005.

F. Hoyle. **A New Model for the Expanding Universe**. *MNRAS*, **108**:372, 1948.

E. Hubble. **A relation between distance and radial velocity among extra-galactic nebulae**. *Proceedings of the National Academy of Sciences*, **15**(3):168–173, 1929.

N. Jackson, E. Xanthopoulos, and I. W. A. Browne. **NICMOS images of JVAS/CLASS gravitational lens systems**. *MNRAS*, **311**:389–396, January 2000.

E. Jullo, J.-P. Kneib, M. Limousin, A. Eliasdottir, P. J. Marshall, and T. Verdugo. **A Bayesian approach to strong lensing modelling of galaxy clusters**. *New J. Phys.*, **9**:pp. 447, 2007.

C. R. Keeton. **A Catalog of Mass Models for Gravitational Lensing**. *ArXiv Astrophysics e-prints*, February 2001.

C.S. Kochanek, E.E. Falco, C. Impey, J. Lehar, B. McLeod, and H.-W. Rix. **CASTLES Survey**. Available from: `http://www.cfa.harvard.edu/castles/`.

C.S. Kochanek and M. White. *ApJ*, **543**:514, 2000.

J. Krist. **Tiny Tim : an HST PSF Simulator**. In R. J. Hanisch, R. J. V. Brissenden, and J. Barnes, editors, *Astronomical Data Analysis Software and Systems II*, **52** of *Astronomical Society of the Pacific Conference Series*, page 536, 1993.

C. R. Lawrence. **Observations Of Lens Systems With Keck I**. In C. S. Kochanek and J. N. Hewitt, editors, *Astrophysical Applications of Gravitational Lensing*, **173** of *IAU Symposium*, page 299, 1996.

J. LEHAR ET AL. *ApJ*, **536**:584, 2000.

G. LEMATRE. **Un Univers homogne de masse constante et de rayon croissant rendant compte de la vitesse radiale des nbuleuses extra-galactiques**. *Annales de la Societe Scientifique de Bruxelles*, **47**:49, 1927.

G. LEMATRE. **L'univers en expansion.** *Annales de la Societe Scietifique de Bruxelles*, **53**, 1933.

O. LODGE. **Gravitation and light.** *Nature*, **104**:334, 1919.

D.W. MURPHY, I.W.A. BROWNE, AND R.A. PERLEY. **VLA observations of a complete sample of core-dominated radio sources**. *MNRAS*, **264**:298, 1993.

S.T. MYERS, C.D. FASSNACHT, S.G. DJORGOVSKI, R.D. BLANDFORD, K. MATTHEWS, G. NEUGEBAUER, T.J. PEARSON, A.C.S. READHEAD, J.D. SMITH, D.J. THOMPSON, ET AL., 1995.

S.T MYERS, N.J JACKSON, I.W.A BROWNE, A.G DE BRUYN, T.J PEARSON, A.C.S READHEAD, P.N WILKINSON, A.D BIGGS, R.D BLANDFORD, C.D FASSNACHT, ET AL. **The Cosmic Lens All-Sky Survey–I. Source selection and observations**. *MNRAS*, **341**(1):1–12, 2003.

D. PARAFICZ AND J. HJORTH. **The Hubble constant inferred from 18 time-delay lenses**. *ApJ*, **712**(2):1378, 2010.

A. R. PATNAIK, I. W. A. BROWNE, P. N. WILKINSON, AND J. M. WROBEL. **Interferometer phase calibration sources. I - The region 35-75 deg**. *MNRAS*, **254**:655–676, 1992.

A.R. PATNAIK, I.W.A. BROWNE, L.J. KING, T.W.B. MUXLOW, D. WALSH, AND P.N. WILKINSON. **B0218+ 35.7-A gravitationally lensed system with the smallest separation**. *MNRAS*, **261**:435–444, 1993.

C. Y. PENG, L. C. HO, C. D. IMPEY, AND H.-W. RIX. **Detailed Decomposition of Galaxy Images. II. Beyond Axisymmetric Models**. *AJ*, **139**:2097–2129, 2010.

S. PERLMUTTER AND A. RIESS. **Cosmological parameters from supernovae: Two groups results agree**. In *AIP Conference Proceedings*, **478**, page 129, 1999.

ROBERT PIESSENS. *The Hankel Transform*, chapter The Hankel Transform. 2000. in (52).

R.W. Porcas, R.S. Booth, I.W.A Browne, D. Walsh, and P.N. Wilkinson. **VLBI structures of the images of the double QSO 0957+ 561**. *Nature*, **289**:758–762, 1981.

Alexander D. Poularikas. *The Transforms and Applications Handbook, Second Edition*. CRC Press, 2000. 110

William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical recipes in C (2nd ed.): the art of scientific computing*. Cambridge University Press, New York, NY, USA, 1992.

S. Refsdal. **The gravitational lens effect**. *MNRAS*, **128**:295, 1964a.

S. Refsdal. **On the possibility of determining Hubble's parameter and the masses of galaxies from the gravitational lens effect**. *MNRAS*, **128**:307, 1964b.

D.H. Roberts, P.E. Greenfield, and B.F. Burke. **The double quasar 0957+ 561- A radio study at 6-centimeters wavelength**. *Science*, **205**(4409):894–896, 1979.

H. P. Robertson. **Kinematics and World-Structure**. *ApJ*, **82**:284, 1935.

P. Schneider. **A new formulation of gravitational lens theory, time-delay, and Fermat's principle**. *A&A*, **143**:413–420, February 1985.

P. Schneider, C. Kochanek, and J. Wambsganss. *Gravitational Lensing: Strong, Weak and Micro: Saas-Fee Advanced Course 33*. Springer, 2006.

P. Schneider and C. Seitz. **Steps towards nonlinear cluster inversion through gravitational distortions. 1: Basic considerations and circular clusters**. *A&A*, **294**:411–431, February 1995.

K. Schwarzschild. **On the Gravitational Field of a Point-Mass, According to Einstein's Theory**. *Sitzungsber. Preuss. Akad. Wiss., Phys. Math. Kl*, **189**, 1916.

M. Stickel, H. Kühr, and J.W. Fried. **Spectroscopy of 1 Jy and S5 radio source identifications. II**. *Astron. Astrophys. Supp. Ser.*, **97**:483–500, 1993.

S. Suyu, P.J. Marshall, M.W. Auger, S. Hilbert, R.D. Blandford, L.V.E. Koopmans, C.D. Fassnacht, and T. Treu. **Dissecting the Gravitational lens B1608+ 656. II. Precision measurements of the hubble constant, spatial curvature, and the dark energy equation of state**. *ApJ*, **711**(1):201, 2010.

REFERENCES

S. H. Suyu, P.J. Marshall, M.P. Hobson, and R.D. Blandford. **A bayesian analysis of regularized source inversions in gravitational lensing**. *MNRAS*, **371**:pp. 983–998, 2006.

Richard C. Tolman. **Effect of inhomogeneity on cosmological models.** *Proceedings of the National Academy of Sciences of the United States of America*, **20.3**:169, 1934.

L. Ubeda et al. **ACS Instrument Handbook, Version 11.0**. *Baltimore: STScI*, 2011.

A. G. Walker. *Proceedings of the London Mathematical Society*, **42**:90, 1936.

D. Walsh, R.F. Carswell, and R.J. Weymann. **0957 + 561 A, B: twin quasistellar objects or gravitational lens?** *Nature*, **279**:381, 1979.

S. J. Warren and S. Dye. **Semilinear Gravitational Lens Inversion**. *ApJ*, **590**:pp. 673682, 2003.

Ray J. Weymann, D. Latham, J. R. P. Angel, R. F. Green, J. W. Liebert, and D. A. Turnshek. **The triple QSO PG1115+08: another probable gravitational lens.** *Nature*, **285**:641, 1980.

T. Wiklind and F. Combes. **CO, HCO+ and HCN absorption in the gravitational lens candidate B0218+ 357 at Z= 0.685.** *A&A*, **299**:382, 1995.

T. Wiklind and F. Combes. **Probing the Interstellar Medium at High Redshift Using Molecular Absorption Lines**. In *Young Galaxies and QSO Absorption-Line Systems*, **114** of *Astronomical Society of the Pacific Conference Series*, page 163, 1997.

H.J. Witt, S. Mao, and C.R. Keeton. **Analytic Time Delays and H0 Estimates for Gravitational Lenses**. *ApJ*, **544**:98, 2000.

O. Wucknitz. **LensClean revisited**. *MNRAS*, **349**(1):1–13, 2004.

O. Wucknitz, A. D. Biggs, and I. W. A. Browne. **Models for the lens and source of B0218+357: a LENSCLEAN approach to determine H_0**. *MNRAS*, **349**:14–30, 2004.

T. York, N. Jackson, I. W. A. Browne, O. Wucknitz, and J. E. Skelton. **The Hubble constant from the gravitational lens CLASS B0218+357 using the Advanced Camera for Surveys**. *MNRAS*, **357**(1):124–134, 2005.

F. Zwicky. *Helv. Phys. Acta*, **6**:110, 1933.

F. ZWICKY. **Nebulae as gravitational lenses.** *Phys. Rev.*, **514**:290, 1937.

F. ZWICKY. **On the Masses of Nebulae and of Clusters of Nebulae.** *ApJ*, **86**:217, 1937.