THE GRAVITATIONAL LENS EFFECT

of

GALAXIES and BLACK HOLES

by

Igor Bray, B.Sc. (Hons.)

A thesis submitted in accordance with the requirements of the Degree of Doctor of Philosophy.

Department of Mathematical Physics
The University of Adelaide
South Australia

January 1986

to my wife Ann

CONTENTS

STATEMENT

This thesis contains no material which has been accepted for the award of any degree, and to the best of my knowledge and belief, contains no material previously published or written by another person except where due reference is made in the text.

The author consents to the thesis being made available for photocopying and loan if applicable if accepted for the award of the degree.

Igor Bray

Acknowledgements

The author wishes to thank Dr. P. Szekeres for his supervision as well as for his encouragement to publish my work. I am also grateful to the Computing Centre of the University of Adelaide for the use of their computing facilities. In particular, I would like to thank Andy Cheel for his continuous assistance in that area. I am indebted to Dr. S. Elhay of the Computing Department of the University of Adelaide for many discussions regarding numerical techniques, and in particular, for his making available to me a most efficient method of numerical integration for integrals which have a singularity of the type $x^{-1/2}$. The author acknowledges the support of the Commonwealth Government of Australia via a Postgraduate Research Award.

ABSTRACT

This thesis examines the multi-imaging aspect of the gravitational lens effect of both galaxies and black holes. Galaxies are modeled by spheroids, the density of which is assumed to have spheroidal symmetry i.e., the density is constant on the surface as well as all the other spheroidal shells that make up the spheroid. If such a spheroid is projected onto a plane then the resultant solid ellipse's density has elliptical symmetry.

It is shown that the equation $z = z_0 - 4DGc^{-2}I^*(z_0)$ which relates the position of an image z_0 with the position of the source z, must satisfy certain symmetry properties. In particular,

$$z = z_0 - 4DGc^{-2}I^*(z_0) \quad \Rightarrow \quad -z = -z_0 - 4DGc^{-2}I^*(-z_0).$$

The solution for $I(z_0)$ presented by Bourassa and Kantowski (1975) violates this property, as their $I(z_0) = I(-z_0)$. Thus a corrected $I(z_0)$ is computed. This contains a proof of the fact that if the projected density has elliptical symmetry, then there is no contribution to the bending angle due to mass lying outside the ellipse formed by the impact parameter. The new $I(z_0)$ is used to present diagrams which show how many images of a single source will result as a function of the position of the source with respect to the galaxy.

Similar diagrams are presented for the case where it is the Kerr black hole that behaves as a gravitational lens. Unlike other massive objects, black holes have r = constant orbits for null geodesics. This implies that a single source of light will have infinitely many images at the observer. These consist of the direct rays and rays which orbit the black hole once, twice, etc. We concern ourselves with the direct rays as these are the ones which are least deviated, hence least deformed and most intense, and therefore, are the ones which are most likely to be observed.

Approximate solutions to the equations of motion for direct null geodesics in the Kerr metric are derived. These are correct up to and including second order terms in m/r_{min} and a/r_{min}, where m is the Kerr mass, a is the Kerr spin and r_{min} the distance of closest approach, and are sufficient for most astronomical purposes. Second order terms are included as spin comes in only at this order and our primary interest is the effect of spin on image multiplicity. It is found that the derived expressions can be combined with numerical integration to give excellent results for rays which have relatively large deviations.

PART I

Spheroidal Gravitational Lenses

§1 Introduction

§1.1 Spherical Gravitational Lenses

The "gravitational lens effect" is the effect of massive objects upon the paths of rays of light. The word lens is used as this effect is comparable to that of a glass optical lens, something which is familiar from classical optics.

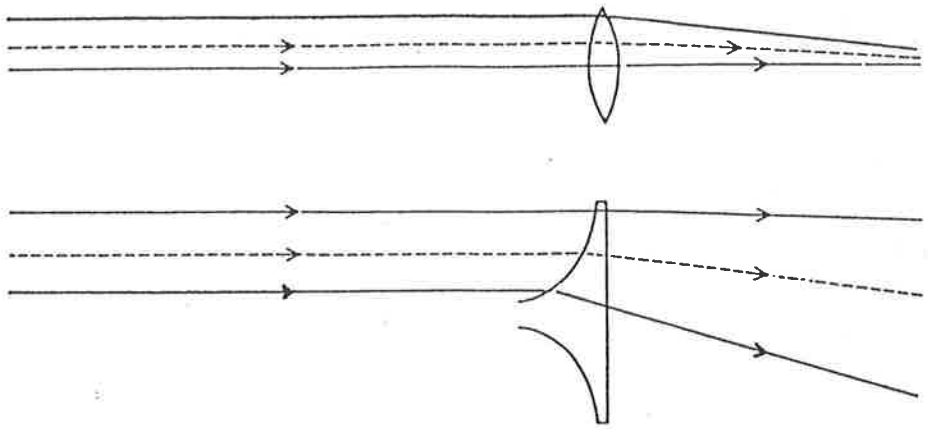

Figure 1.1 Optical model of a point-mass lens. At the top, is an ordinary positive lens where the closer the light passes to the centre, the less it is bent. In a point-mass gravitational lens the light does the opposite: the closer it passes to the centre the more it is bent. At the bottom, is an optical analogue of a gravitational lens.

There are many aspects of the gravitational lens effect that have been of interest to scientists since Einstein proposed his theory of General Relativity back in 1916. The first aspect of this effect studied was the deviation of a ray of light due to a spherically symmetric massive body. Questions regarding size of deviation, shape and intensification of the image were carefully studied. Other topics of interest such as multi-imaging and time delays between images were also investigated. It is the multi-imaging aspect of the gravitational lens effect that is of most interest to us.

Gravitational lens effects of spherical deflectors have been thoroughly investigated by Einstein (1936), Zwicky (1937), Tikhov (1937), Liebes (1964) and Refsdal (1964a). Many applications have been made to objects seen around galaxies by

Zwicky (1957), Metzner (1963), Klimov (1963), Refsdal (1964b, 1966), Noonan (1971) and Barnothy and Barnothy (1972).

As early as 1937 Zwicky suggested that some galaxies at different distances might appear to overlap in the sky. He concluded: "The gravitational fields of a number of foreground nebulae may therefore be expected to deflect the light coming to us from certain background nebulae. The observations of such gravitational lens effects promises to furnish us with the simplest and most accurate determination of nebular masses."

Einstein recognized the possibility that the image of a distant star could be split by the gravitational field of an intervening one, but he concluded that "there is no great chance of observing this phenomenon." However he did not know of the existance of quasars, which were discovered 25 years later.

This prediction held for a long time and was only broken on March 29^{th} of 1979, when Dennis Walsh of the Jodrell Bank radio-astronomy observatory, Robert F. Carswell of the University of Cambridge and Ray J. Weymann of the University of Arizona observed what turned out to be the first recognized observation of multi-imaging due to a gravitational lens. In fact the quasar twins, 0957+561 A and B, as they have been officially designated, had been observed and photographed back in the 1950's with the 1.2 metre Schmidt telescope on Palomar Mountain. From the photograph in fig.1.2 it is clear that the blurred image is that due to two objects, however, no investigation of these objects was done for more than twenty years.

On the above-mentioned date, Walsh, Carswell and Weymann obtained spectra indicating that each of the two images, which are 5.7 seconds of arc apart, are the image of a quasar. For each of the twins they measured a red shift of 1.405 and found extraordinary similarity between their spectra.

There are 1500 known quasars, which are uniformly distributed in the sky. On average, one quasar is found every 30 square degrees. Therefore, it was a great surprise to find two quasars so close together.

The spectra of quasars are characterised by broad lines emitted by atoms that have been stripped of at least one electron by the intense radiation field to which they are subjected. Many such emission lines are possible, and each quasar has

its unique collection of them. Not only do the twins have the same set of emission lines, but the strength of any particular line in one spectrum is the same as that of the corresponding line in the other, see fig.1.3. Moreover, the discoverers found all the lines had the same red shift, within experimental error.

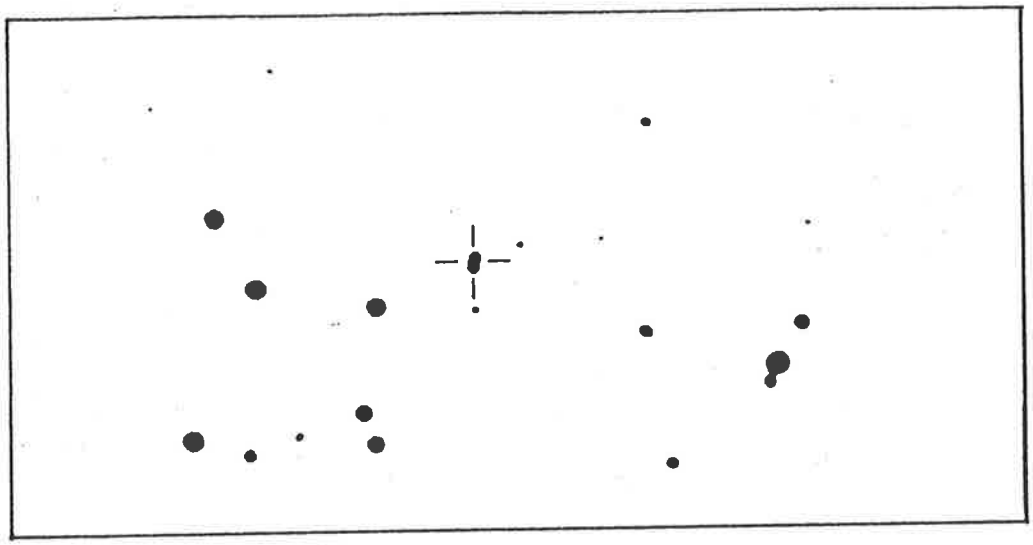

Figure 1.2 First photograph of the quasars. The fused image of the twins appears in this negative print of a photograph made in the early 1950's with the 1.2-metre Schmidt telescope on Palomar Mountain. The image is fused because of the motion of the earth's atmosphere. If the telescope had been above the atmosphere, it could have resolved objects 60 times closer together than the twins. The photograph indicates the difficulty of examining the space between the twins with optical instruments.

Like many other quasars, the twins also have in their spectra sharp absorption lines that probably are caused by cool gas lying between the quasar and our galaxy. The red shift of these absorption lines usually differs from that of the broad emission lines, and in some quasars the absorption lines show more than one red shift, indicating that the radiation has passed through a corresponding number of clouds.

Walsh, Carswell and Weymann found that the absorption lines in the spectrum of quasar A have the same red shift, indicating that the radiation passes through only one intervening cloud. The red shift of the absorption lines differs from that of the emission lines by only 0.6 of a percent, suggesting that the in-

tervening cloud is close to the quasar itself. Furthermore, the absorption lines in the spectrum of quasar B have the same red shift as the absorption lines in the spectrum of quasar A, to within experimental uncertainty; see fig.1.4.

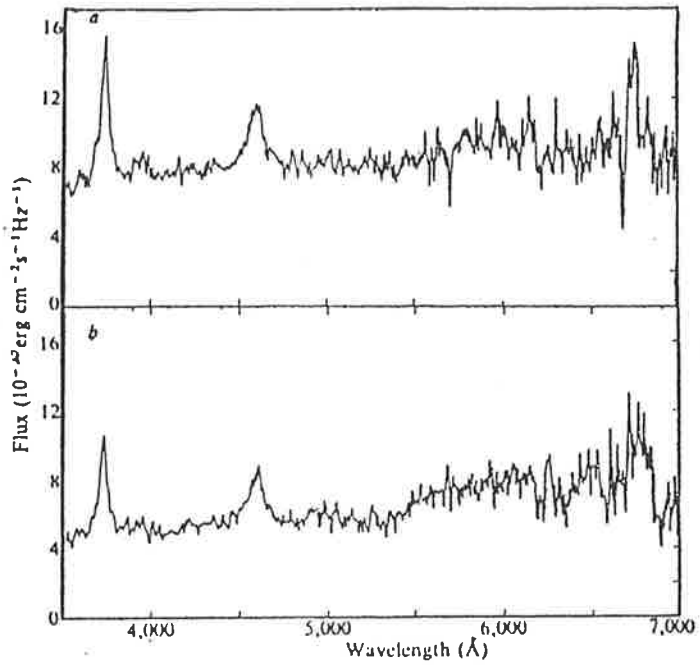

Figure 1.3 IIDS scans of 0957+561 A and B The data is smoothed over 10Å and the spectral resolution is 16Å. Nature 279, 31 May 1979

Thus there was seemingly overwhelming evidence in favour of the gravitational lens hypothesis. Nevertheless, other possible explanations for the remarkable similarity of the spectra had to be considered, as the subject was still regarded to be of a speculative nature. In fact, Frederic H. Chaffee, Jr., who later joined the Walsh group, said "...the subject of gravitational lenses has had the odor of witchcraft for the notoriously conservative astronomical community."

It is well known that there are binary stars that travel in orbit around each other. Could the twins simply be the first example of a binary quasar? Clearly they are physically related, and if they are a binary quasar, the similarities in their spectra could arise naturally from the fact that the twins originated and evolved in the same cosmic environment. This explanation seems to be rather ad hoc, and yet it divided the astronomical community Putting aside the ad hoc aspect of the binary explanation, the evidence against it rests chiefly on the interpretation of the

absorption lines, particularly on the similarity of the red shifts of the absorption lines.

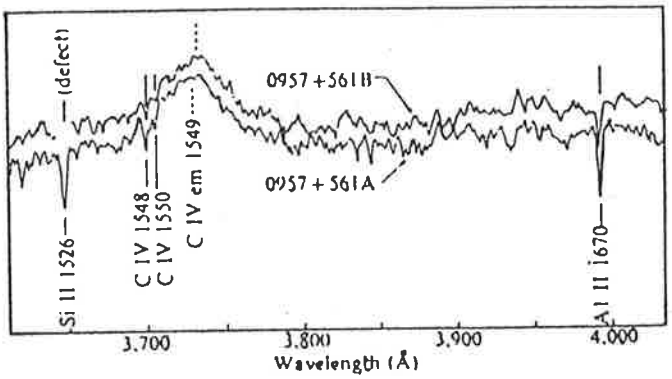

Figure 1.4 Microdensitometer tracings of portions of the 0957+561 A and B. The solid lines mark the position of absorption features in the two quasars and the dashed lines mark the adopted centres of the CIV emission line.

Suppose the cloud around the twins resulted from the violent ejection of matter from one of them, say quasar A, at some time in the past. Figure 1.5 shows the current geometrical relations between the quasars and the cloud, which is shown as an expanding spherical shell with an expansion velocity v. The component of v toward us, where the line of sight to quasar A intersects it, is equal to the difference between the velocity represented by the red shift of the emission lines of quasar A and the velocity represented by the red shift of the absorption lines of quasar A. Red shifts yield velocities along the line of sight only, thus we expect the cloud velocity measured towards B to be somewhat less than v, see fig.1.5

Since we know the angular separation as well as the red shift, the separation of the quasars can be calculated to be 200,000 light-years, nearly three times the diameter of a typical galaxy. Now, experimental error in the difference between the absorption line red shifts was shown to be no more than 150 km/s, quite good, considering that the cloud is receding at 210,000 km/s. Thus the minimum radius of the cloud, necessary to cause the maximum possible difference in the red shifts, can be calculated to be 575,000 light-years.

This distance, combined with the measured strength of the absorption lines, makes it possible to calculate that the minimum mass of the shell is 10^{12} times the

mass of the sun. The energy required to eject such a massive shell is equivalent to 10^{28} suns expelled in one violent explosion. As this is physically unlikely, immediate doubt is cast upon the binary quasar hypothesis.

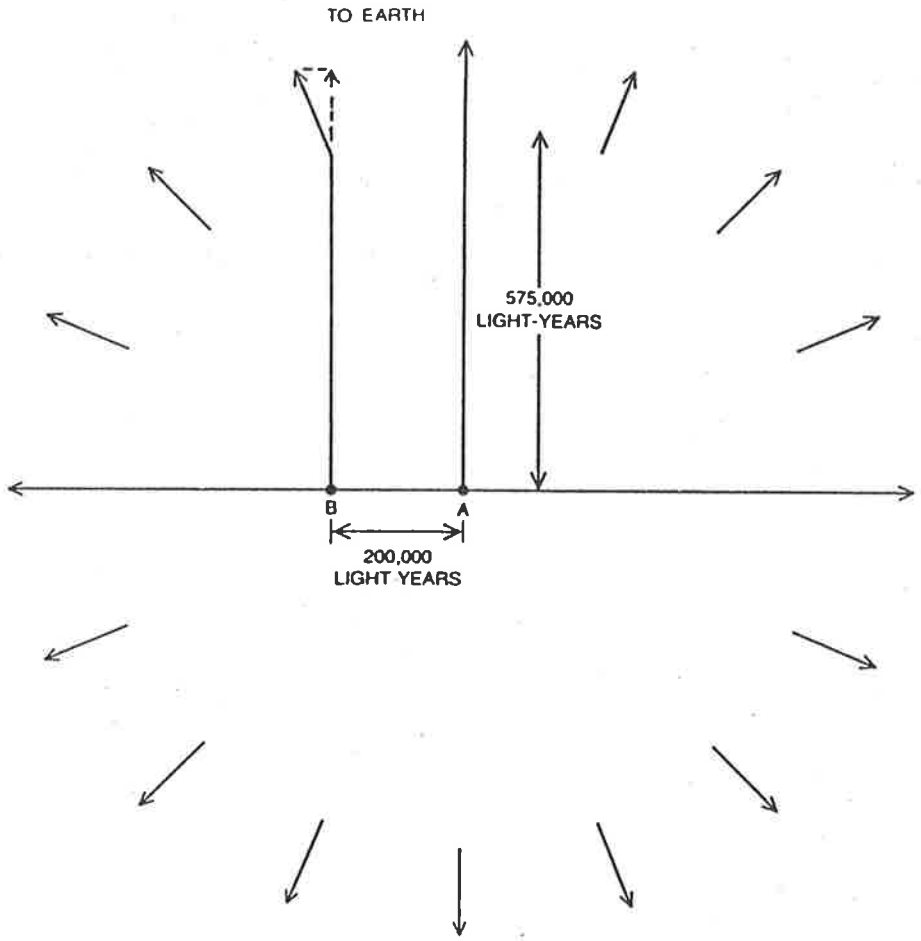

Figure 1.5 Binary quasar hypothesis. This claims that the twins are not dual images of a single quasar, but are actually two quasars. The arrows indicate the expanding cloud.

This hypothesis was made further untenable by subsequent observations. Three weeks after the initial discovery, Walsh, Carswell and Weynmann joined Marc Davis of the Harvard College Observatory, Nathaniel P. Carleton and Chaffee at the then new multiple-mirror telescope (MMT) of the Smithsonian Astrophysical Observatory and the University of Arizona. The measured spectra were of such high quality that they concluded that the absorption line velocities did not differ by more than 15 km/s.

The only way to hold on to the binary quasar hypothesis now, was to suggest that instead of the gas cloud being an enormous sphere, that it was small and

ejected in the direction of our galaxy only. As well, any rotational motion of the quasars would have to be perpendicular to the line of sight. Thus as the binary quasar hypothesis required too many coincidences to hold, the gravitational lens hypothesis was favoured at that stage.

But the real "proof" of the gravitational lens hypothesis would be to observe the lens itself i.e., the intervening massive object. This was no simple task, as can be seen from the closeness of the two images in the photograph in fig.1.2.

David H. Roberts, Perry E. Greenfield and Bernard F. Burke of the Massachusetts Institute of Technology obtained data on the twins using the Very Large Array radio telescope. This data made it possible to achieve a resolution of 0.8 arc seconds, 7.5 times better than the resolution of the optical photograph.

The radio picture revealed two radio images of equal brightness, to the northeast of the northern quasar image, and no radio emission from between the twins, see fig.1.6. From the position of the two radio blobs, the astronomers inferred that both were due to the northern image. However, the gravitational lens hypothesis expected to have similar radio emission associated with the southern image of the quasar. This, together with no radio emission from the area in between the quasar images led the M.I.T. workers to conclude that the images were due to a binary quasar system. The astronomical community was divided into two factions: one consisted of the optical astronomers who supported the gravitational lens hypothesis, since any conventional explanation leads to unphysical energies, and the other consisting of the radio astronomers who favoured the conventional binary quasar hypothesis, and avoided the large energies by having the quasars eject only a small cloudlet directly toward our galaxy, and have the plane of rotation perpendicular to the line of sight. The radio astronomers cited the absence of a second image of the northeast radio blobs as sufficient evidence to rule out the gravitational lens hypothesis. Clearly there was a need for further observations.

On November 15 of 1979 the Hale group consisting of Peter Young, James E. Gunn, Jerome Kristian, J. Beverley Oke and James A. Westphal took a two hour CCD (charge-coupled device) exposure of the twins with a five metre telescope. Subsequent analysis of the exposure showed that the southern image was elongated to the north by one arc second and that the northern image had the circular shape

characteristic of a single object.

On November 28 of the same year Alan N. Stockton of the University of Hawaii at Manoa, took a series of one minute photographic exposures with a 2.2 metre telescope atop the 4,200 metre peak of Mauna Kea on the island of Hawaii. His photographs turned up a small patch of fuzz to the north of the south image. The Hale astronomers were quick to recognize that their elongated image and Stockton's fuzz were the same object. They realized that this fuzz was in fact the galaxy they were looking for. The image of the galaxy is almost coincident with the southern quasar image, and so the spectrum of the southern image is undoubtedly contaminated by the light of the much fainter galaxy. A close study of the results revealed a spectrum of the galaxy superimposed on the spectrum of the southern image. The Hale observers concluded that the galaxy has a red shift of 0.4, indicating that it lies about halfway between the quasar and our galaxy.

The news was very exciting, but nagging questions remained. The proponents of the gravitational lens hypothesis had thought that for the lens to create dual images of almost equal brightness it would have to be almost exactly on the line of sight to the quasar. Moreover, it should appear to be halfway between the twin images. If the galaxy really is a gravitational lens, why is it so close to the southern image? And where are the dual images of the northeast radio blobs?

The answer to these questions was simply to point out that the assumption that the gravitational lens is a spherically symmetric point mass is incorrect. And instead one has to use the theory of extended gravitational lenses which have quite different properties to the spherical ones and are more likely to model the realistic gravitational lenses.

Using extended mass models of the gravitational lens the Hale group concluded that the reason for no dual imaging of the radio blobs is that their origin is too far out of line of the deflecting galaxy. They also concluded that the southern image consists of two images which are too close to be resolved and hence appears just as bright as the northern image, see fig.1.6.

It is these observations that are of interest to the author. Namely, since there are regions where multiplicity of images occur, it is interesting to find the shape and the extent of these regions. Also, since the geometry of the deflector effects the

number of images, we wish to know just how it does so, and how many images one may obtain from a realistic model of a galaxy which behaves like a gravitational lens.

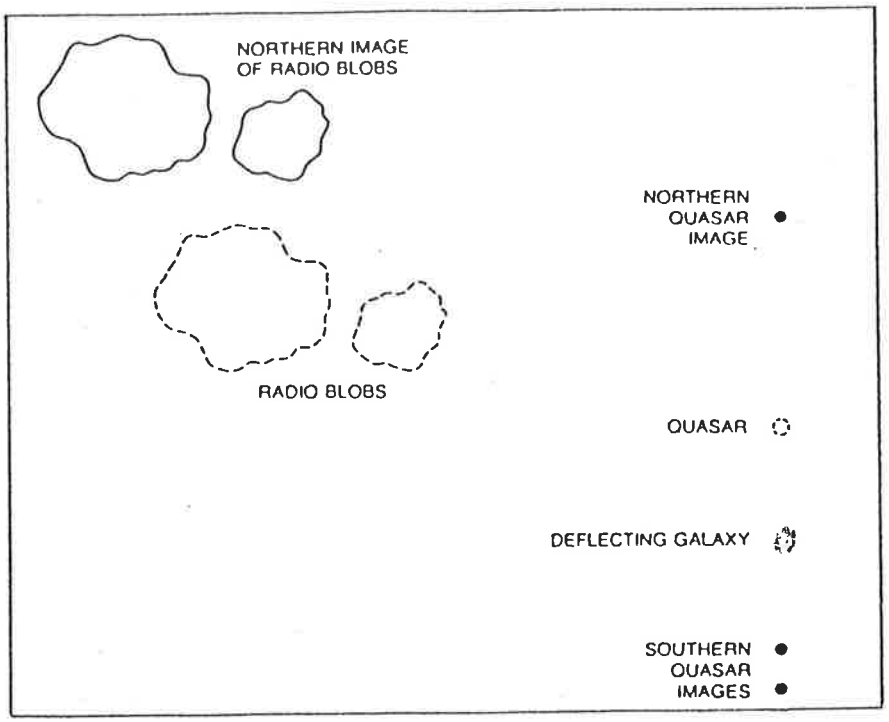

Figure 1.6 The Final Picture. The solid lines show the apparent position of the radio wave sources. The broken lines show the actual position of the radio sources and and the position of the quasar itself. The galaxy is sufficiently close to the line of sight of the quasar to create three images of it, but not sufficiently close to the source of the radio waves to create more than one image of them. The two southern images are too close to be resolved, but their brightness adds up to give a similar brightness to that of the northern image.

Since the discovery of the first gravitational lens, another five have been observed. The last of these differed from the others in that the intervening galaxy is very much closer to the Earth. Infact, unlike the other examples where the existance of adjacent quasar images led to the search for the intervening galaxy, it was the galaxy which was first observed. It was only a routine examination of the spectrum of the galaxy that revealed a single quasar image directly in line with

the centre of the galaxy, see fig.1.7.

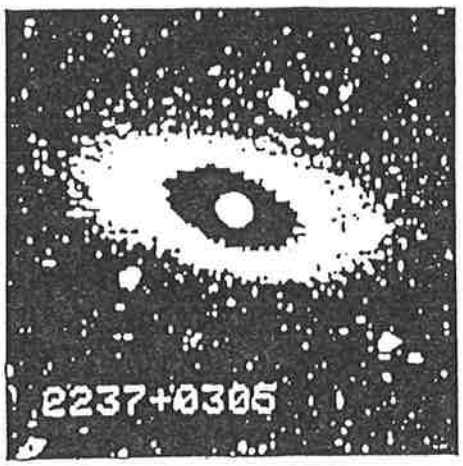

Figure 1.7 Nearest Gravitational Lens. The image has been processed so that the outer faint parts of the galaxy form the large white oval, and the intermediate region has been deliberately inverted to show the position of the quasar with respect to the galaxy.

The galaxy has a red shift of 0.039, indicating that it is about 300 million light years away, while the quasar has a red shift of 1.7, which implies that it is 12×10^9 light years away.

The surprising aspect of this gravitational lens is that only one image is observed, which led to the suggestion that this was due to the unique position of the quasar with respect to the lens. However, in an earlier publication, resulting from the first part of this thesis, we showed multiplicity of images is greatest when the source and the lens are aligned. Thus, in this case there must be more than one image, probably three or five. The reason that these are not observed is that they are too close together. This is due to the fact that the ratio of distances: observer to lens, and observer to source is 1:40, whereas for maximal effect the ratio should be 1:1.

Having obtained observational data of gravitational lens effects, astronomers began to make detailed mathematical models to explain the data. The discovery of the first gravitational lens was very exciting to the astronomical community, as for years they had known that such a discovery would lead to great insights into the workings of the universe.

In 1964, Sjur Refsdal showed that by observing time delays between images

of a single source, one can attempt to calculate the Hubble constant H_0. This calculation was valid for sources with small red shifts and thus was not valid for quasars. Thus, in 1965 he included the higher order terms necessary for larger redshifts and found that these depended on the choice of the cosmological theory that one adopted. This led him to consider the possibility of testing cosmological theories from the gravitational lens effect. But there are more fundamental problems that this effect helps to solve.

Quasars in general have been a puzzle since their discovery 25 years ago. Their redshifts are in excess of any other known object. As well, despite being much smaller than a galaxy they generate 100 times more radiation than the brightest galaxy known. This created doubt in the astronormers' minds as to whether the redshifts of the quasars were indeed cosmological. Perhaps, it was thought, quasars are much closer than their redshifts indicate. The observation of the quasar twins and the intervening galaxy implies that the redshifts are indeed cosmological. Since we know the distance of the galaxy, the separation between the images, and that the galaxy is a typical galaxy i.e., we can make an estimate of its mass, we can independantly estimate the distance of the quasar itself. This proves to be consistent with the cosmological interpretation of redshifts for quasars. Having assumed that the redshifts are cosmological, we can make more accurate estimates of the intervening mass.

Thus, mathematical modelling of gravitational lenses became of prime interest once more.

§1.2 Spheroidal Gravitational Lenses

The foundations behind gravitational lens models for galaxies have been laid down by the complex formalism of Bourassa, Kantowski and Norton (1973) and Bourassa and Kantowski (1975). In this formalism galaxies are modelled by spheroids, the density of which is assumed to have spheroidal symmetry. For greater generality it is assumed that the gravitational lens may be transparent.

Using the notation of Bourassa and Kantowski (1975) suppose that the source is a distance D_s from the observer and that the deflecting galaxy is somewhere in between the source and the observer, a distance of D_d away from the observer and a distance of D_{ds} away from the source, see fig.1.8.

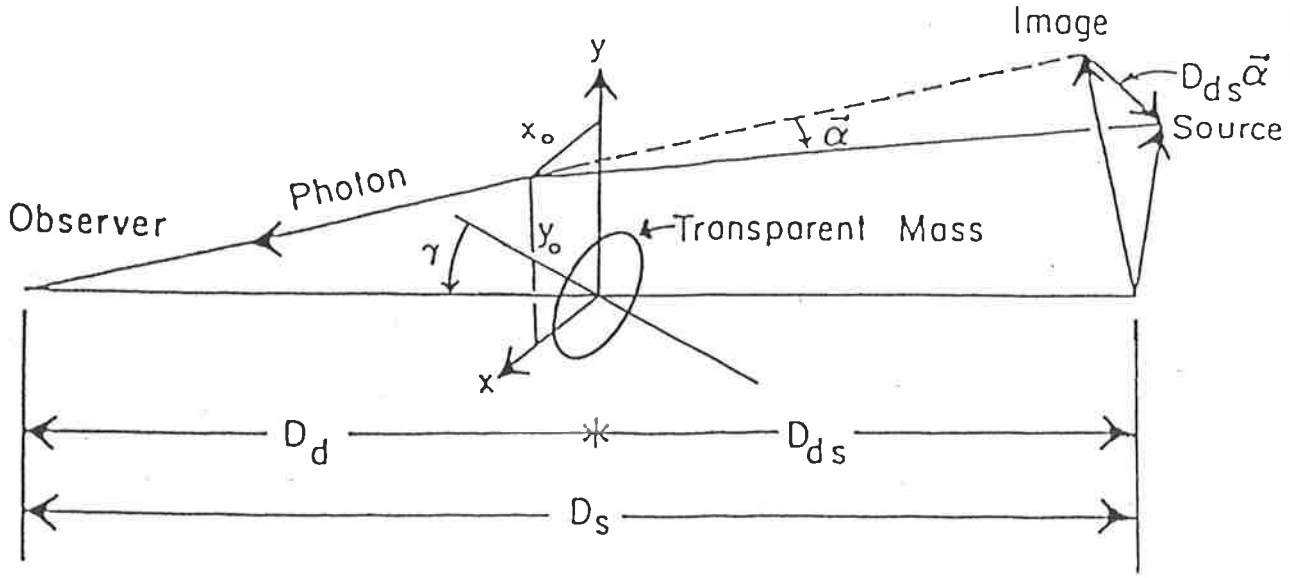

Figure 1.8 Parameters of spheroidal gravitational lenses. The mass is assumed to be an oblate spheroid, whose axis forms an angle γ with the line joining the observer and the origin of the x y plane. The plane is thought of as a complex plane with $z = x + iy$. The bending angle α is assumed to be small. The ray intersects the plane at $z_0 = x_0 + iy_0$.

Consider the plane through the centre of the spheroid and perpendicular to the line connecting the centre of the spheroid and the observer as a complex plane. The point where the ray of light cuts this plane we call the impact parameter and write it as $x_0 + iy_0 = z_0$. The position of the source is projected onto this plane as is the spheroid, see fig.1.9.

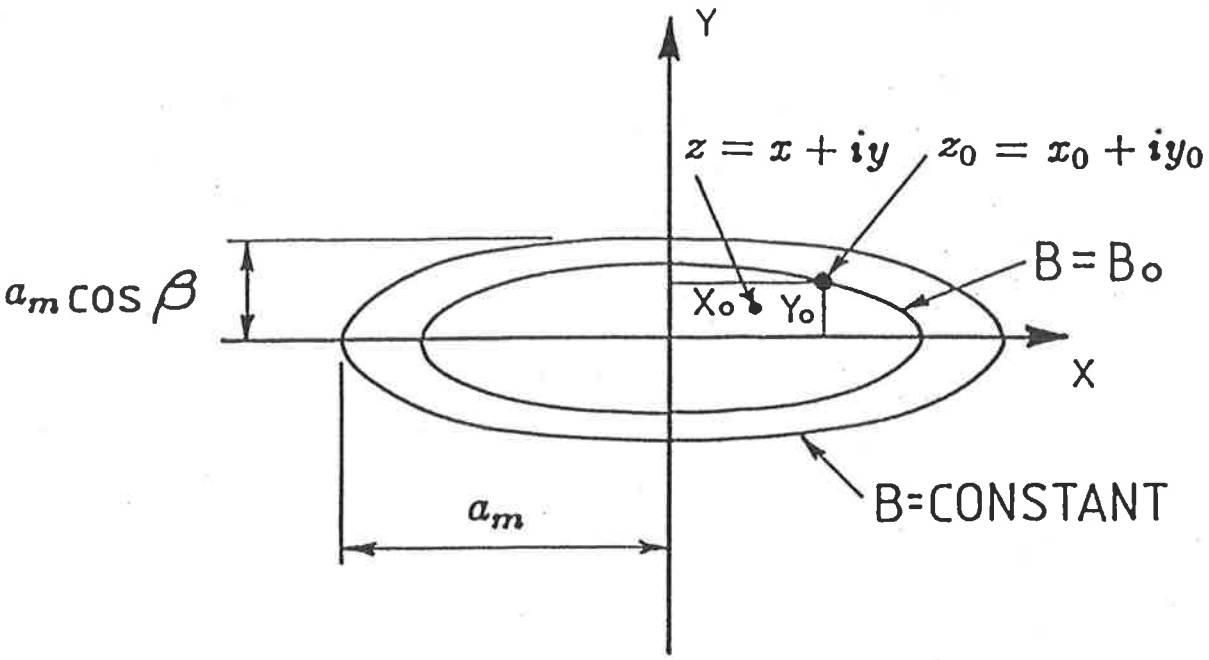

Figure 1.9 Projected spheroid. On projection, the spheroid becomes a solid ellipse. If the eccentricity of the spheroid is e, then the eccentricity of the solid ellipse is $e \sin\gamma = \sin\beta$, which defines β. The length a_m is the semi-major axis of the spheroid, and hence of the solid ellipse. The impact parameter is z_0, and the resultant projected position of the source is z.

The gravitational field is assumed to be very weak and so we use Einstein's linearised theory for the bending of a ray of light about a point mass. Assuming that the extension of the spheroid is negligible when compared with D_d and D_{ds} the bending angle for each z_0 can be calculated using the solid ellipse formed by the projected spheroid. This ellipse is taken to consist of many point masses, the contribution of each to the bending angle is summed to give the total bending angle.

Einstein's linearised theory states that the bending angle for a ray of light about a point mass is

$$\alpha = -\frac{4GM}{c^2 r},$$

or in vector form

$$\vec{\alpha} = -\frac{4GM}{c^2 r}\frac{\vec{r}}{r},$$

and in complex form

$$\alpha = -\frac{4GMr}{c^2 r r^*} = -\frac{4GM}{c^2 r^*}. \qquad (1.1)$$

So finally, the bending angle for a ray of light passing through z_0 is

$$\alpha = -\frac{4G}{c^2}\int_{\mathcal{E}} \frac{dM(\omega)}{z_0^* - \omega^*} = -\frac{4G}{c^2} I^*(z_0), \tag{1.2}$$

where the integration is carried out over all of the points ω of the solid ellipse $\mathcal{E}$. By construction of fig.1.8 and fig.1.9, we can see that

$$z = z_0 - \frac{4DG}{c^2} I^*(z_0), \tag{1.3}$$

where $D = D_{ds}D_d/D_s$.

Suppose we model the galaxy which is acting as a gravitational lens by an oblate spheroid of eccentricity e, semi-major axis a_m, and whose axis forms an angle γ with the line joining the centre of the spheroid with the observer. The following calculations will work for a prolate spheroid if e is replaced by ie. The solid ellipse which results from projecting the spheroid onto the plane mentioned above, then has eccentricity $e\sin\gamma$ and semi-major axis a_m. For convenience define an angle β by $\sin\beta = e\sin\gamma$. If the spheroid's density has spheroidal symmetry then the solid ellipse's density has elliptical symmetry. By spheroidally symmetric density we mean that the density is constant on each of the concentric spheroidal shells that make up the solid spheroid. By elliptically symmetric density we mean that the density is constant on each of the concentric ellipses which make up the solid ellipse. In fact if $\rho(a)$ is the spheroid's density and $\sigma(B)$ is the solid ellipse's density, then the two are related via

$$\sigma(B) = \frac{2(1-e^2)^{1/2}}{\cos\beta}\int_B^{a_m} \frac{\rho(a)a\,da}{(a^2 - B^2)^{1/2}}, \tag{1.4}$$

where a is the semi-major axis of a spheroidal shell with $0 \leq a \leq a_m$, and B is the semi-major axis of an ellipse inside the solid ellipse with $0 \leq B \leq a_m$. Due to the inherent symmetries in fig.1.9, it is clear that if z_0 is reflected in the real axis then so should the resultant z, similarly for the imaginary axis. Stating this

more precisely we have

$$\begin{aligned} x+iy &= x_0+iy_0-\frac{4DG}{c^2}I^*(x_0+iy_0) \\ \Rightarrow \quad -x-iy &= -x_0-iy_0-\frac{4DG}{c^2}I^*(-x_0-iy_0) \\ \Rightarrow \quad x-iy &= x_0-iy_0-\frac{4DG}{c^2}I^*(x_0-iy_0) \\ \Rightarrow \quad -x+iy &= -x_0+iy_0-\frac{4DG}{c^2}I^*(-x_0+iy_0). \end{aligned} \tag{1.5}$$

The expression for $I(z_0)$ derived by Bourassa and Kantowski (1975) is

$$I(z_0) = 2\pi\cos\beta\int_0^{\min(a_m,B_0)}\frac{\sigma(B)B\,dB}{(z_0^2-B^2\sin^2\beta)^{1/2}}, \tag{1.6}$$

where B_0 is defined by ${x_0}^2+{y_0}^2\cos^{-2}\beta={B_0}^2$.

Clearly since $I(z_0)=I(-z_0)$ the symmetry principle is violated, as from the first two lines we deduce that $z=z_0$.

In the next chapter we derive the correct expression for $I(z_0)$ which is

$$I(z_0) = \mathrm{sign}(z_0)2\pi\cos\beta\int_0^{\min(a_m,B_0)}\frac{\sigma(B)B\,dB}{(z_0^2-B^2\sin^2\beta)^{1/2}}, \tag{1.7}$$

where B_0 is defined as above, and where we define $\mathrm{sign}(z_0)=\sqrt{z_0^2}/z_0$. The derivation contains the proof of the fact that the mass outside of the solid ellipse with semi-major axis B_0 does not contribute to the bending angle.

§2 Derivation of $I(z_0)$

Equation (1.2) defined $I(z_0)$ to be

$$I(z_0) = \int_{\mathcal{E}} \frac{dM(\omega)}{z_0 - \omega}, \tag{2.1}$$

where the region of integration, $\mathcal{E}$, is the surface of the ellipse of surface density $\sigma(B)$, semi-major axis a_m and eccentricity $e \sin\gamma = \sin\beta$. Therefore,

$$I(z_0) = \int_{-a_m}^{a_m} dx \int_{-\cos\beta(a_m^2 - x^2)^{1/2}}^{\cos\beta(a_m^2 - x^2)^{1/2}} dy \frac{\sigma((x^2 + y^2\cos^{-2}\beta)^{1/2})}{x_0 - x + i(y_0 - y)}, \tag{2.2}$$

where we let $z_0 = x_0 + iy_0$ and $\omega = x + iy$. Let $x = B\cos\theta$ and $y = B\cos\beta\sin\theta$, then

$$I(z_0) = \cos\beta \int_0^{a_m} dB\sigma(B)B \int_0^{2\pi} \frac{d\theta}{z_0 - B\cos\theta - iB\cos\beta\sin\theta}. \tag{2.3}$$

Consider the latter integral of (2.3)

$$\int_0^{2\pi} \frac{d\theta}{z_0 - B\cos\theta - iB\cos\beta\sin\theta}. \tag{2.4}$$

To evaluate (2.4) we make the substitution $t = \tan(\theta/2)$ and employ complex integration theory. Equation (2.4) becomes

$$2\int_{-\infty}^{\infty} \frac{dt}{(z_0 + B)t^2 - i2B\cos\beta t + z_0 - B}. \tag{2.5}$$

The integrand has two poles at

$$t_{\pm} = \frac{iB\cos\beta \pm i(z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B}. \tag{2.6}$$

Write $t_{\pm}$ as $u_{\pm} + iv_{\pm}$, and taking the standard semi-circle contour in the upper half of the plane we have

$$\int_0^{2\pi} \frac{d\theta}{z_0 - B\cos\theta - iB\cos\beta\sin\theta}$$

$$= \frac{2\pi}{(z_0^2 - B^2\sin^2\beta)^{1/2}} \qquad \text{if } v_+ > 0 \text{ and } v_- < 0$$

$$= -\frac{2\pi}{(z_0^2 - B^2\sin^2\beta)^{1/2}} \qquad \text{if } v_- > 0 \text{ and } v_+ < 0$$

$$= 0 \qquad \text{if } v_+ v_- > 0. \tag{2.7}$$

The case where $v_+ = 0$ or $v_- = 0$ will be handled later.

In order to prove that the mass outside the ellipse of semi-major axis B_0 does not contribute to the bending angle, we need to prove that $v_+ v_- > 0$ for $B > B_0$. Now,

$$u_+ + iv_+ = \frac{iB\cos\beta + i(z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B}$$

$$\Rightarrow u_+ - iv_+ = \frac{-iB\cos\beta - i(z_0^{*2} - B^2\sin^2\beta)^{1/2}}{z_0^* + B}$$

$$\Rightarrow 2iv_+ = \frac{iB\cos\beta + i(z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B} + \frac{iB\cos\beta + i(z_0^{*2} - B^2\sin^2\beta)^{1/2}}{z_0^* + B}.$$

Similarly

$$2iv_- = \frac{iB\cos\beta - i(z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B} + \frac{iB\cos\beta - i(z_0^{*2} - B^2\sin^2\beta)^{1/2}}{z_0^* + B}.$$

So we have

$$4v_+ v_- = \left(\frac{B\cos\beta + (z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B} + \frac{B\cos\beta + (z_0^{*2} - B^2\sin^2\beta)^{1/2}}{z_0^* + B}\right)$$

$$\times \left(\frac{B\cos\beta - (z_0^2 - B^2\sin^2\beta)^{1/2}}{z_0 + B} + \frac{B\cos\beta - (z_0^{*2} - B^2\sin^2\beta)^{1/2}}{z_0^* + B}\right)$$

$$= \frac{B^2\cos^2\beta - (z_0^2 - B^2\sin^2\beta)}{(z_0 + B)^2} + \frac{B^2\cos^2\beta - (z_0^{*2} - B^2\sin^2\beta)}{(z_0^* + B)^2}$$

$$+\frac{(B\cos\beta+(z_0^{*2}-B^2\sin^2\beta)^{1/2})(B\cos\beta-(z_0^2-B^2\sin^2\beta)^{1/2})}{|z_0+B|^2}$$

$$+\frac{(B\cos\beta+(z_0^2-B^2\sin^2\beta)^{1/2})(B\cos\beta-(z_0^{*2}-B^2\sin^2\beta)^{1/2})}{|z_0+B|^2}$$

$$=\frac{B^2\cos^2\beta-(z_0^2-B^2\sin^2\beta)}{(z_0+B)^2}+\frac{2B^2\cos^2\beta-2|z_0^2-B^2\sin^2\beta|}{|z_0+B|^2}$$

$$+\frac{B^2\cos^2\beta-(z_0^{*2}-B^2\sin^2\beta)}{(z_0^*+B)^2}$$

$$=\frac{(z_0^*+B)^2(B^2-z_0^2)+2|z_0+B|^2(B^2\cos^2\beta-|z_0^2-B^2\sin^2\beta|)}{|z_0+B|^4}$$

$$+\frac{(z_0+B)^2(B^2-z_0^{*2})}{|z_0+B|^4}$$

$$=\frac{(z_0^{*2}+2z_0^*B+B^2)(B^2-z_0^2)+2|z_0+B|^2(B^2\cos^2\beta-|z_0^2-B^2\sin^2\beta|)}{|z_0+B|^4}$$

$$+\frac{(z_0^2+2Bz_0+B^2)(B^2-z_0^{*2})}{|z_0+B|^4}$$

$$=\frac{-2|z_0|^4+2B(z_0^*+z_0)(B^2-|z_0|^2+2B^4)}{|z_0+B|^4}$$

$$+\frac{2|z_0+B|^2(B^2\cos^2\beta-|z_0^2-B^2\sin^2\beta|)}{|z_0+B|^4}.$$

On replacing z_0 by x_0+iy_0 where convenient, we have

$$2|z_0+B|^4v_+v_-=-(x_0^2+y_0^2)^2+2Bx_0(B^2-(x_0^2+y_0^2)+B^4)$$
$$+|z_0+B|^2(B^2\cos^2\beta-|z_0^2-B^2\sin^2\beta|)$$

$$=B^3(2x_0+B)-(x_0^2+y_0^2)(x_0^2+y_0^2+2Bx)+$$
$$+((x_0+B)^2+y_0^2)(B^2\cos^2\beta-|z_0^2-B^2\sin^2\beta|)$$

$$= 2B^3(x_0 + B) - (x_0^2 + y_0^2 - B^2)(x_0^2 + y_0^2 + 2Bx_0)$$
$$- ((x_0 + B)^2 + y_0^2)(B^2 \sin^2\beta + |z_0^2 - B^2 \sin^2\beta|)$$

$$= 2B^3(x_0 + B) - (x_0^2 + y_0^2 - B^2)(x_0^2 + y_0^2 + 2Bx_0)$$
$$-((x_0 + B)^2 + y_0^2)(B^2 \sin^2\beta + \left((x_0^2 - y_0^2 - B^2 \sin^2\beta)^2 + 4x_0^2 y_0^2\right)^{1/2}).$$

Using the relation $B_0^2 = x_0^2 + y_0^2 \cos^{-2}\beta$ we substitute for y_0^2, then above becomes

$$= 2B^3(x_0 + B) - (\sin^2\beta x_0^2 + \cos^2\beta B_0^2 - B^2)$$
$$\times (\sin^2\beta x_0^2 + \cos^2\beta B_0^2 + 2Bx_0)$$

$$- ((x_0 + B)^2 + \cos^2\beta(B_0^2 - x_0^2))$$
$$\times(B^2 \sin^2\beta + \left((x_0^2 - \cos^2\beta(B_0^2 - x_0^2) - B^2 \sin^2\beta)^2 + 4x_0^2 \cos^2\beta(B_0^2 - x_0^2)\right)^{1/2}).$$

Now define k and c by the relations $x_0 = B_0 k$ and $B_0 = cB$. Therefore, $-1 \leq k \leq 1$, $0 \leq c < \infty$, and we now wish to prove that $v_+ v_- > 0$ for $0 \leq c < 1$. So now,

$$\frac{2|z_0 + B|^4 v_+ v_-}{B^4} = 2(kc + 1) - (c^2 p - 1)(c^2 p + 2kc)$$
$$- ((kc + 1)^2 + c^2 \cos^2\beta(1 - k^2))$$
$$\times (\sin^2\beta + \left((2k^2c^2 - c^2 p - \sin^2\beta)^2 + 4k^2c^4 \cos^2\beta(1 - k^2)\right)^{1/2}),$$

where $p = 1 - \sin^2\beta(1 - k^2)$,

$$= 2(kc + 1) - (c^2 p - 1)(c^2 p + 2kc)$$
$$- (2kc + 1 - c^2(\sin^2\beta(1 - k^2) - 1))$$

$$\times(\sin^2\beta + \big((2k^2c^2 - c^2p - \sin^2\beta)^2 + 4k^2c^4(1-\sin^2\beta)(1-k^2)\big)^{1/2})$$

$$= 2(kc+1) - (c^2p-1)(c^2p+2kc) - \sin^2\beta(c^2p+2kc+1)$$
$$-(c^2p+2kc+1)\big((2k^2c^2 - c^2p - \sin^2\beta)^2 + 4k^2c^4(1-\sin^2\beta)(1-k^2)\big)^{1/2}$$

$$= 2(kc+1) - c^2p(c^2p+2kc) + c^2p + 2kc - \sin^2\beta(c^2p+2kc+1)$$
$$-(c^2p+2kc+1)\big((2k^2c^2 - c^2p - \sin^2\beta)^2 + 4k^2c^4(1-\sin^2\beta)(1-k^2)\big)^{1/2}$$

$$= 2(c^2p+2kc+1) - (c^2p+\sin^2\beta)(c^2p+2kc+1)$$
$$-(c^2p+2kc+1)\big((2k^2c^2 - c^2p - \sin^2\beta)^2 + 4k^2c^4(1-\sin^2\beta)(1-k^2)\big)^{1/2}$$

$$= (c^2p+2kc+1)\Big(2 - \sin^2\beta - c^2p$$
$$-\big(4k^4c^4 + c^4p^2 + \sin^4\beta - 4k^2c^4p - 4k^2c^2\sin^2\beta + 2c^2p\sin^2\beta + 4k^2c^4$$
$$(1 - \sin^2\beta - k^2 + \sin^2\beta k^2)\big)^{1/2}\Big)$$

$$= (c^2p+2kc+1)\Big(2 - \sin^2\beta - c^2p$$
$$-\big(4k^4c^4 + c^4p^2 + \sin^4\beta - 4k^2c^4p - 4k^2c^2\sin^2\beta + 2c^2p\sin^2\beta + 4k^2c^4(p-k^2)\big)^{1/2}\Big)$$

$$= (c^2p+2kc+1)\Big(2 - \sin^2\beta - c^2p$$
$$-\big((c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta\big)^{1/2}\Big)$$

$$= \frac{2|z_0+B|^4 v_+ v_-}{B^4}. \tag{2.8}$$

Since $-1 \le k \le 1$ and $0 \le \sin^2\beta < 1$, it follows that

$$0 < p = 1 - \sin^2\beta(1-k^2) \le 1$$

and that

$$c^2p + 2kc + 1 = (kc+1)^2 + c^2(1-\sin^2\beta)(1-k^2) > 0.$$

Thus, by equation (2.8)

$$v_+v_- > 0 \iff 2 - \sin^2\beta - c^2p - \left((c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta\right)^{1/2} > 0. \quad (2.9)$$

In order to prove our hypothesis, $v_+v_- > 0$ for $0 \leq c < 1$, we use the following method. We find a function $h(c)$ which satisfies the following inequality

$$h(c) \leq 2 - \sin^2\beta - c^2p - \left((c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta\right)^{1/2}, \quad (2.10)$$

and then we show that $h(c) > 0$ for $0 \leq c < 1$, thus proving our hypothesis.

To find $h(c)$ we replace the irrational part of (2.10) by a rational function, i.e. Let

$$h(c) = 2 - \sin^2\beta - c^2p - g(c),$$

where g(c) is some rational function. As the right hand side of (2.10) is zero for $c = 1$, we must have that $h(1) = 0$, which implies that $g(1) = 1 - k^2\sin^2\beta$. Furthermore, in order that (2.10) be satisfied, $g(c)$ must obey the following inequality

$$0 \leq \left((c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta\right)^{1/2} \leq g(c). \quad (2.11)$$

Let

$$g(c) = \sin^2\beta + c^2(1 - \sin^2\beta(k^2+1)).$$

Claim: $g(c)$ satisfies (2.11).

Proof:

$$(c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta - g^2(c)$$

$$= (c^2p+\sin^2\beta)^2 - 4k^2c^2\sin^2\beta - (\sin^2\beta + c^2(1-\sin^2\beta(k^2+1)))^2$$

$$= c^4p^2 + \sin^4\beta + 2c^2p\sin^2\beta - 4k^2c^2\sin^2\beta - \sin^4\beta$$
$$- c^4(1 - k^2\sin^2\beta - \sin^2\beta)^2 - 2c^2\sin^2\beta(1 - k^2\sin^2\beta - \sin^2\beta)$$

$$= c^4p^2 + 2c^2p\sin^2\beta - 4k^2c^2\sin^2\beta - c^4(p - 2k^2\sin^2\beta)^2$$
$$- 2c^2\sin^2\beta(p - 2k^2\sin^2\beta)$$

$$= c^4p^2 - 4k^2c^2\sin^2\beta(1 - \sin^2\beta) - c^4p^2 + 4c^4pk^2\sin^2\beta - 4c^4k^4\sin^4\beta$$

$$= -4k^2c^2\sin^2\beta(1 - \sin^2\beta) - 4c^4k^2\sin^2\beta(p - k^2\sin^2\beta)$$

$$= -4c^2k^2\sin^2\beta(1 - \sin^2\beta)(1 + c^2) \leq 0.$$

Thus,

$$g^2(c) \geq (c^2p + \sin^2\beta)^2 - 4k^2c^2\sin^2\beta,$$

and therefore

$$g(c) \geq \left((c^2p + \sin^2\beta)^2 - 4k^2c^2\sin^2\beta\right)^{1/2}. \qquad \text{QED}$$

We now show that $h(c) > 0$ for $0 \leq c < 1$, which proves our hypothesis.

$$\begin{aligned} h(c) &= 2 - \sin^2\beta - c^2p - \sin^2\beta - c^2(1 - \sin^2\beta(k^2 + 1)) \\ &= (1 - \sin^2\beta)(2 - c^2) - c^2p + c^2k^2\sin^2\beta \\ &= (1 - \sin^2\beta)(2 - c^2) - c^2(1 - \sin^2\beta(1 - k^2)) + c^2k^2\sin^2\beta \\ &= 2(1 - c^2)(1 - \sin^2\beta) \\ &> 0 \quad \text{for } 0 \leq c < 1, \text{ as } 0 \leq \sin^2\beta < 1. \end{aligned} \qquad (2.12)$$

Thus, by equations (2.12), (2.10) and (2.9) we have that $v_+v_- > 0$ for $0 \leq c < 1$.

As c was defined by $B_0 = cB$, this implies that

$$v_+ v_- > 0 \quad \text{for} \quad B > B_0. \qquad \text{QED}$$

This, with the assistance of equation (2.7), proves that the mass outside the solid ellipse of semi-major axis B_0 does not contribute to the bending angle of a ray of light with the impact parameter z_0, where B_0 and z_0 are related via $x_0^2 + y_0^2 \cos^{-2}\beta = B_0^2$.

We have considered the case $B > B_0$, now for the case $B = B_0$. Putting $B = B_0$ in equation (2.6) we find that

$$t_\pm = \frac{iB_0 \cos\beta \pm i(z_0^2 - B_0^2 \sin^2\beta)^{1/2}}{z_0 + B_0},$$

substituting for B_0^2 and z_0^2,

$$= \frac{iB_0 \cos\beta \pm i(x_0^2 - y_0^2 + i2x_0y_0 - \sin^2\beta(x_0^2 + y_0^2\cos^{-2}\beta))^{1/2}}{z_0 + B_0},$$

assuming that $(z_0^2)^{1/2} = z_0$ and not $-z_0$,

$$= \frac{iB_0 \cos\beta \pm i(x_0 \cos\beta + iy_0 \cos^{-1}\beta)}{z_0 + B_0}$$

$$= \frac{(iB_0 \cos\beta \pm i(x_0 \cos\beta + iy_0 \cos^{-1}\beta))(x_0 + B - iy_0)}{(x_0 + B_0)^2 + y_0^2}$$

$$= \cos\beta \frac{y_0(B_0 \pm x_0 \mp (x_0 + B_0)\cos^{-2}\beta) + i((B_0 + x_0)(B_0 \pm x_0) \pm (B_0^2 - x_0^2))}{(x_0 + B_0)^2 + y_0^2}.$$

So for $(z_0^2)^{1/2} = z_0$,

$$t_+ = -\cos\beta \frac{y_0(x_0 + B_0)\tan^2\beta - i2B_0(B_0 + x_0)}{(x_0 + B_0)^2 + y_0^2} \qquad (2.13)$$

$$t_- = y_0 \cos\beta \frac{B_0 - x_0 + (x_0 + B_0)\cos^{-2}\beta}{(x_0 + B_0)^2 + y_0^2}. \qquad (2.14)$$

If $(z_0^2)^{1/2} = -z_0$ then t_+ and t_- simply interchange.

We see that in the case $B = B_0$, either $v_- = \mathrm{Im}(t_-) = 0$ or $v_+ = \mathrm{Im}(t_+) = 0$. Thus there is a real singularity in the case $B = B_0$. Since this singularity is of the first order i.e., it is of the form $(t - u_\pm)^{-1}$, no problems arise, as the principle value of $\int_{-\infty}^{\infty} \frac{dt}{t-u_\pm} = 0$. The reason that the real singularity arises is that in the case $B = B_0$ we integrate through the point z_0. For $B < B_0$ this can't happen and therefore we infer that $v_\pm \neq 0$, $\forall\ B \in [0, B_0)$. As the integral is neither divergent nor zero for $B = B_0$ we shall evaluate it for completeness. However, we leave this till later for reasons which shall become apparent.

Thus, having shown that $I(z_0) = 0$ for $B > B_0$, equation (2.3) becomes

$$I(z_0) = \cos\beta \int_0^{\min(a_m, B_0)} dB\sigma(B)B \int_0^{2\pi} \frac{d\theta}{z_0 - B\cos\theta - iB\cos\beta\sin\theta},$$

where now B ranges from 0 to B_0, if $B_0 < a_m$, rather than to a_m as before. Using the result of (2.7) we have

$$I(z_0) = 2\pi\cos\beta \int_0^{\min(a_m, B_0)} \frac{\sigma(B)B\,dB}{(z_0^2 - B^2\sin^2\beta)^{1/2}} \qquad \text{if } v_+ > 0 \text{ and } v_- < 0$$

$$= -2\pi\cos\beta \int_0^{\min(a_m, B_0)} \frac{\sigma(B)B\,dB}{(z_0^2 - B^2\sin^2\beta)^{1/2}} \qquad \text{if } v_- > 0 \text{ and } v_+ < 0.$$

The contribution to the total bending angle due to the ellipse of semi-major axis B and thickness dB is

$$\pm 2\pi\cos\beta \frac{\sigma(B)B\,dB}{(z_0^2 - B^2\sin^2\beta)^{1/2}},$$

where the sign is determined by v_+ and v_-. As the impact parameter z_0 lies on or outside the ellipse, this contribution is non-zero for $B \in (0, B_0]$. Given a fixed z_0, we expect this contribution to be a continuous function of B. So as this contribution is never zero, the sign must be independent of B. Thus, we can select

the appropriate sign by evaluating v_+ and v_- in the limit B goes to 0. Now, by equation (2.6)

$$\lim_{B\to 0} t_\pm = u_\pm + iv_\pm = \pm\frac{i(z_0^2)^{1/2}}{z_0},$$

so

$$\lim_{B\to 0} v_+ = \quad 1 \qquad \text{if } (z_0^2)^{1/2} = \quad z_0$$

$$= -1 \qquad \text{if } (z_0^2)^{1/2} = -z_0,$$

whilst

$$\lim_{B\to 0} v_- = -1 \qquad \text{if } (z_0^2)^{1/2} = \quad z_0$$

$$= \quad 1 \qquad \text{if } (z_0^2)^{1/2} = -z_0.$$

Therefore we can write

$$I(z_0) = \operatorname{sign}(z_0) 2\pi \cos\beta \int_0^{\min(a_m, B_0)} \frac{\sigma(B) B\, dB}{(z_0^2 - B^2 \sin^2\beta)^{1/2}}, \qquad (2.15)$$

where we defined $\operatorname{sign}(z_0)$ by

$$\operatorname{sign}(z_0) = \frac{(z_0^2)^{1/2}}{z_0}.$$

Equation (2.15) satisfies the symmetry requirements of §1.2.

Now, for completeness, we evaluate equation (2.5) for $B = B_0$.

$$2\int_{-\infty}^{\infty} \frac{dt}{(z_0 + B_0)t^2 - i2B_0 \cos\beta t + z_0 - B_0} = 2\int_{-\infty}^{\infty} \frac{dt}{(z_0 + B_0)(t - t_-)(t - t_+)},$$

where

$$t_\pm = \frac{iB_0 \cos\beta \pm i(z_0^2 - B_0^2 \sin^2\beta)^{1/2}}{z_0 + B_0},$$

with t_- being real if $(z_0^2)^{1/2} = z_0$ and t_+ being real if $(z_0^2)^{1/2} = -z_0$. The reason we left this calculation till this point is that we needed the above continuity argument to lead to the definition of the function $\text{sign}(z_0)$ which occurs naturally here. Thus, t_- is real if $\text{sign}(z_0) = 1$ and t_+ is real if $\text{sign}(z_0) = -1$. Now,

$$2\int_{-\infty}^{\infty} \frac{dt}{(z_0 + B_0)(t - t_-)(t - t_+)}$$

$$= \frac{2}{(z_0 + B_0)} \int_{-\infty}^{\infty} \frac{dt}{(t_- - t_+)} \left(\frac{1}{t - t_-} - \frac{1}{t - t_+} \right)$$

$$= \frac{-2}{(z_0 + B_0)(t_- - t_+)} \int_{-\infty}^{\infty} \frac{dt}{t - t_+} \qquad \text{for } \text{sign}(z_0) = \quad 1$$

$$= \frac{2}{(z_0 + B_0)(t_- - t_+)} \int_{-\infty}^{\infty} \frac{dt}{t - t_-} \qquad \text{for } \text{sign}(z_0) = -1.$$

Now, $(z_0 + B_0)(t_- - t_+) = -2i(z_0^2 - B_0^2 \sin^2\beta)^{1/2}$, and since the imaginary part of the pole in either case is always positive, see equation (2.13) , $\int_{-\infty}^{\infty} \frac{dt}{(t - t_\pm)} = 2\pi i$. Thus

$$2\int_{-\infty}^{\infty} \frac{dt}{(z_0 + B_0)(t - t_-)(t - t_+)} = \frac{2\pi}{(z_0^2 - B_0^2 \sin^2\beta)^{1/2}} \qquad \text{sign}(z_0) = \quad 1$$

$$= \frac{-2\pi}{(z_0^2 - B_0^2 \sin^2\beta)^{1/2}} \qquad \text{sign}(z_0) = -1$$

$$= \frac{\text{sign}(z_0) 2\pi}{(z_0^2 - B_0^2 \sin^2\beta)^{1/2}},$$

which is cosistent with equation (2.15).

The only difference between the new $I(z_0)$ defined by equation (2.15) and that of Bourassa and Kantowski is that of a minus sign whenever $\text{sign}(z_0) = -1$.

Whilst $I(z_0)$ of Bourassa and Kantowski has z_0 appear as z_0^2 only, their subsequent evaluations of $I(z_0)$ for the various density models have $I(z_0) \neq I(-z_0)$. Thus, in the next chapter we evaluate our own expressions for $I(z_0)$ for the various densities, directly from equation (2.15).

§3 Numerical Investigation

§3.1 Evaluations of $I(z_0)$

Recall that it was the misconception that the gravitational effect would lead to two images of equal brightness only if the deflector is between the images, and that any radio emissions associated with the source should be effected in the same way, that caused the radio astronomers to reject the gravitational lens hypothesis. Later, it was shown that the gravitational lens effect may cause many more images than was previously thought, and that the number of images depends on the angular separation of the source and the deflector. As a result, the intensity of images was no longer as simple to predict as was previously thought.

It is the realisation that multiplicity of images depends on the angular separation of the source and the deflecting mass that attracted our interest. How does this aspect of the gravitational lens effect manifest itself for the various models of the deflecting mass? What do the regions, where a source has many images look like? These and associated questions we answer in this chapter.

In the previous chapters we found that for an oblate spheroid of eccentricity e, orientation angle γ, semi-major axis a_m and density $\rho(a)$, the projected source position z and the projected image position z_0 are related via

$$z = z_0 - \frac{4DG}{c^2} I^*(z_0), \tag{3.1}$$

where

$$I(z_0) = \operatorname{sign}(z_0) 2\pi \cos\beta \int_0^{\min(a_m, B_0)} \frac{\sigma(B) B \, dB}{(z_0^2 - B^2 \sin^2\beta)^{1/2}} \tag{3.2}$$

and

$$\sigma(B) = \frac{2(1-e^2)^{1/2}}{\cos\beta} \int_B^{a_m} \frac{\rho(a) a \, da}{(a^2 - B^2)^{1/2}}. \tag{3.3}$$

The function $\operatorname{sign}(z_0)$ is defined as $\operatorname{sign}(z_0) = \sqrt{z_0^2}/z_0$ and β is defined by $\sin\beta = e \sin\gamma$.

The choice of the density function $\rho(a)$ has been the subject of a great deal of study: Perek (1962), King (1966), Schmidt (1965). Many models of star clusters

are available, however it is not our intention to choose one over another, but instead, compare them with each other. We do this as we are interested in the gravitational lens effect of any massive object that can be suitably modelled by an oblate spheroid, not only a galaxy. The models for densities that we will choose are the Schmidt model where $\rho(a) \propto a^{-1}$, the truncated King model, favoured by most to be the most suitable model for a galaxy, and the simple model where $\rho(a) = \rho_0$, a constant. We evaluate expressions for $\sigma(B)$ and $I(z_0)$ for the above models in order of simplicity. In subsequent expressions we assume that $B_0 < a_m$, thus if $B_0 \geq a_m$ then all occurences of B_0 should be replaced by a_m.

For the constant density model

$$\begin{aligned} \rho_c(a) &= \frac{3M}{4\pi(1-e^2)^{1/2}a_m^3} && a \leq a_m \\ &= 0 && a > a_m \end{aligned} \tag{3.4}$$

Therefore, using equation (3.3)

$$\begin{aligned} \sigma_c(B) &= \frac{3M}{2\pi\cos\beta a_m^3}\int_B^{a_m} \frac{a\,da}{\sqrt{a^2-B^2}} \\ &= \frac{3M}{2\pi\cos\beta a_m^3}\sqrt{a_m^2-B^2}. \end{aligned} \tag{3.5}$$

To find $I_c(z_0)$, the $I(z_0)$ function for the constant density model, we need to consider two possible cases. The general case when $0 < \sin\beta < 1$, and the case when $\sin\beta = 0$. The latter corresponds to the projected mass being a solid circle i.e., the spheroid is either a sphere, or that it is observed along its axis in which case it appears to be a sphere.

Suppose that $\sin\beta = 0$, substituting equation (3.5) into (3.2) we have that

$$I_c(z_0) = \frac{\mathrm{sign}(z_0)3M}{a_m^3}\int_0^{B_0} \frac{B\sqrt{a_m^2-B^2}\,dB}{\sqrt{z_0^2}}$$

$$= -\frac{M}{z_0 a_m^3}\left(a_m^2 - B^2\right)^{\frac{3}{2}}\Big]_0^{B_0}$$

$$= \frac{M}{z_0}\left(1 - \left(1 - \frac{B_0^2}{a_m^2}\right)^{\frac{3}{2}}\right). \qquad (3.6a)$$

For $\sin\beta \neq 0$ define $K = \text{sign}(z_0)3M/a_m^3$, then

$$I_c(z_0) = K\int_0^{B_0} \frac{B\sqrt{a_m^2 - B^2}\,dB}{\sqrt{z_0^2 - \sin^2\beta B^2}}. \qquad \text{Let } x = B^2, \text{ then this becomes}$$

$$= \frac{K}{2}\int dx \frac{\sqrt{a_m^2 - x}}{\sqrt{z_0^2 - \sin^2\beta x}}$$

$$= \frac{K}{2}\int dx \frac{a_m^2 - x}{\sqrt{\sin^2\beta x^2 - (a_m^2\sin^2\beta + z_0^2)x + a_m^2 z_0^2}}$$

$$= -\frac{K}{4\sin\beta}\int dx \frac{2x - 2a_m^2 - a_m^2 - \frac{z_0^2}{\sin^2\beta} + a_m^2 + \frac{z_0^2}{\sin^2\beta}}{\sqrt{x^2 - \left(a_m^2 + \frac{z_0^2}{\sin^2\beta}\right)x + \frac{a_m^2 z_0^2}{\sin^2\beta}}}$$

$$= -\frac{K}{2\sin\beta}\sqrt{x^2 - \left(a_m^2 + \frac{z_0^2}{\sin^2\beta}\right)x + \frac{a_m^2 z_0^2}{\sin^2\beta}}\Bigg]^x$$

$$+\frac{K}{4\sin\beta}\int dx \frac{a_m^2 - \frac{z_0^2}{\sin^2\beta}}{\sqrt{x^2 - \left(a_m^2 + \frac{z_0^2}{\sin^2\beta}\right)x + \frac{a_m^2 z_0^2}{\sin^2\beta}}}$$

$$= -\frac{K}{2\sin\beta}\sqrt{B^4 - \left(a_m^2 + \frac{z_0^2}{\sin^2\beta}\right)B^2 + \frac{a_m^2 z_0^2}{\sin^2\beta}}\Bigg]_0^{B_0}$$

$$+\frac{K\left(a_m^2 - \frac{z_0^2}{\sin^2\beta}\right)}{4\sin\beta}$$

$$\times \ln\left(2\sqrt{B^4 - \left(a_m^2 + \frac{z_0^2}{\sin^2\beta}\right)B^2 + \frac{a_m^2 z_0^2}{\sin^2\beta}} + 2B^2 - a_m^2 - \frac{z_0^2}{\sin^2\beta}\right)\Bigg]_0^{B_0}$$

$$= -\frac{\text{sign}(z_0)3M}{2a_m^3 \sin\beta}\left(\sqrt{(a_m^2 - B_0^2)\left(\frac{z_0^2}{\sin^2\beta} - B_0^2\right)} - \frac{a_m\sqrt{z_0^2}}{\sin\beta}\right) \tag{3.6b}$$

$$+\frac{\text{sign}(z_0)3M\left(a_m^2 - \frac{z_0^2}{\sin^2\beta}\right)}{4a_m^3\sin\beta}\ln\left(\frac{a_m^2 + \frac{z_0^2}{\sin^2\beta} - 2B_0^2 - 2\sqrt{(a_m^2 - B_0^2)\left(\frac{z_0^2}{\sin^2\beta} - B_0^2\right)}}{\left(a_m - \frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right).$$

It is interesting to note that the term $\sqrt{z_0^2}$ arises once again from the limit $B \to 0$. This is consistent with equation (3.2), as it's integrand has z_0 appearing as z_0^2 only.

In the Schmidt model

$$\begin{aligned} \rho_s(a) &= \frac{M}{2\pi(1-e^2)^{1/2}a_m^2}\left(\frac{1}{a}\right) && a \le a_m \\ &= 0 && a > a_m \end{aligned} \tag{3.7}$$

Therefore

$$\begin{aligned} \sigma_s(B) &= \frac{M}{\pi\cos\beta a_m^2}\int_B^{a_m}\frac{da}{\sqrt{a^2 - B^2}} \\ &= \frac{M}{\pi\cos\beta a_m^2}\ln\left(\frac{a_m}{B}\left(1 + \sqrt{1 - \frac{B^2}{a_m^2}}\right)\right). \end{aligned} \tag{3.8}$$

For $\sin\beta = 0$, substituting $\sigma_s(B)$ into equation (3.2), we find that

$$I_s(z_0) = \frac{\text{sign}(z_0)2M}{a_m^2}\int_0^{B_0}\frac{\ln\left(\frac{a_m}{B}\left(1 + \sqrt{1 - \frac{B^2}{a_m^2}}\right)\right)B\,dB}{\sqrt{z_0^2}}$$

$$= \frac{M}{z_0 a_m^2}\left(B_0^2\left(\ln\left(\frac{a_m}{B_0}\right) + \frac{1}{2}\right) + \int_0^{B_0} dB\,\ln\left(1 + \sqrt{1 - \frac{B^2}{a_m^2}}\right)\frac{d}{dB}B^2\right)$$

$$= \frac{M}{z_0 a_m^2}\left(B_0^2\left(\ln\left(\frac{a_m}{B_0}\right)+\frac{1}{2}\right)+B_0^2\ln\left(1+\sqrt{1-\frac{B_0^2}{a_m^2}}\right)+\int_0^{B_0}\frac{dB\,B\left(1-\sqrt{1-\frac{B^2}{a_m^2}}\right)}{\sqrt{1-\frac{B^2}{a_m^2}}}\right)$$

$$= \frac{M}{z_0 a_m^2}\left(B_0^2\ln\left(\frac{a_m}{B_0}\left(1+\sqrt{1-\frac{B_0^2}{a_m^2}}\right)\right)+a_m^2\left(1-\sqrt{1-\frac{B_0^2}{a_m^2}}\right)\right). \qquad (3.9a)$$

For $\sin\beta \neq 0$, define $K = \mathrm{sign}(z_0)2M/a_m^2$, and then we have

$$I_s(z_0) = K\int_0^{B_0}\frac{\ln\left(\frac{a_m}{B}\left(1+\sqrt{1-\frac{B^2}{a_m^2}}\right)\right)B\,dB}{\sqrt{z_0^2-\sin^2\beta B^2}}$$

$$= \frac{-K}{\sin^2\beta}\int_0^{B_0} dB\,\ln\left(\frac{a_m}{B}\left(1+\sqrt{1-\frac{B^2}{a_m^2}}\right)\right)\frac{d}{dB}\sqrt{z_0^2-\sin^2\beta B^2}$$

$$= \frac{-K}{\sin^2\beta}\ln\left(\frac{a_m}{B}\left(1+\sqrt{1-\frac{B^2}{a_m^2}}\right)\right)\sqrt{z_0^2-\sin^2\beta B^2}\Bigg]_0^{B_0}$$

$$+\frac{K}{\sin^2\beta}\int_0^{B_0}\frac{dB\,B\sqrt{z_0^2-\sin^2\beta B^2}}{1+\sqrt{1-\frac{B^2}{a_m^2}}}\frac{d}{dB}\left(\frac{1+\sqrt{1-\frac{B^2}{a_m^2}}}{B}\right),$$

for convenience denote the first term by K,

$$= K+\frac{K}{\sin^2\beta}\int_0^{B_0}\frac{dB\,B\sqrt{z_0^2-\sin^2\beta B^2}}{1+\sqrt{1-\frac{B^2}{a_m^2}}}\frac{\left(-\frac{B^2}{a_m^2}\left(1-\frac{B^2}{a_m^2}\right)^{-\frac{1}{2}}-1-\sqrt{1-\frac{B^2}{a_m^2}}\right)}{B^2}$$

$$= K-\frac{K}{\sin^2\beta}\int_0^{B_0} dB\sqrt{z_0^2-\sin^2\beta B^2}\left(\frac{B}{a_m^2\left(1+\sqrt{1-\frac{B^2}{a_m^2}}\right)\sqrt{1-\frac{B^2}{a_m^2}}}+\frac{1}{B}\right)$$

$$= K-\frac{K}{\sin^2\beta}\int_0^{B_0}\frac{dB\sqrt{z_0^2-\sin^2\beta B^2}}{B\sqrt{1-\frac{B^2}{a_m^2}}}, \qquad \text{let } x = B^2,$$

$$= K-\frac{K}{2\sin^2\beta}\int^{x}\frac{dx}{x}\frac{z_0^2-\sin^2\beta x}{\sqrt{\left(1-\frac{x}{a_m^2}\right)\left(z_0^2-\sin^2\beta x\right)}}$$

$$= K - \frac{K}{2\sin^2\beta}\int^{x} \frac{z_0^2\,dx}{x\sqrt{\frac{\sin^2\beta}{a_m^2}x^2 - \left(\frac{z_0^2}{a_m^2} + \sin^2\beta\right)x + z_0^2}}$$

$$+\frac{K}{2}\int^{x} \frac{dx}{\sqrt{\frac{\sin^2\beta}{a_m^2}x^2 - \left(\frac{z_0^2}{a_m^2} + \sin^2\beta\right)x + z_0^2}}$$

Now, the latter integral

$$\frac{K}{2}\int^{x} \frac{dx}{\sqrt{\frac{\sin^2\beta}{a_m^2}x^2 - \left(\frac{z_0^2}{a_m^2} + \sin^2\beta\right)x + z_0^2}}$$

$$= \frac{a_m K}{2\sin\beta}\ln\left(\frac{2\sin\beta}{a_m}\sqrt{\left(\frac{z_0^2}{a_m^2} - \frac{\sin^2\beta B^2}{a_m^2}\right)(a_m^2 - B^2)} + \frac{2\sin^2\beta B^2}{a_m^2} - \frac{z_0^2}{a_m^2} - \sin^2\beta\right)\Bigg]_0^{B_0}$$

$$= \frac{a_m K}{2\sin\beta}\ln\left(\frac{\frac{z_0^2}{\sin^2\beta} + a_m^2 - 2B_0^2 - 2\sqrt{\left(\frac{z_0^2}{\sin^2\beta} - B_0^2\right)(a_m^2 - B_0^2)}}{\left(a_m - \frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right).$$

Consider the former integral, which is

$$-\frac{K}{2\sin^2\beta}\int^{x} \frac{z_0^2\,dx}{x\sqrt{\frac{\sin^2\beta}{a_m^2}x^2 - \left(\frac{z_0^2}{a_m^2} + \sin^2\beta\right)x + z_0^2}}.$$ Let $t = 1/x$ then it becomes

$$= \frac{Kz_0^2}{2\sin^2\beta}\int^{t} \frac{dt}{\sqrt{z_0^2t^2 - \left(\frac{z_0^2}{a_m^2} + \sin^2\beta\right)t + \frac{\sin^2\beta}{a_m^2}}}$$

$$= \frac{K\sqrt{z_0^2}}{2\sin^2\beta}\ln\left(2\sqrt{z_0^2}\sqrt{(z_0^2t - \sin^2\beta)\left(t - \frac{1}{a_m^2}\right)} + 2z_0^2t - \frac{z_0^2}{a_m^2} - \sin^2\beta\right)\Bigg]^{t}$$

$$= \frac{K\sqrt{z_0^2}}{2\sin^2\beta}\ln\left(2\sqrt{z_0^2}\sqrt{\left(\frac{z_0^2}{B_0^2} - \sin^2\beta\right)\left(\frac{1}{B_0^2} - \frac{1}{a_m^2}\right)} + \frac{2z_0^2}{B_0^2} - \frac{z_0^2}{a_m^2} - \sin^2\beta\right)$$

$$-\frac{K\sqrt{z_0^2}}{\sin^2\beta}\lim_{B\to 0}\ln\left(\frac{2\sqrt{z_0^2}}{B}\right).$$

Putting these results together, we have that

$$I_s(z_0) = \frac{-K}{\sin^2\beta}\ln\left(\frac{a_m}{B}\left(1+\sqrt{1-\frac{B^2}{a_m^2}}\right)\right)\sqrt{z_0^2-\sin^2\beta B^2}\Bigg]_0^{B_0} - \frac{K\sqrt{z_0^2}}{\sin^2\beta}\lim_{B\to 0}\ln\left(\frac{2\sqrt{z_0^2}}{B}\right)$$

$$+\frac{a_m K}{2\sin\beta}\ln\left(\frac{\frac{z_0^2}{\sin^2\beta}+a_m^2-2B_0^2-2\sqrt{\left(\frac{z_0^2}{\sin^2\beta}-B_0^2\right)(a_m^2-B_0^2)}}{\left(a_m-\frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right)$$

$$+\frac{K\sqrt{z_0^2}}{2\sin^2\beta}\ln\left(2\sqrt{z_0^2}\sqrt{\left(\frac{z_0^2}{B_0^2}-\sin^2\beta\right)\left(\frac{1}{B_0^2}-\frac{1}{a_m^2}\right)}+\frac{2z_0^2}{B_0^2}-\frac{z_0^2}{a_m^2}-\sin^2\beta\right)$$

$$=\frac{-\text{sign}(z_0)2M}{a_m^2\sin^2\beta}\left[\ln\left(\frac{a_m}{B_0}\left(1+\sqrt{1-\frac{B_0^2}{a_m^2}}\right)\right)\sqrt{z_0^2-\sin^2\beta B_0^2}+\sqrt{z_0^2}\ln\left(\frac{\sqrt{z_0^2}}{a_m}\right)\right]$$

$$+\frac{\text{sign}(z_0)M}{a_m\sin\beta}\ln\left(\frac{\frac{z_0^2}{\sin^2\beta}+a_m^2-2B_0^2-2\sqrt{\left(\frac{z_0^2}{\sin^2\beta}-B_0^2\right)(a_m^2-B_0^2)}}{\left(a_m-\frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right) \qquad (3.9b)$$

$$+\frac{z_0 M}{a_m^2\sin^2\beta}\ln\left(2\sqrt{z_0^2}\sqrt{\left(\frac{z_0^2}{B_0^2}-\sin^2\beta\right)\left(\frac{1}{B_0^2}-\frac{1}{a_m^2}\right)}+\frac{2z_0^2}{B_0^2}-\frac{z_0^2}{a_m^2}-\sin^2\beta\right).$$

Finally, for the truncated King model

$$\rho_k(a) = \frac{\rho_0}{\left(1+\frac{a^2}{r_c^2}\right)^{\frac{3}{2}}} \qquad \frac{a}{r_c} \le n$$

$$= 0 \qquad \frac{a}{r_c} > n \qquad (3.10)$$

where

$$\rho_0 = \frac{M}{4\pi\delta r_c^3\sqrt{1-e^2}},$$

and where

$$\delta = \ln(\sqrt{1+n^2}+n) - \frac{n}{\sqrt{1+n^2}}.$$

The parameter r_c refers to the core radius, whilst n determines the cut-off radius of the galaxy i.e., $a_m = r_c n$. Therefore, using equations (3.3) and (3.10)

$$\sigma_k(B) = \frac{M}{2\pi\cos\beta r_c^3\delta}\int_B^{nr_c}\frac{a\,da}{\left(1+\frac{a^2}{r_c^2}\right)^{\frac{3}{2}}\sqrt{a^2-B^2}}, \qquad \text{let } x = a^2,$$

$$= \frac{M}{4\pi\cos\beta r_c^3\delta}\int^{x}\frac{dx}{\left(1+\frac{x}{r_c^2}\right)^{\frac{3}{2}}\sqrt{x-B^2}}$$

$$= \frac{M}{4\pi\cos\beta r_c\delta}\int^{x}\frac{dx}{(x+r_c^2)\sqrt{\frac{x^2}{r_c^2}+\left(1-\frac{B^2}{r_c^2}\right)x-B^2}}, \qquad \text{let } t = (x+r_c^2)^{-1},$$

$$= -\frac{M}{4\pi\cos\beta r_c\delta}\int^{t}\frac{dt}{\sqrt{\frac{1}{r_c^2}+\left(1-\frac{B^2}{r_c^2}-\frac{2r_c^2}{r_c^2}\right)t+\left(r_c^2-r_c^2\left(1-\frac{B^2}{r_c^2}\right)-B^2\right)t^2}}$$

$$= -\frac{M}{4\pi\cos\beta r_c\delta}\int^{t}\frac{dt}{\sqrt{\frac{1}{r_c^2}-\left(1+\frac{B^2}{r_c^2}\right)t}}$$

$$= \frac{M}{2\pi\cos\beta r_c\delta\left(1+\frac{B^2}{r_c^2}\right)}\sqrt{\frac{1}{r_c^2}-\left(1+\frac{B^2}{r_c^2}\right)t}\Bigg]^{t}$$

$$= \frac{M}{2\pi\cos\beta\delta\left(r_c^2+B^2\right)}\sqrt{1-\frac{r_c^2+B^2}{a^2+r_c^2}}\Bigg]_{a=B}^{a=nr_c}$$

$$= \frac{M}{2\pi\cos\beta\delta r_c\sqrt{1+n^2}}\frac{\sqrt{r_c^2n^2-B^2}}{(r_c^2+B^2)} \qquad (3.11)$$

Using equations (3.2) and (3.11), we now find $I_k(z_0)$. For $\sin\beta = 0$ we have that

$$I_k(z_0) = \frac{\text{sign}(z_0)M}{r_c\delta\sqrt{1+n^2}}\int_0^{B_0}\frac{dB\,B\sqrt{r_c^2n^2-B^2}}{(r_c^2+B^2)\sqrt{z_0^2}},$$

$$= \frac{M}{2z_0 r_c \delta\sqrt{1+n^2}} \int_0^{B_0^2} \frac{dx\,\sqrt{r_c^2 n^2 - x}}{(r_c^2 + x)}, \qquad \text{where } x = B^2,$$

$$= \frac{M}{2z_0 r_c \delta\sqrt{1+n^2}} \left(2\sqrt{r_c^2 n^2 - x}\Big]_0^{B_0^2} + r_c^2(1+n^2) \int_0^{B_0^2} \frac{dx}{(r_c^2+x)\sqrt{r_c^2 n^2 - x}} \right)$$

$$= \frac{M}{2z_0 r_c \delta\sqrt{1+n^2}} \left(2\sqrt{r_c^2 n^2 - B_0^2} - 2r_c n \right.$$

$$\left. + r_c\sqrt{1+n^2} \ln \left(\frac{\sqrt{r_c^2 n^2 - x} - r_c\sqrt{1+n^2}}{\sqrt{r_c^2 n^2 - x} + r_c\sqrt{1+n^2}} \right) \Big]_0^{B_0^2} \right)$$

$$= \frac{M}{2z_0 r_c \delta\sqrt{1+n^2}} \left(2\sqrt{r_c^2 n^2 - B_0^2} - 2r_c n \right.$$

$$\left. + r_c\sqrt{1+n^2} \ln \left(\left(n + \sqrt{1+n^2}\right)^2 \frac{r_c\sqrt{1+n^2} - \sqrt{r_c^2 n^2 - B_0^2}}{r_c\sqrt{1+n^2} + \sqrt{r_c^2 n^2 - B_0^2}} \right) \right). \quad (3.12a)$$

For $\sin\beta \neq 0$, define $K = \text{sign}(z_0) M/\delta$, then

$$I_k(z_0) = \frac{K}{r_c\sqrt{1+n^2}} \int_0^{B_0} \frac{dB\, B\sqrt{r_c^2 n^2 - B^2}}{(r_c^2 + B^2)\sqrt{z_0^2 - \sin^2\beta B^2}}, \qquad \text{let } x = B^2,$$

$$= \frac{K}{2r_c\sqrt{1+n^2}} \int^{x} \frac{dx\,\sqrt{r_c^2 n^2 - x}}{(r_c^2+x)\sqrt{z_0^2 - \sin^2\beta}}$$

$$= \frac{K}{2r_c \sin\beta\sqrt{1+n^2}} \int^{x} \frac{dx\,\left(r_c^2 n^2 - x\right)}{(r_c^2+x)\sqrt{\left(\frac{z_0^2}{\sin^2\beta}\right)\left(r_c^2 n^2 - x\right)}}$$

$$= \frac{K}{2r_c \sin\beta\sqrt{1+n^2}} \int^{x} dx\, \frac{-1 + \frac{r_c^2(1+n^2)}{r_c^2+x}}{\sqrt{x^2 - \left(r_c^2 n^2 + \frac{z_0^2}{\sin^2\beta}\right)x + r_c^2 n^2 \frac{z_0^2}{\sin^2\beta}}}$$

$$= I_1 + I_2.$$

We first evaluate I_1.

$$I_1 = \frac{-K}{2r_c \sin\beta\sqrt{1+n^2}} \int^{x} \frac{dx}{\sqrt{x^2 - \left(r_c^2 n^2 + \frac{z_0^2}{\sin^2\beta}\right) x + r_c^2 n^2 \frac{z_0^2}{\sin^2\beta}}}$$

$$= \frac{-K}{2r_c \sin\beta\sqrt{1+n^2}} \ln\left(2\sqrt{(r_c^2 n^2 - x)\left(\frac{z_0^2}{\sin^2\beta} - x\right)} + 2x - r_c^2 n^2 - \frac{z_0^2}{\sin^2\beta}\right)\Bigg]_{x=0}^{x=B_0^2}$$

$$= \frac{-K}{2r_c \sin\beta\sqrt{1+n^2}} \ln\left(\frac{r_c^2 n^2 + \frac{z_0^2}{\sin^2\beta} - 2B_0^2 - 2\sqrt{(r_c^2 n^2 - B_0^2)\left(\frac{z_0^2}{\sin^2\beta} - B_0^2\right)}}{\left(r_c n - \frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right).$$

Inorder to evaluate I_2 we make the substitution $t = (r_c^2 + x)^{-1}$, then

$$I_2 = \frac{K r_c \sqrt{1+n^2}}{2\sin\beta} \int^{x} \frac{dx}{(r_c^2 + x)\sqrt{x^2 - \left(r_c^2 n^2 + \frac{z_0^2}{\sin^2\beta}\right) x + r_c^2 n^2 \frac{z_0^2}{\sin^2\beta}}}$$

$$= \frac{-K r_c \sqrt{1+n^2}}{2\sin\beta} \int^{t} \frac{dt}{\sqrt{1 - \left(r_c^2 n^2 + \frac{z_0^2}{\sin^2\beta} + 2r_c^2\right) t + r_c^2 (1+n^2)\left(r_c^2 + \frac{z_0^2}{\sin^2\beta}\right) t^2}}$$

$$= \frac{-K}{2\sin\beta\sqrt{r_c^2 + \frac{z_0^2}{\sin^2\beta}}}$$

$$\times \ln\left(2r_c\sqrt{(1+n^2)\left(r_c^2 + \frac{z_0^2}{\sin^2\beta}\right)\left(1 - \left(r_c^2 + \frac{z_0^2}{\sin^2\beta}\right) t\right)} (1 - r_c^2 (1+n^2)\, t)\right.$$

$$\left. + 2r_c^2 (1+n^2)\left(r_c^2 + \frac{z_0^2}{\sin^2\beta}\right) t - r_c^2 n^2 - \frac{z_0^2}{\sin^2\beta} - 2r_c^2\right)\Bigg]^{t}$$

$$= \frac{-K}{2\sin\beta\sqrt{r_c^2 + \frac{z_0^2}{\sin^2\beta}}}$$

$$\times \ln\left(\frac{\frac{2C}{r_c^2 + B_0^2}\left(C + \sqrt{\left(B_0^2 - \frac{z_0^2}{\sin^2\beta}\right)(B_0^2 - r_c^2 n^2)}\right) - r_c^2 n^2 - \frac{z_0^2}{\sin^2\beta} - 2r_c^2}{\left(\frac{\sqrt{z_0^2}}{\sin\beta}\sqrt{1+n^2} + n\sqrt{r_c^2 + \frac{z_0^2}{\sin^2\beta}}\right)^2}\right),$$

where

$$C = r_c\sqrt{(1+n^2)\left(r_c^2 + \frac{z_0^2}{\sin^2\beta}\right)}.$$

Thus, combining I_1 and I_2 we have that

$$I_k(z_0) = \frac{-\text{sign}(z_0)M}{2r_c\delta\sin\beta\sqrt{1+n^2}}\ln\left(\frac{r_c^2n^2 + \frac{z_0^2}{\sin^2\beta} - 2B_0^2 - 2\sqrt{(r_c^2n^2 - B_0^2)\left(\frac{z_0^2}{\sin^2\beta} - B_0^2\right)}}{\left(r_cn - \frac{\sqrt{z_0^2}}{\sin\beta}\right)^2}\right)$$

$$- \frac{\text{sign}(z_0)M}{2\delta\sin\beta\sqrt{r_c^2 + \frac{z_0^2}{\sin^2\beta}}} \qquad (3.12b)$$

$$\times \ln\left(\frac{\frac{2C}{r_c^2+B_0^2}\left(C + \sqrt{\left(B_0^2 - \frac{z_0^2}{\sin^2\beta}\right)(B_0^2 - r_c^2n^2)}\right) - r_c^2n^2 - \frac{z_0^2}{\sin^2\beta} - 2r_c^2}{\left(\frac{\sqrt{z_0^2}}{\sin\beta}\sqrt{1+n^2} + n\sqrt{r_c^2 + \frac{z_0^2}{\sin^2\beta}}\right)^2}\right),$$

where C is defined as above.

Recall that the above expressions are valid for $B_0 < a_m$. If $B_0 \geq a_m$ then all occurences of B_0 should be replaced by a_m, with $a_m = r_cn$ in the latter model. These are the external solutions i.e., the rays of light pass outside the spheroid. The above expressions are considerably simplified in this event.

§3.2 Numerical Techniques

Having found all of the necessary mathematical expressions we proceed as follows. Recall that z and z_0 are the projections of the source position and the image position onto the complex plane, which was defined in fig.1.8 and fig.1.9. By equation (3.1), our functions are of the form $z = f(z_0)$, and so it is the impact parameter i.e., z_0, that is the free variable. We think of z_0 as a position of an image of the source. Thus, we evaluate $z = f(z_0)$ for z_0 in the vicinity of the deflecting mass, and investigate the results.

We consider the complex plane as a fine grid. Each point of the grid is a complex number z_0 i.e., an impact parameter. The point $z_0 = 0 + i0$ corresponds to the ray that goes through the centre of the solid ellipse, which was formed by projecting the spheroid onto the plane. This ray is not deflected, and therefore $z = 0 + i0 = f(0 + i0)$. We evaluate the function $z = f(z_0)$ for all points of the grid which are not "too far away" from the central point 0+i0. A point z_0 would be "too far away" if $z = f(z_0) \approx z_0$ i.e. the ray suffers little or no deviation. In particular, the following diagrams were generated with $|z_0| \leq 20$ units of length. This quantity is determined by the extension and the size of the gravitational lens.

Our aim is to produce diagrams which show the distribution of the multiplicity of images of a source, as a function of the position of the source. Also, we are interested in the positions of the images, their distorsions, amplifications, orientations, and other related properties. The above mentioned grid is formed by two Fortran "DO" loops. Both range from -20 to 20, with the increments being 0.01. Thus there are a total of 16 million possible images, z_0. Some of these, those in the "corners", are disqualified as they have $|z_0| > 20$ units. This leaves about $\pi/4$ of the original images i.e., slightly over 12.5 million. The function $z = f(z_0)$ is evaluated for each of these points, and the results stored in the following manner.

The projected source z is assumed to be a circle of radius 0.25 units, whose centre lies on the intersection of integral or half-integral lines of the grid. Thus the centre of a source may be 1+1i, or 1+0.5i, or 0.5-0.5i, etc. Supposing that $3.26 - 2.03i = f(7.38 - 5.22i)$ to three significant figures, then a source at position $3.5 - 2i$ has an image at $7.38 - 5.22i$. However, if say $3.26 - 2.23i = f(7.38 - 5.53i)$,

then this result is ignored, as there is no source consistent with the above definition which contains the point $3.26 - 2.23i$, see fig.3.1 below.

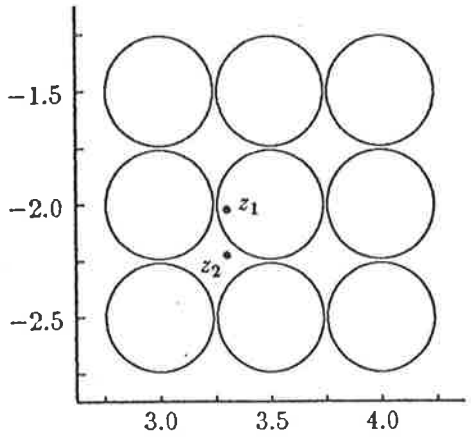

Figure 3.1 Possible sources. Note that whilst the complex number $3.26 - 2.03i$ lies inside a source centred at $3.5 - 2i$, the point $3.26 - 2.23i$ does not. Indeed it does not lie inside any of the sources.

Since the grid is so fine, many adjacent values of z_0 will be mapped onto the same z. These yield the shape of the image. Since intensity is given by the ratio of the area of the image to the area of the source, Etherington (1933), the relative intensity of the image can be immediately found. These adjacent points are rounded off in the same way as were the sources, and the results are stored in the following way.

We are to present diagrams which show the multiplicity of images as a function of the position of the projected source with respect to the deflecting mass. As the projected source always lies between the impact parameter and the centre of the mass, the dimensions of the region of interest do not need to be as large as those of the impact parameters. In particular, most sources which have more than one image in our examples have their real and imaginary parts between -10 and 10. So there are 41 possible sources along both the vertical and the horizontal axes.

Most of the information is held in the $41 \times 41 \times 100$ array of complex numbers called IMAGE. Thus, the first two indices of IMAGE correspond to the position of the source, and the third index corresponds to the number of the image that was mapped on to the source. For example, IMAGE(1,2,3) is the third image that

was mapped onto the source at -10-9.5i. The actual mapping of z onto K and L of IMAGE(K,L,M) is K=INT(REAL(z)*2)+21 and L=INT(AIMAG(z)*2)+21.

Now, if the third index of the above array is M, it does not mean that the source, corresponding to K and L, has M images, where images is used in the astronomical sense. If, astronomically speaking, a source has only one magnified image, then M may be a 100 or more. All this array tells us is which values z_0 were mapped onto the source z, it does not differentiate between images in the astronomical sense. This is the purpose of another $41 \times 41 \times 10$ array of complex numbers called IMAGE1. This array determines where one image ends and another begins. This is simple to do when the real images are far apart in the sky, but becomes difficult when the images are close together e.g., when two images coalesce to form one. Thus, we define IMAGE(K,L,M) to be a new image if M=1, or if $|IMAGE(K,L,M) - IMAGE(K,L,N)| > 0.75 \quad \forall \quad N \in [1, M-1]$. The results are stored in IMAGE1(K,L,M), where M is now the pointer to the M^{th} image, the value of which is the arithmetic average of all the points that make up the image. This array contains all the information regarding the multiplicity of images as a function of the position of the source. As, for example, if the image pointer IMAGECT1(K,L) of IMAGE1(K,L,IMAGECT1(K,L)) is 5, then the source whose position is mapped onto K and L has 5 images.

This formalism is very convenient, as it is very easy to make thorough investigations of the gravitational lens effect due to a known spheroid after just one run of the program. The program allows the user to interactively examine the number of images any source has, as well as their positions, orientations and intensities.

In the following section we investigate the gravitational lens effect of spheroidal deflectors for the three models discussed in section 1 of this chapter.

§3.3 Numerical Results

For our first example of a gravitational lens, we choose a typical E7 elliptical galaxy (Page 1962). We take its mass to be 5×10^{11} $M_\odot$, axis ratio $\sqrt{1-e^2}$ to be 3/10, and its semi-major axis to be 8 kpc, see fig.3.3. For the King model we take r_c to be 2 kpc, and hence n is 4. We assume that the source and the galaxy are situated such that $D = 500$ Mpc, and that the galaxy is seen edge on i.e., $\gamma = \pi/2$.

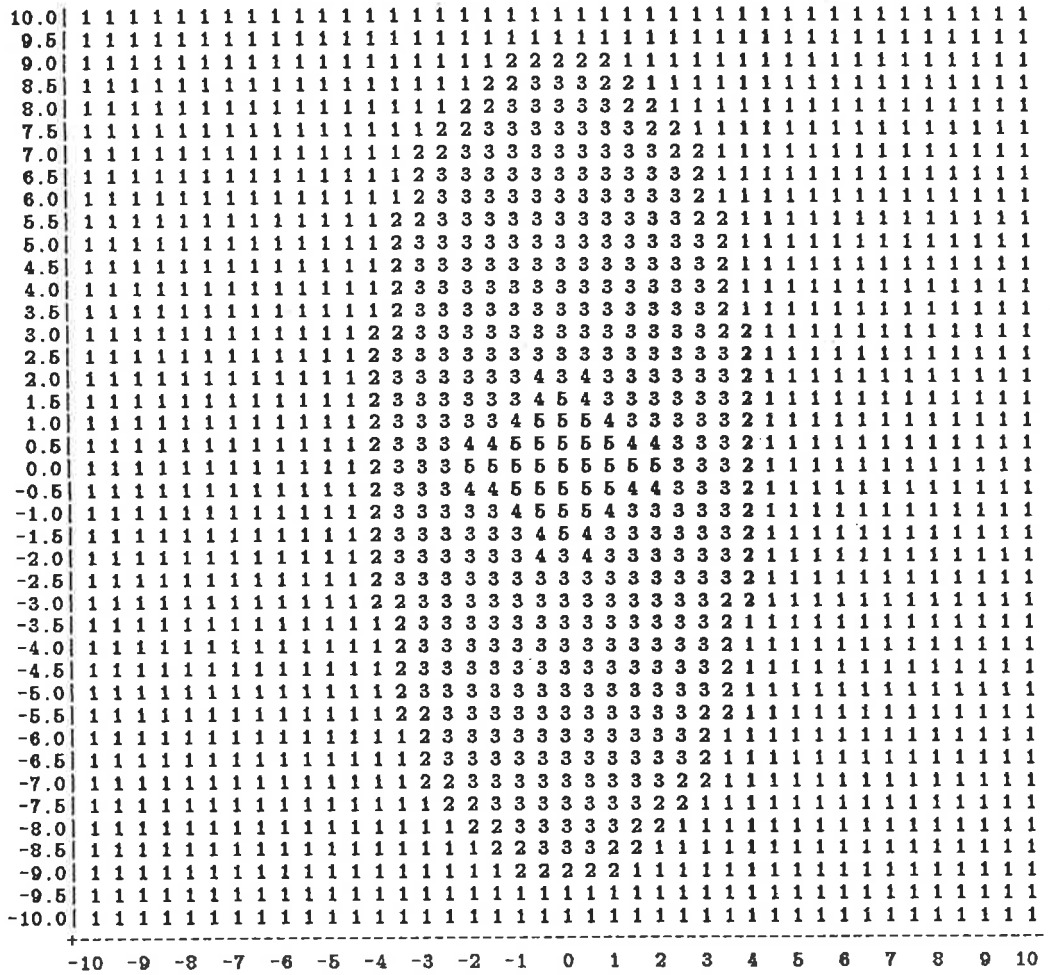

Figure 3.2. Multiplicity of images as a function of the position of the source, for a typical E7 elliptical galaxy modelled by the Schmidt density distribution. The above coordinate plane is a complex plane, with each entry corresponding to the centre of a source whose radius is 0.25 units of length. The units of length are determined by the spheroid, and hence are kilo-parsecs. The origin corresponds to the centre of the projected spheroid, whose orientation can be seen in fig.3.3. A clearer presentation of this diagram is fig.3.4.

For the Schmidt model, we get the above table of multiplicity of images as a function of the position of the source. Firstly we note that the results reflect the symmetry of the projected spheroid, see fig.3.3.

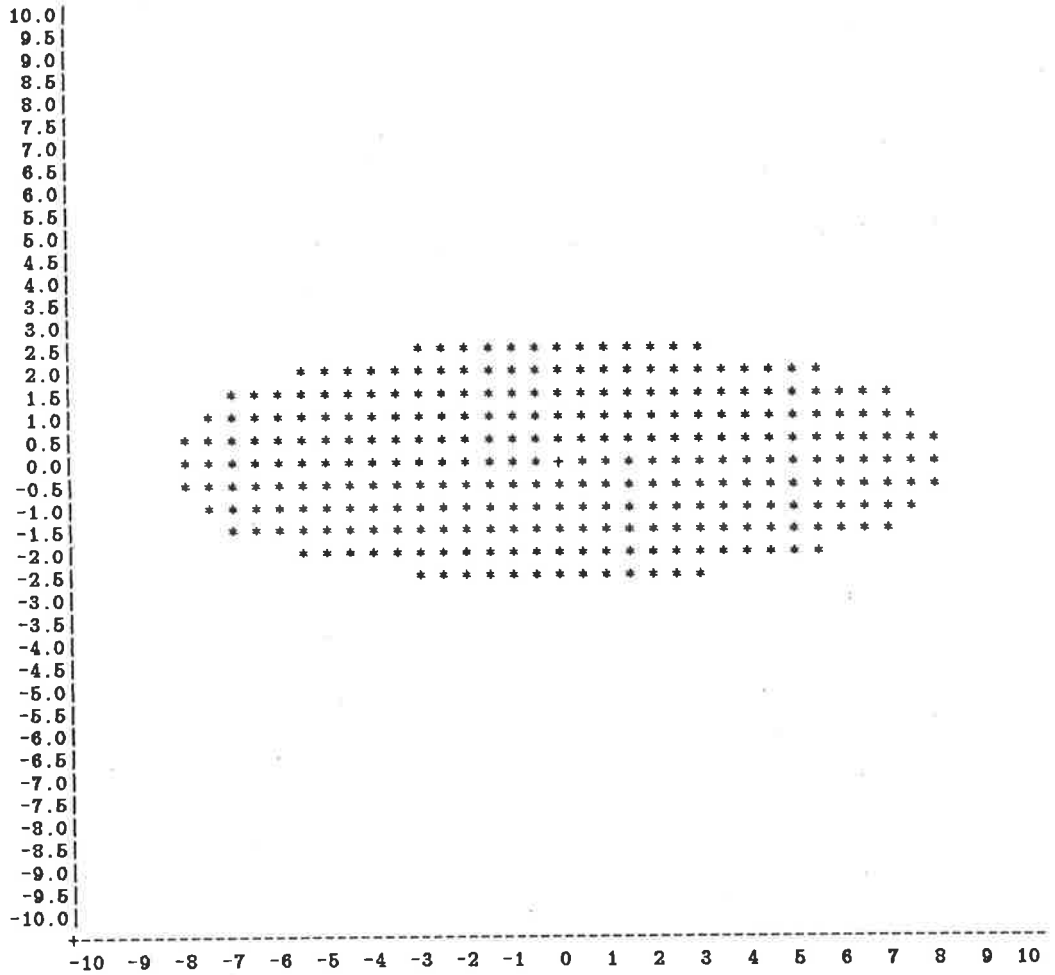

Figure 3.3. The solid ellipse formed by projecting the spheroid. This table may be superimposed on that of fig.3.2. The ellipse was generated by finding which impact parameters z_0 have $B_0 \leq a_m$. In this case of a typical E7 elliptical galaxy, $a_m = 8$ kpc, as can be seen above. As we assumed that the spheroid is seen edge on i.e., $\gamma = \pi/2$, the axis ratio of approximately 3/10 is evident. For z_0 outside the ellipse i.e., for $B_0 > a_m$, the considerably simpler "external" solution is used to find the source position z. The "external" solution corresponds to setting $B_0 = a_m$ in the evaluations of $I(z_0)$.

By Burke (1981), we know that a transparent lens may only have sources produce an odd number of images. Thus, it is at first surprising to find a thin elliptic region of sources which have two images, and small regions of sources which have four images. This however, is not a contradiction, as it is in these regions

that two images coalesce to form one. As we had to choose some finite resolution, the same resolution as that of the sources, two images were interpreted as one in these regions. As these regions are not very important, as well as for clarity of presentation, we represent them by blanks. Thus, fig.3.2 becomes

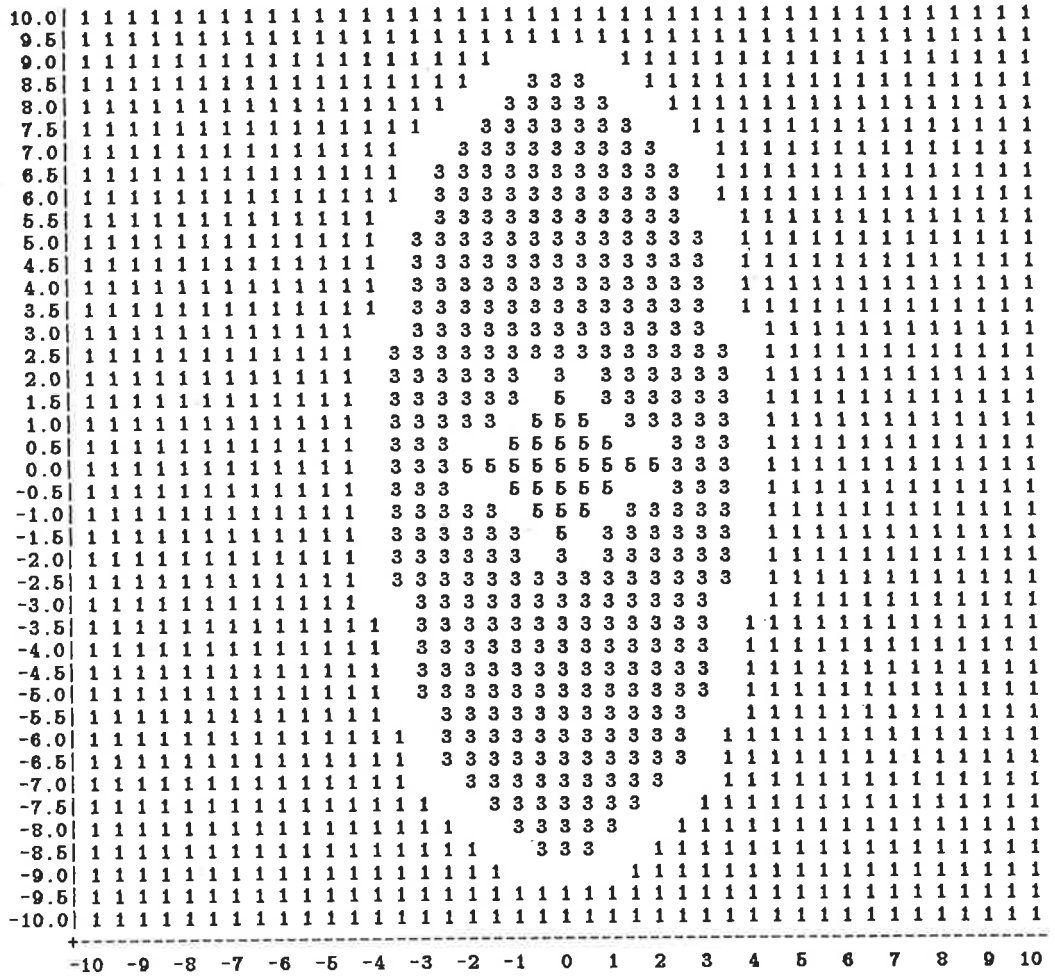

Figure 3.4. A clearer presentation of fig.3.2. This is just a reproduction of fig.3.2, but with the even numbers blanked out for clarity. It is in the blank regions where two images coalesce to form one.

So far we have only talked of the multiplicity of images, but not their positions nor their shapes. Sources near the edges of the above diagrams have images outside the dimensions of the above diagrams as the projected position of the source is always between the image and the centre of the lens. Recall that the source positions z of fig.3.4 were generated by $z_0 \leq 20$. However, as these projected sources are far from the centre of the lens, their images are not as interesting as

those due to z being near the origin. We first examine the direct results for the source which happens to be nearly perfectly aligned with the centre of the lens and the observer. By direct results we mean, that we present all impact parameters z_0 that were mapped onto the specified source, see fig.3.5 below.

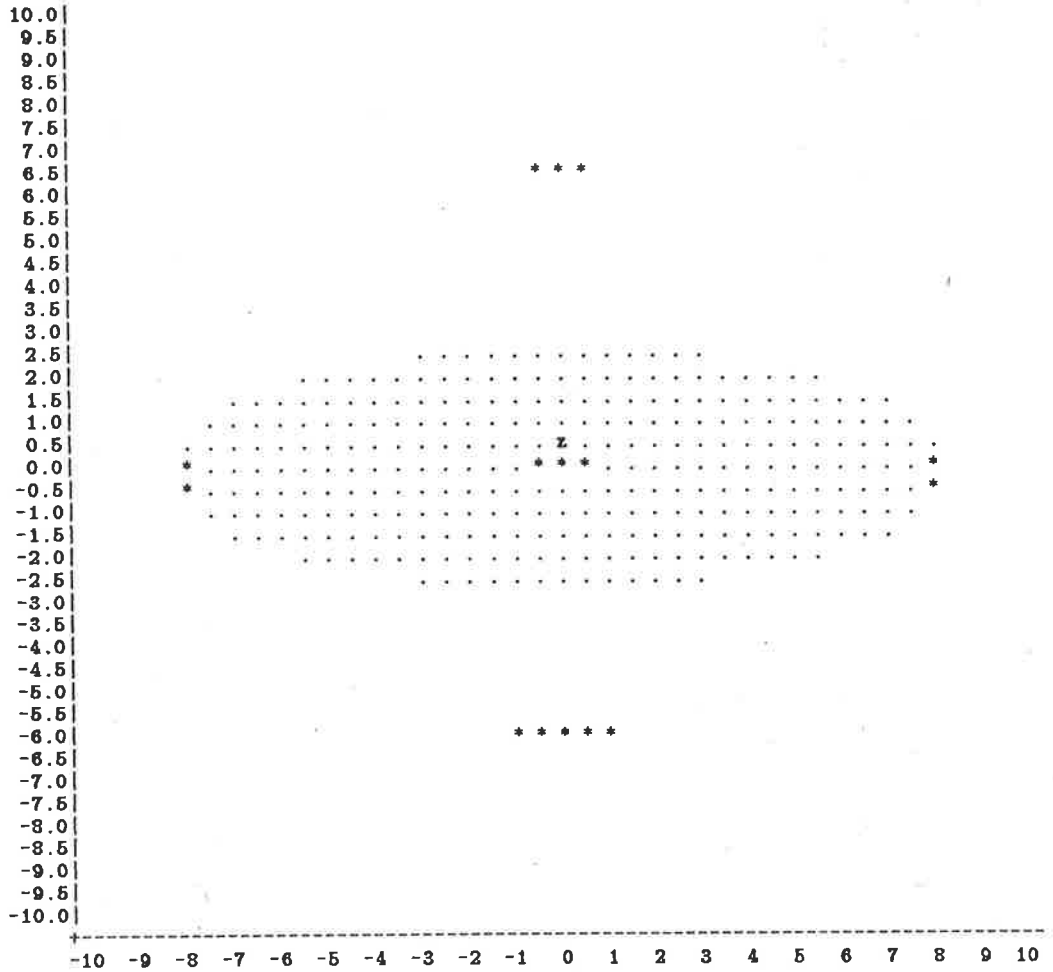

Figure 3.5. Impact parameters for $z = 0+0.5i$. The stars are the impact parameters z_0 which are mapped onto $z = 0+0.5i$ by equation (3.1), for the constant density model. Apart from showing the positions and the orientations of the images, this diagram also shows their approximate amplifications. The dots show the projected spheroid.

We can see from the diagram above that there are indeed five real images associated with the source $z = 0 + 0.5i$. The bottom image has the largest cross sectional area and therefore has the greatest magnification. The resolution of the output does not allow for exact results of magnification, however, once the position of the image is known, it is trivial to increase the resolution in that region to obtain greater detail if necessary.

Recall once again our motivation for producing the diagrams such as fig.3.4. It was the fact, that the radio blobs associated with the double quasar did not have a second image, that led the radio astronomers to reject the gravitational lens hypothesis. However, there is an angular separation in the position of the quasar and its associated radio sources. We see from fig.3.4 that depending on the orientation of the galaxy, a small angular separation between two sources, may change the multiplicity of images from five to one. The size of the angular separation required to produce a reduction in the number of images depends on the model of the galaxy, see corresponding results for the King and the constant density models.

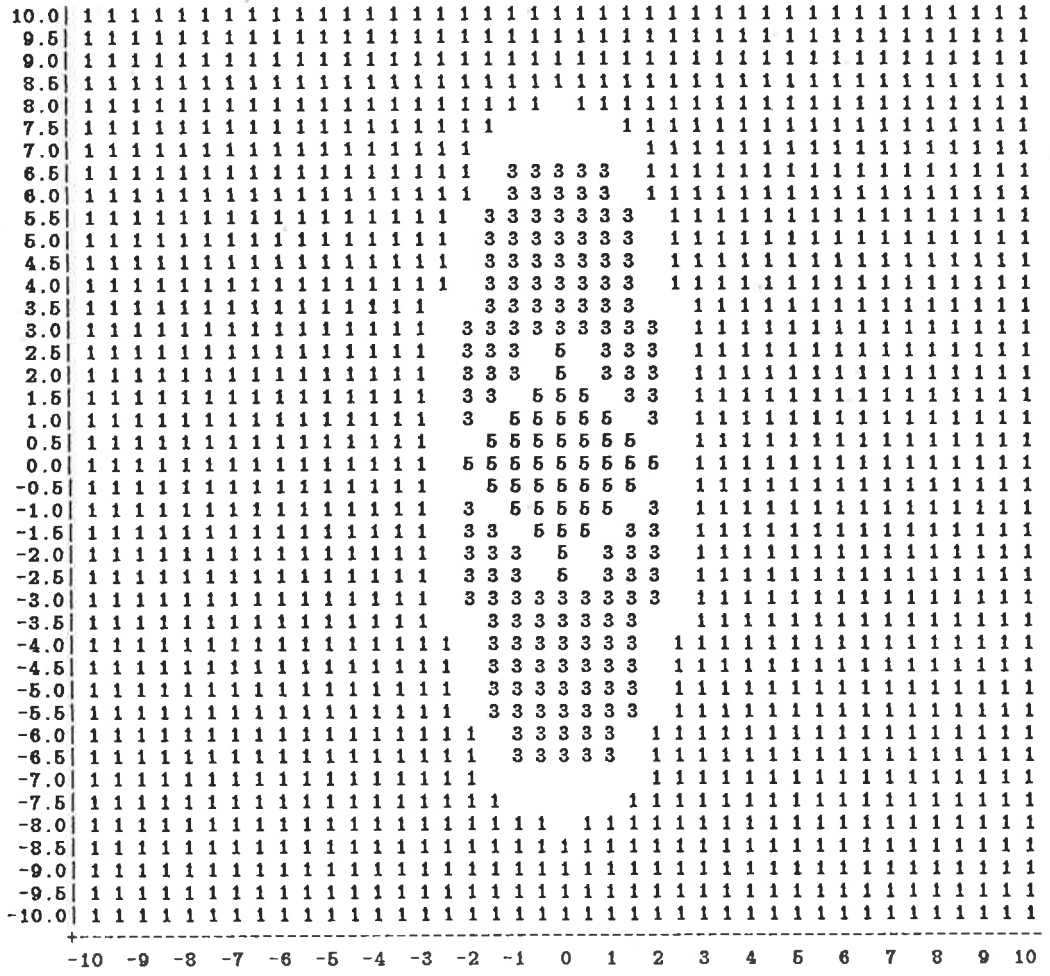

Figure 3.6. Multiplicities of images as a function of the position of the source for the constant density model.

There is more useful information that can be immediately obtained. As we

know the position of every image for each source of fig.3.4, we can study the path that images take as the projected source moves along any path across the table. Thus, we are able to perform a dynamic study of the variation of images due to a real source passing near the line of sight of the lensing mass. This we do at a later stage.

At this stage we can compare fig.3.4, results for the Schmidt model, with fig.3.6, results for the constant density model, and fig.3.7, due to the King model.

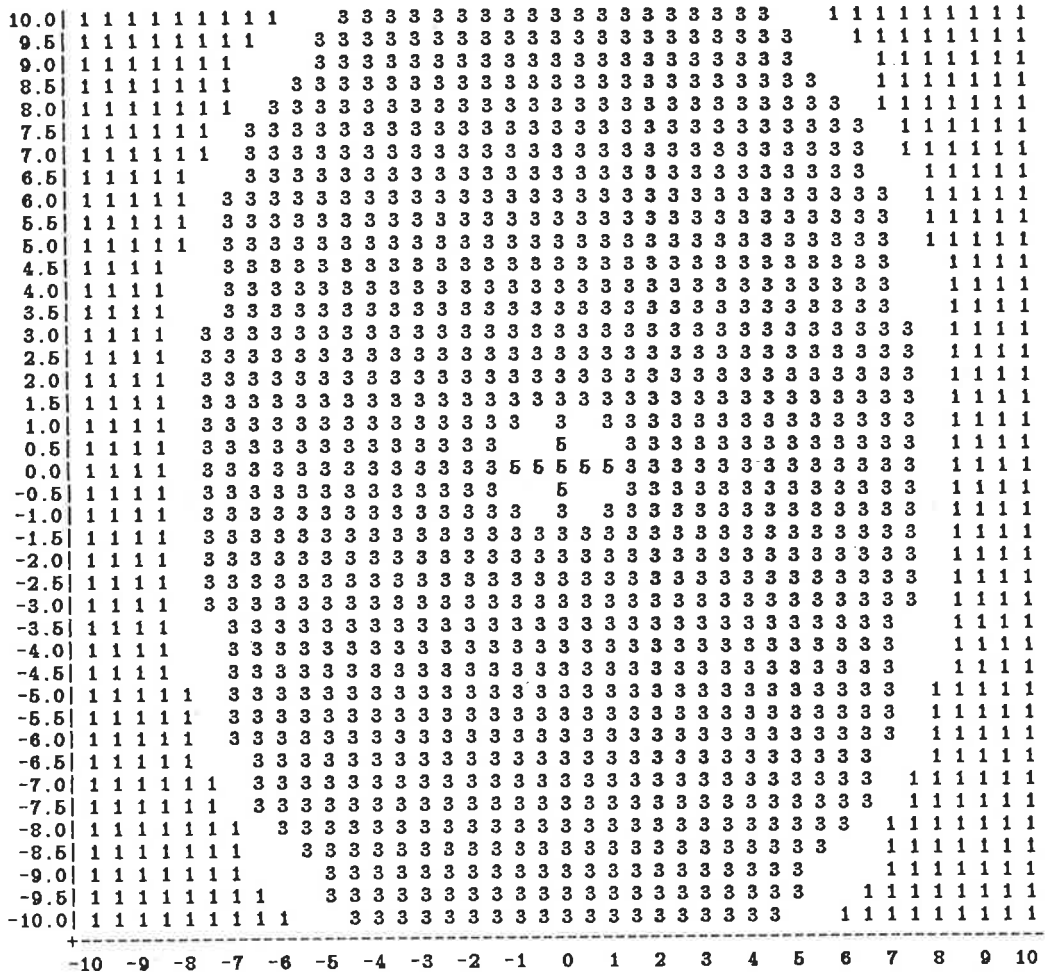

Figure 3.7 Multiplicities of images as a function of the position of the source for the King density model.

In general, the results are similar. However, there are notable differences. We see that the size of the region where a source generates five images is largest for the constant density model, and smallest for the King density model. However, the size of the region in which a source generates three images is largest in the

King model and least in the constant density model. The latter observation also applies to the situation where γ is zero i.e., when the spheroid appers spherical, see figures 3.8, 3.10 and 3.11.

Now, as by changing the density model we do not change the symmetry of the lens, the results are similar i.e., changing the model merely alters the shape of the corresponding regions of image multiplicities. Thus, the choice of the density function determines the size of the corresponding regions of image multiplicities.

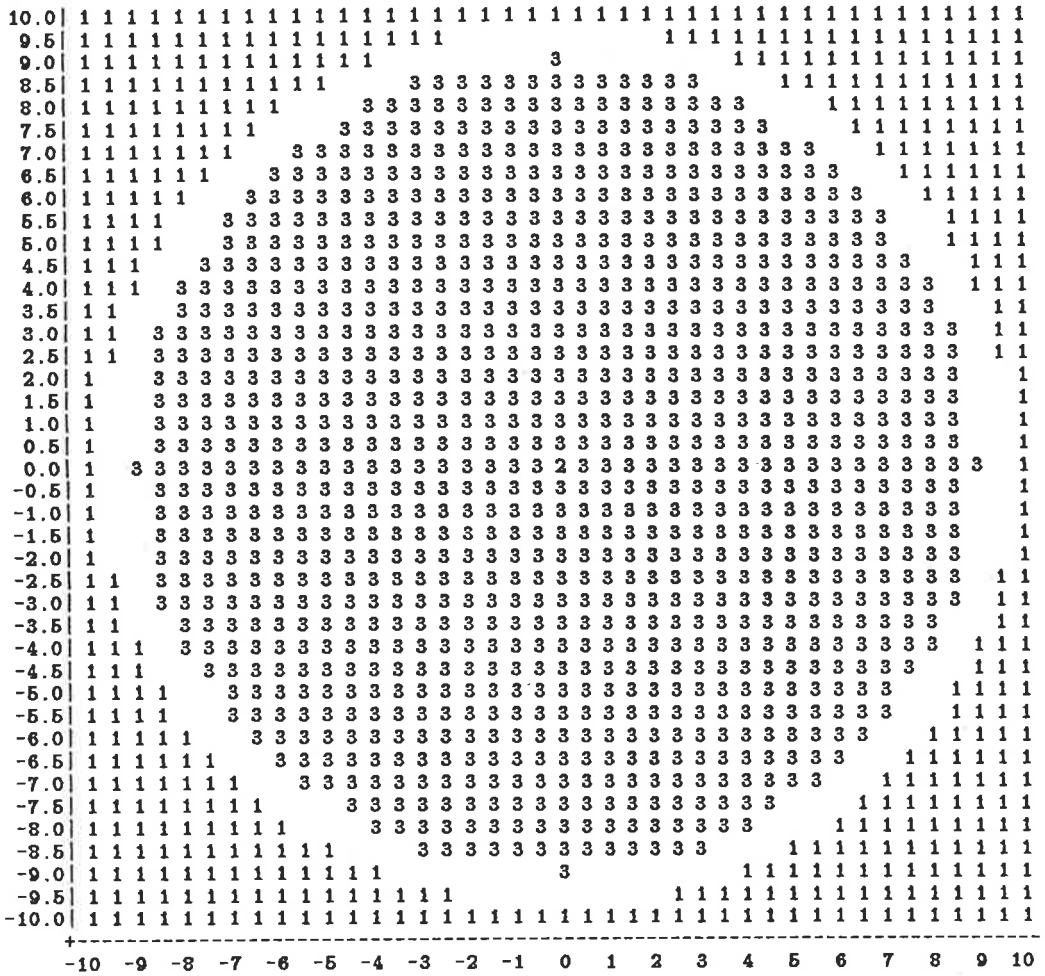

Figure 3.8 Multiplicities of images as a function of the position of the source for the Schmidt density model with γ being zero. Note that the central source i.e., the one that is perfectly aligned with the centre of the spheroid and the observer, has two images. This is due to the fact that as the source comes into such perfect alignment, two of its three images coalesce to form a ring, see fig.3.9.

If the impact parameters are outside the solid ellipse formed by projecting the spheroid, then the external solutions are used, this gives similar results for

all three models. This is because as the ray is external to the lens, it is less sensitive to the mass distribution. It is when the impact parameters fall inside the solid ellipse that the internal solutions are used, which are more sensitive to the mass distribution. This is most evident as one traces the images for each model for a particular source path. When a source is near the boundary of the above diagrams, not surprisingly, the image it generates is the same in all models. In general, the "external" solutions give very similar results in all three models, whilst the "internal" solutions are critically dependant on the mass distribution.

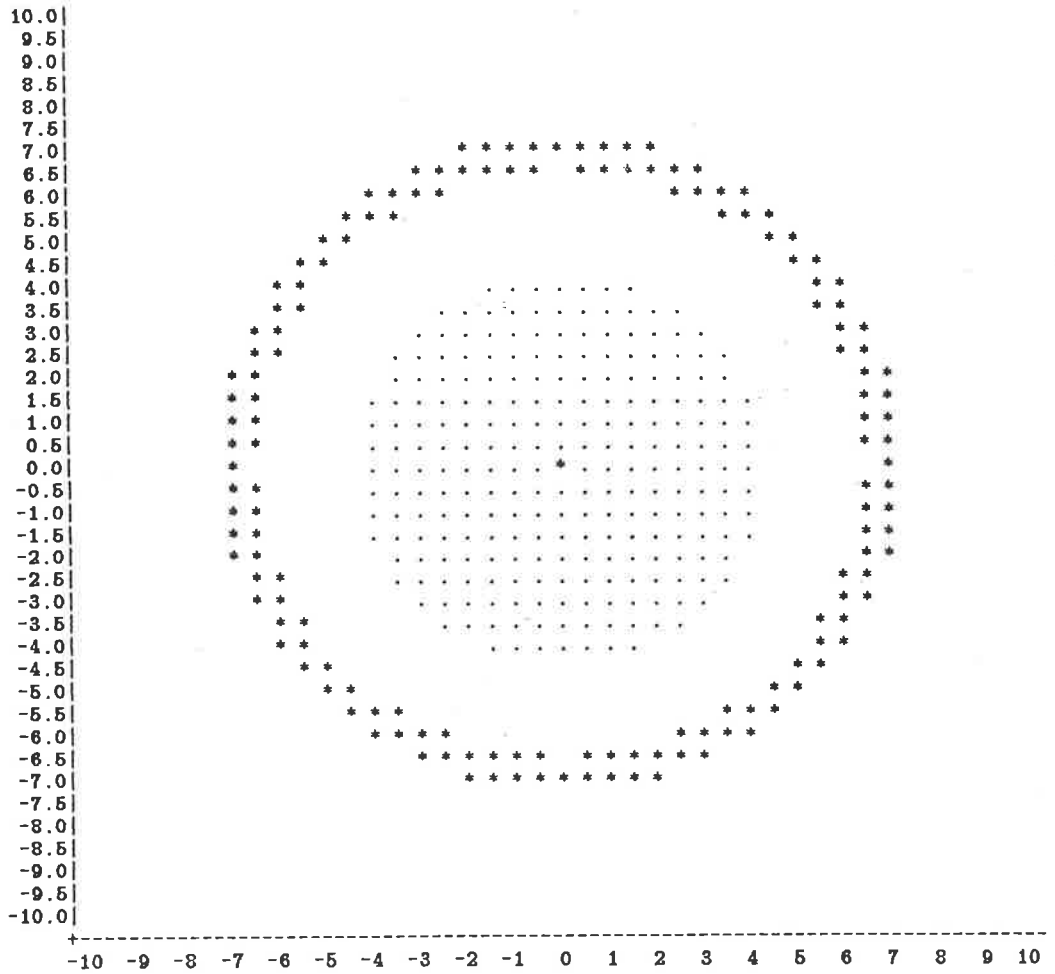

Figure 3.9 Projected spheroid and the ring image. Note that the spheroid is either a sphere or is seen along its axis. The ring image is due to the projected source being at $0 + 0i$, thus, it is perfectly aligned with the centre of the spheroid and the observer. A projected source at $0 + 0i$, by symmetry, always has an image at $0 + 0i$. The ring was generated by the constant density model, however, the other models give a similar result.

Now we compare the results for $\gamma = 0$. This means that the spheroid is seen along its axis of symmetry. We saw in our equations that $\gamma = 0$ is equivalent to $e = 0$. Thus a spheroid of arbitrary eccentricity seen along its axis generates the same results as a spheroid which is a sphere, with all other parameters being equal. This is due to the assumption we made at the start, namely that the extension of the spheroid is negligible when compared to the total distance travelled by the ray.

We left all parameters the same as for the spheroidal mass, except that we halved a_m for ease of presentation. For the King model we halved r_c to become 1 kpc hence, n is unaltered. The resultant projected spheroid can be seen above in fig.3.9.

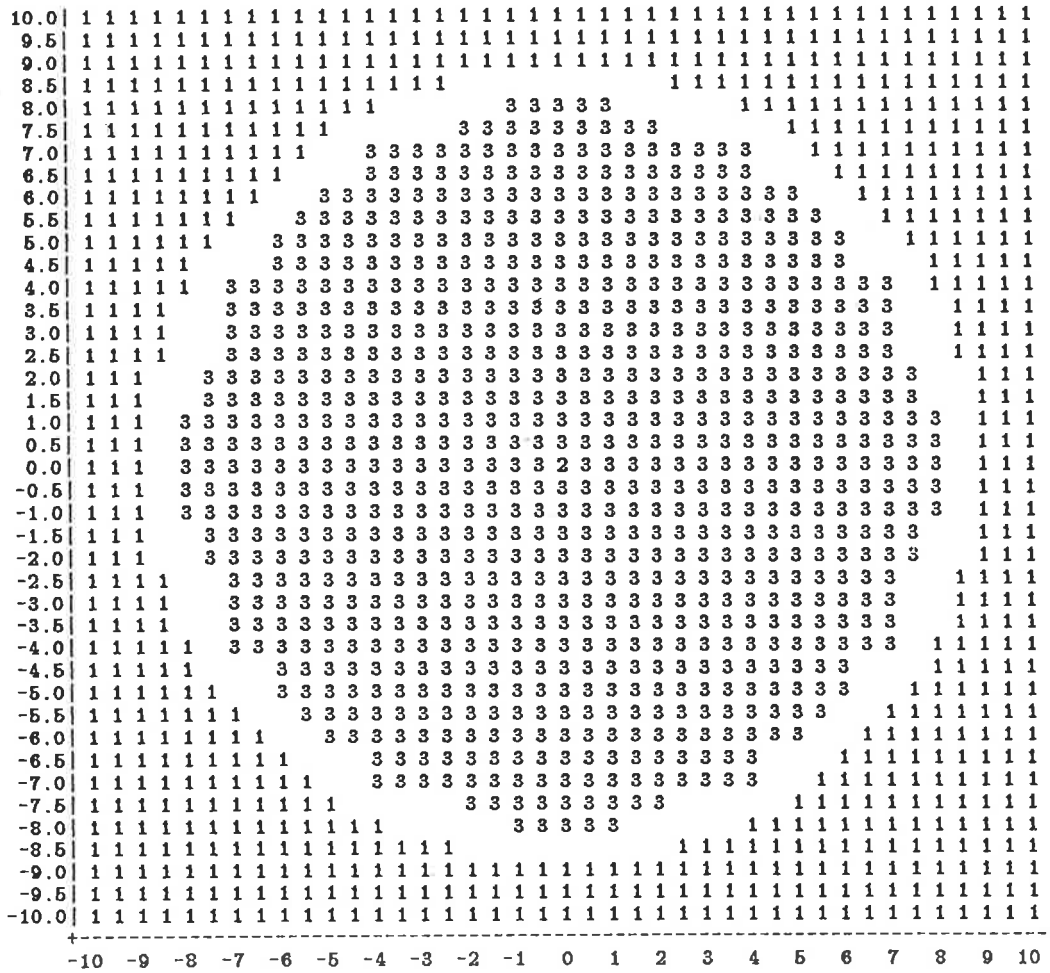

Figure 3.10 Multiplicities of images as a function of the position of the source for the constant density model with γ being zero.

We see that image multiplicities are extremely similar for the Schmidt density model and the constant density model, figures 3.8 and 3.9. The King density model, however, differs from the other two considerably, fig.3.11.

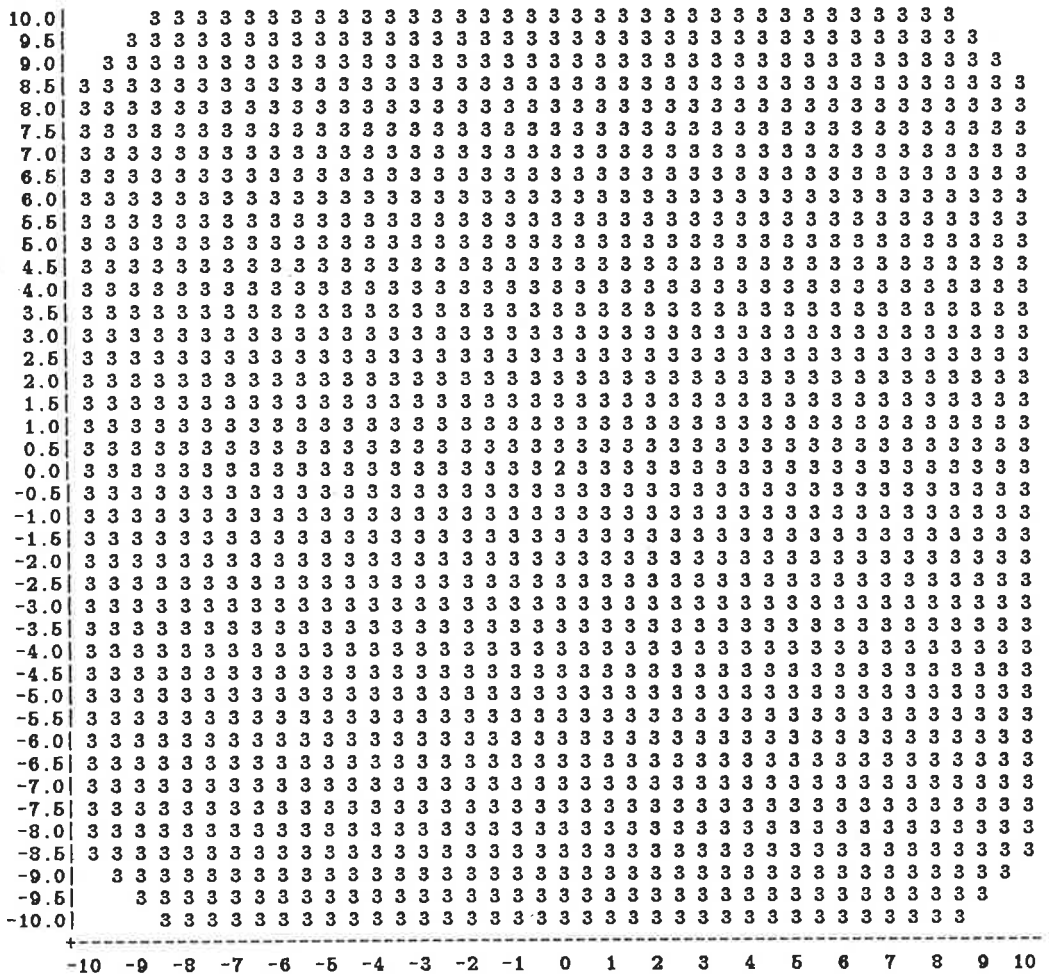

Figure 3.11 Multiplicities of images as a function of the position of the source for the King density model with γ being zero. The corners are blank as the number of images a source generates there is two.

In the King model the circular region inside which a source generates three images has a radius of one and a half times the radius of the corresponding regions in the other models. We wonder at the cause of this and consider if the value of r_c has an effect on the size of the three image region. Thus, we set $r_c = 2$, and hence $n = 2$ also, as $a_m = 4 = nr_c$.

We see from fig.3.12 that the value of r_c plays a role in determining the size of the three image region.

The best comparison between models is to trace the images of a source, as its projection follows a path across the projected spheroid. We do so in figures 3.13 to 3.18, where we compare the King model with the Schmidt model for the original oblate spheroid.

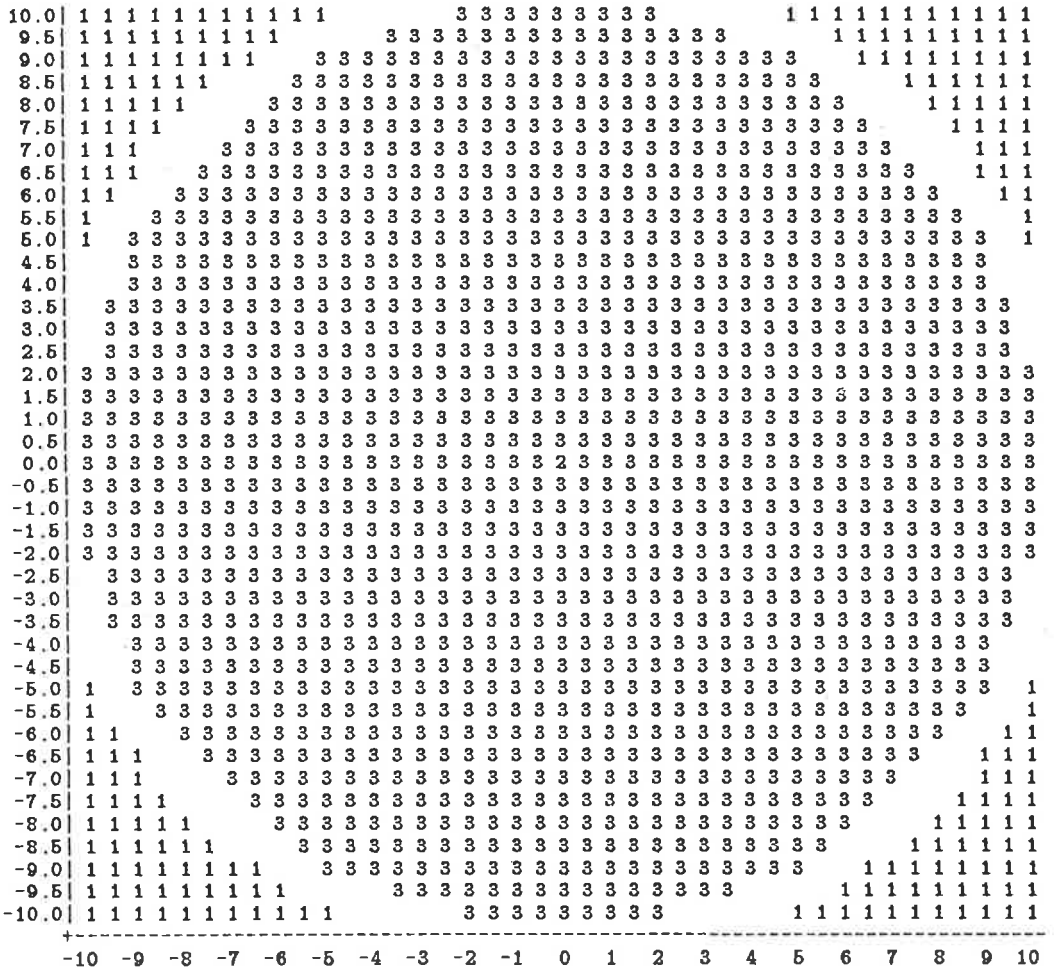

Figure 3.12 Multiplicities of images as a function of the position of the source for the King density model with γ being zero, but with $n = r_c = 2$. The size of the three image region is reduced by the larger value of r_c.

We suppose that the projected source takes the path parallel to the real axis, with IM(z)=0.5. At large distances from the projected spheroid the two models generate the same single image. This image is external i.e., it lies outside the projected spheroid, and it stays external as the source moves across the projected spheroid. Thus, there is always at least one image in one model which has a corresponding image in another model. The odd image theorem states that a

transparent gravitational lens causes an odd number of images of a single source. Therefore, new images are created in pairs and we may also compare the corresponding pairs of images in the two models. We move the projected source from large negative $\mathrm{RE}(z)$ to $\mathrm{RE}(z) = 0$, and take snapshots of the images when $z = -4+0.5i$, $z = -1+0.5i$ and $z = 0+0.5i$. The positions of images for positive $\mathrm{RE}(z)$ are mirror images of those for negative $\mathrm{RE}(z)$.

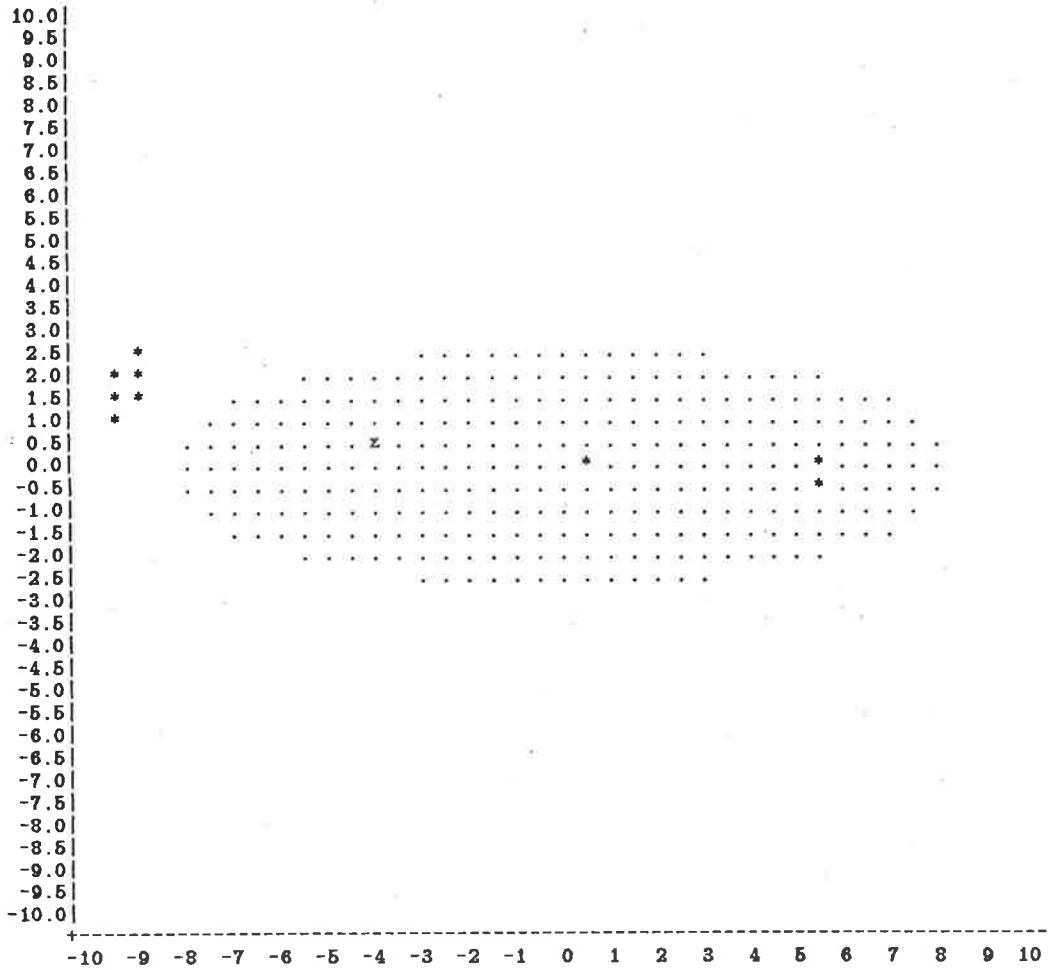

Figure 3.13 Impact parameters for $z = -4 + 0.5i$, **using the King density model.**

We see from fig.3.4, image multiplicities for the Schmidt model, that as the projected source moves along $\mathrm{IM}(z) = 0.5$ from left to right, the last position where it generates only one image is $z = -4.5 + 0.5i$. Similarly, in the King model, fig.3.7, the last source position that generates only one image along this path is $z = -8.5 + 0.5i$. However, the external images of the two models differ, up to the displayed resolution, for the first time when z reaches $-4.5 + 0.5i$.

Our first snapshot is at $z = -4 + 0.5i$, fig.3.13 for the King model and fig.3.14 for the Schmidt model. We see that the external images are very similar in position, shape and amplification, with the most noticeable difference being that the Schmidt image is a little higher. According to fig.3.4, the projected source $z = -4 + 0.5i$, generates two images. This is not a contradiction to the odd image theorem as we can see from fig.3.14. The internal image is really two images that are so close together, that they cannot be resolved using our resolution of half a kiloparsec.

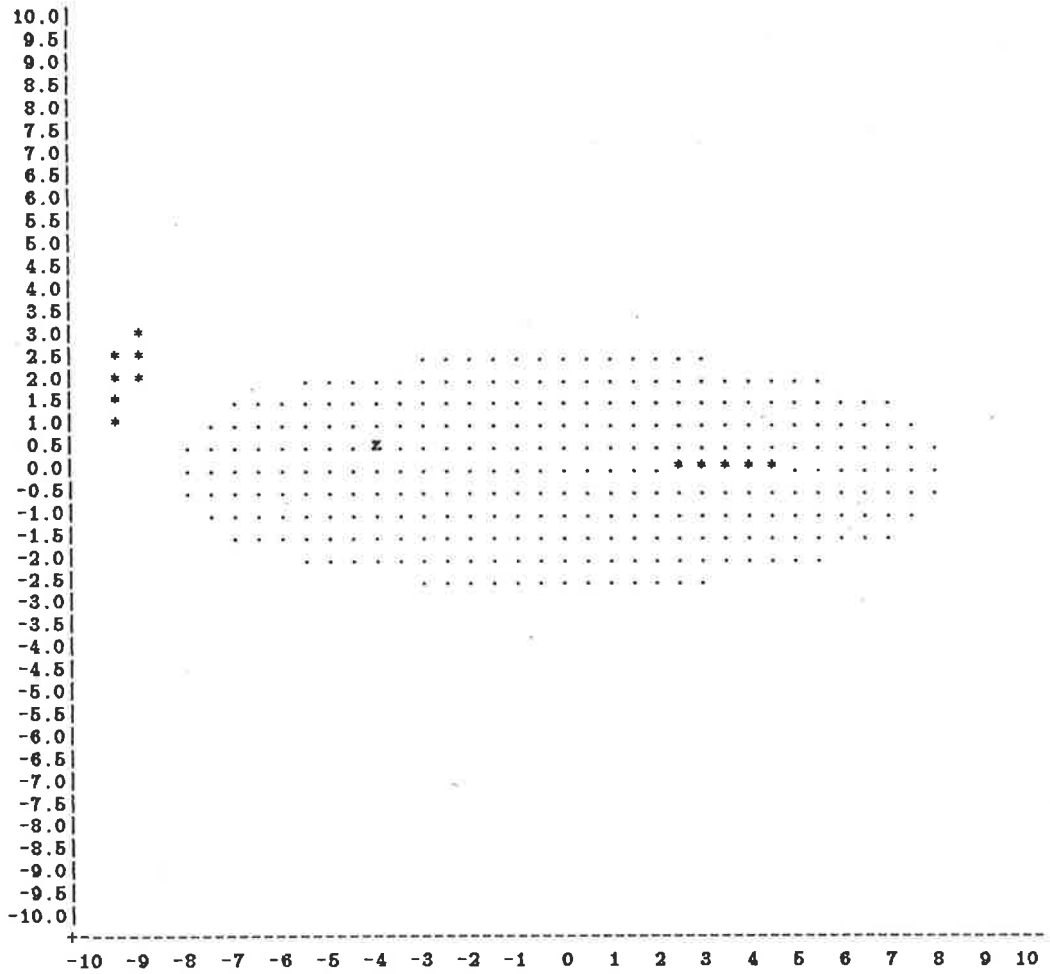

Figure 3.14 Impact parameters for $z = -4 + 0.5i$, using the Schmidt density model.

We see that the internal images differ markedly for the two models. The King model was able to generate internal images, when the projected source was four kiloparsecs further away.

As the projected source moves closer to the centre, the internal images move

further apart, and the external image moves in a clockwise direction. Our next snapshot is at $z = -1 + 0.5i$, fig.3.15 for the King model and fig.3.16 for the Schmidt model. We see that a new pair of images has come into existence. This pair has large surface area, hence large magnification, yet a mere half a kiloparsec earlier it did not exist. This is an indication of large instability i.e., a small distance travelled by the source has a large effect on this pair of images. The other pair of images seems very stable at these values of z, and the original, external image is fairly stable for all values of z.

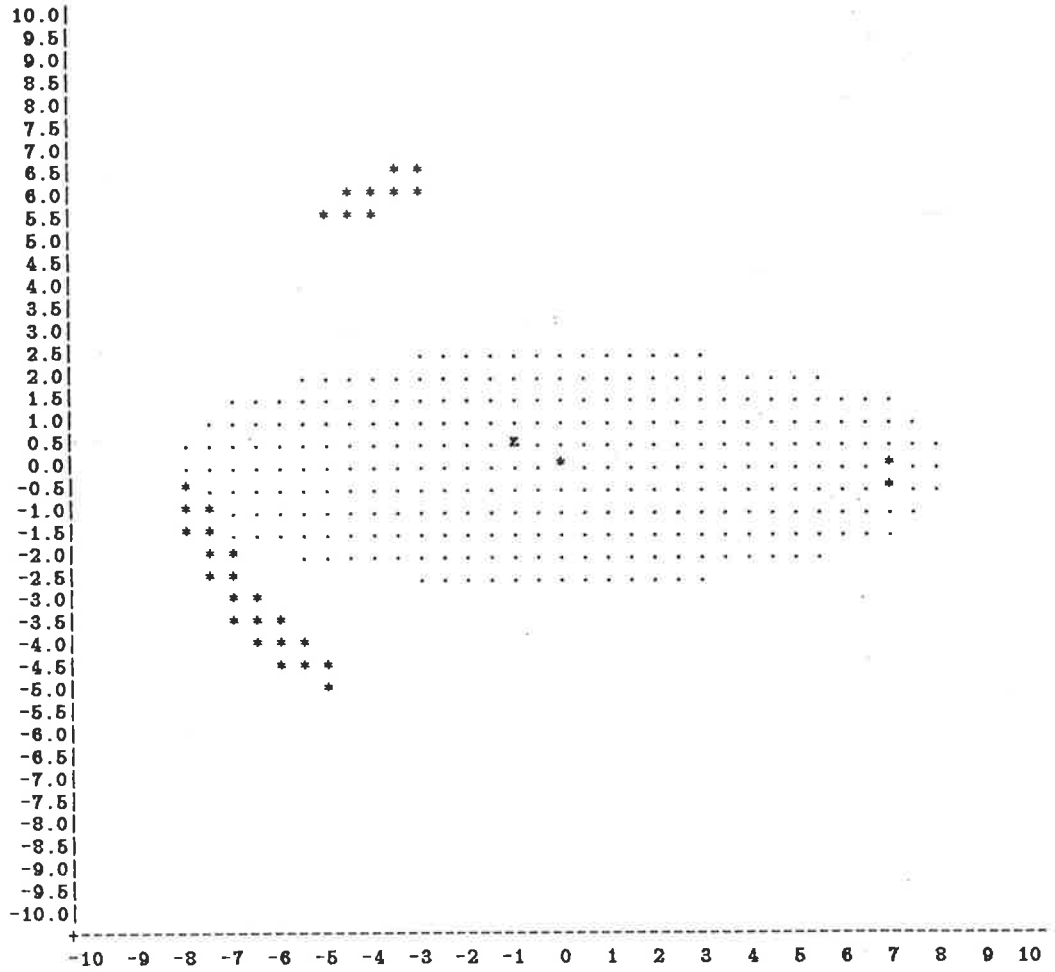

Figure 3.15 Impact parameters for $z = -1 + 0.5i$, using the King density model.

Once again we see that the images in the newly created pair of images are so close together, that they appear to be a single image. It is interesting to note that as the projected source moves towards the centre, it is the King model which is the first to generate three images of the source, whilst it is the Schmidt model

which is the first to generate five images of the source. Comparing corresponding images we see that the original, external images are similar in position, shape and amplification. Their mean position varies by less than one kiloparsec. The amplifications are roughly nine and seven for the top external images of the King and Schmidt models respectively, where the amplification is the ratio of image area to source area, which can be estimated by the number of points that make up the image.

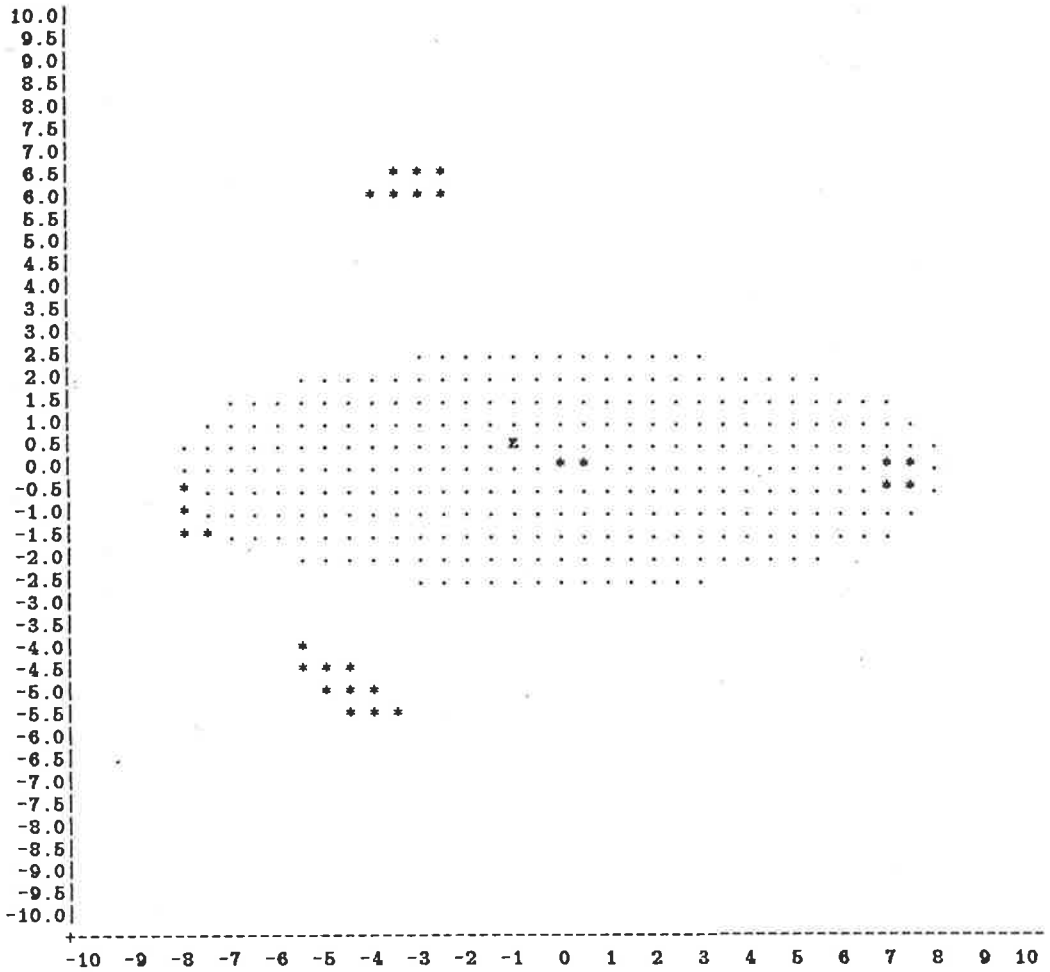

Figure 3.16 Impact parameters for $z = -1 + 0.5i$, **using the Schmidt density model.**

Note that this procedure becomes less reliable if there are few points that make up the image, however, once the general position is known, it is trivial to increase the resolution of that particular region to enable more detailed study. The cross sectional area of the source is equivalent to one unit.

The last snapshot we take is at $z = 0 + 0.5i$. The results are in figures 3.17

and 3.18 due to the King and Schmidt models respectively. These may also be compared with fig.3.5 which is a snapshot of the images when $z = 0 + 0.5i$ for the constant density model. The first thing to note is the remarkable similarity in the corresponding positions of all five images in each model. The main difference appears to be in the amplifications of the two external images. These have greatest amplification in the King model where they have about twice the magnification of the corresponding images in the other two models.

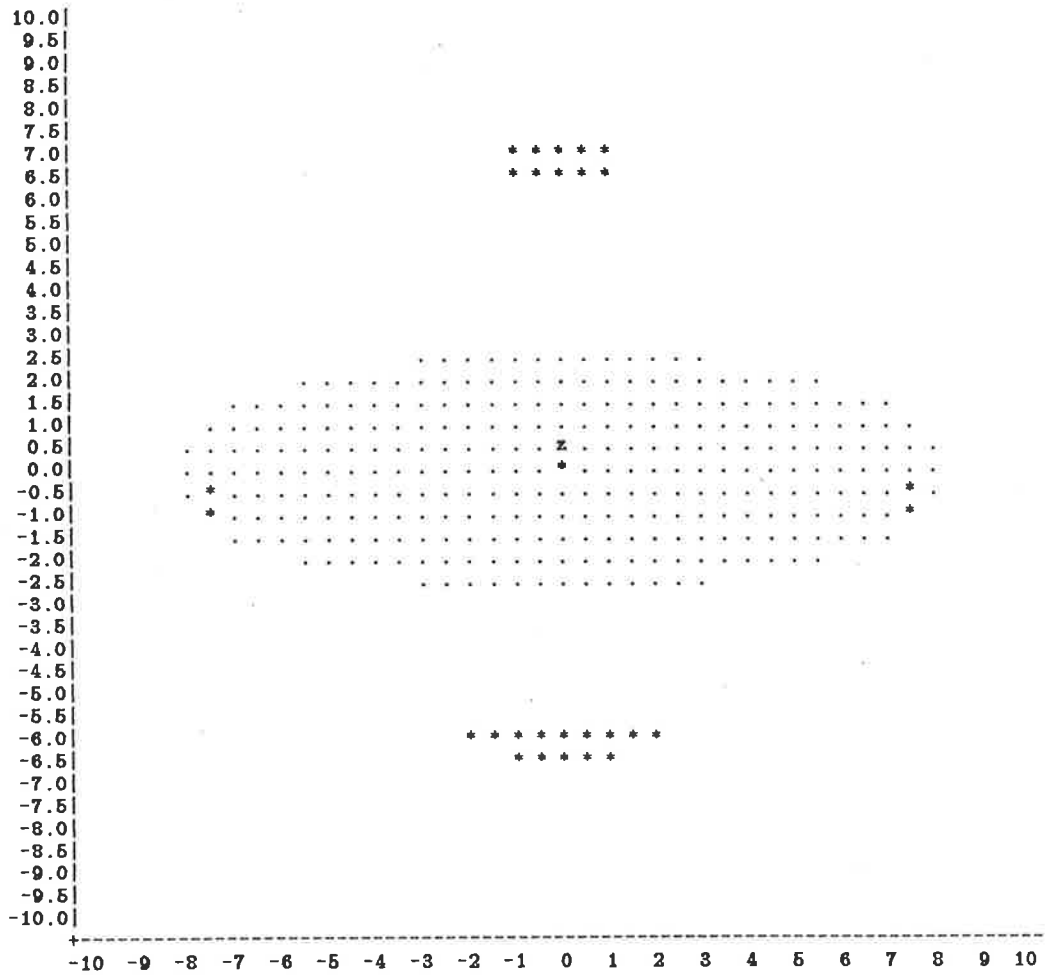

Figure 3.17 Impact parameters for $z = 0 + 0.5i$, using the King density model.

It is difficult to make wide-spread generalizations from the results that we have obtained in this section. Essentially, we have only worked with one spheroid, whose properties we have examined in detail. We have found that the choice of density model does not drastically alter the spheroid's behaviour as a gravitational lens. In particular, the positions of the brightest, external images hardly vary for

the three models. However, depending on the value of r_c, the King model has a larger region where projected sources generate more than one image. Throughout, every diagram has had the same symmetry as the spheroid that generated them, namely, they all are symmetric about the real and the imaginary axis. This is why we did not choose a path for the source to be on either axis, as then redundant information would be displayed.

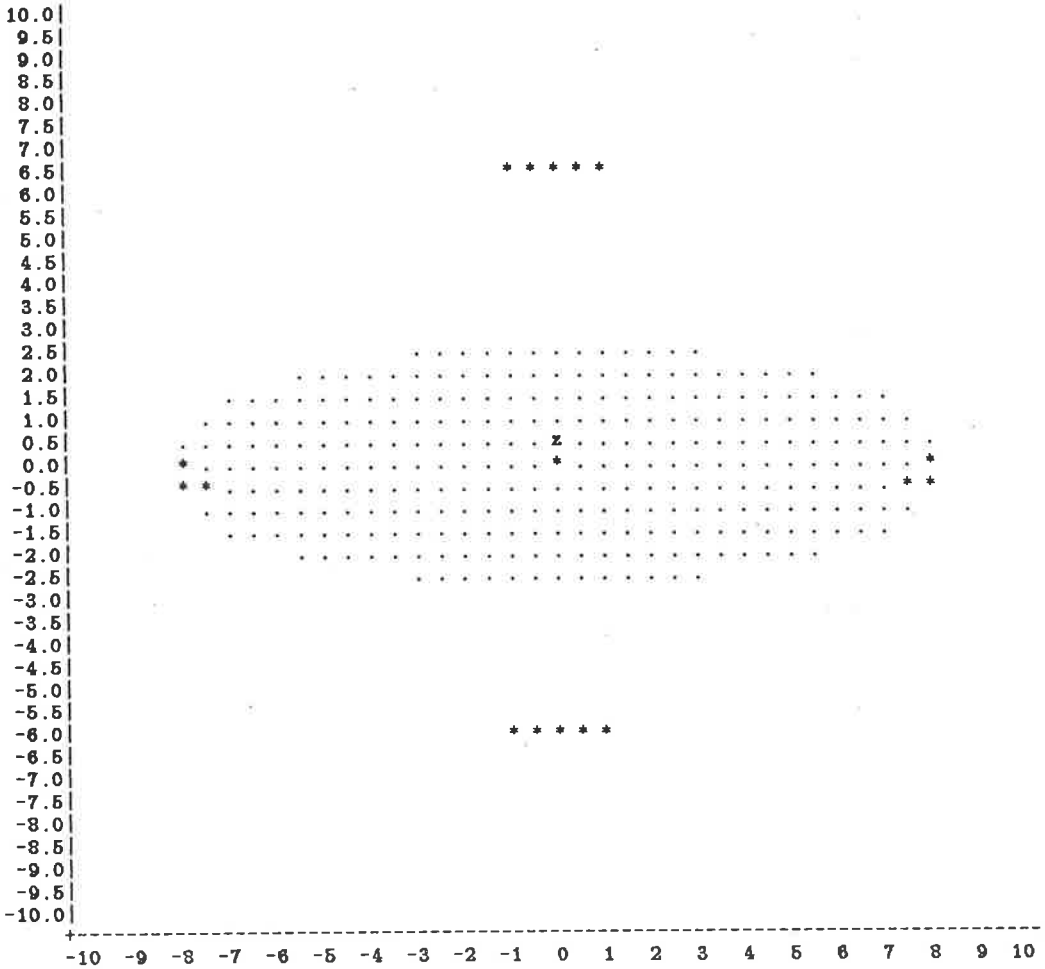

Figure 3.18 Impact parameters for $z = 0 + 0.5i$, using the Schmidt density model.

There is one more feature of the equations we derived that we would like to discuss. We found the position of the source as a function of the impact parameter by summing the contributions due to the individual masses that made up the gravitational lens. So far we have assumed that these masses constituted an arbitrary but single spheroid. However, it is equally valid to think of the gravitational lens as being best represented as a combination of spheroids, with each spheroid

having its own density distribution. As an example, we may model a spiral galaxy by two concentric spheroids: the first of large eccentricity and large radius, and the second of zero eccentricity i.e., spherical, and of the small core radius. We give the results of such an attempt in figures 3.19 and 3.20. In order that we may compare these results with those found earlier, we keep the total mass the same i.e., $5 \times 10^{11} M_{\odot}$. We distribute the mass equally between the two spheroids, choose the axis ratio to be 1/10 for the 8 kpc radius spheroid, and take the radius of the sphere to be 1.5 kpc.

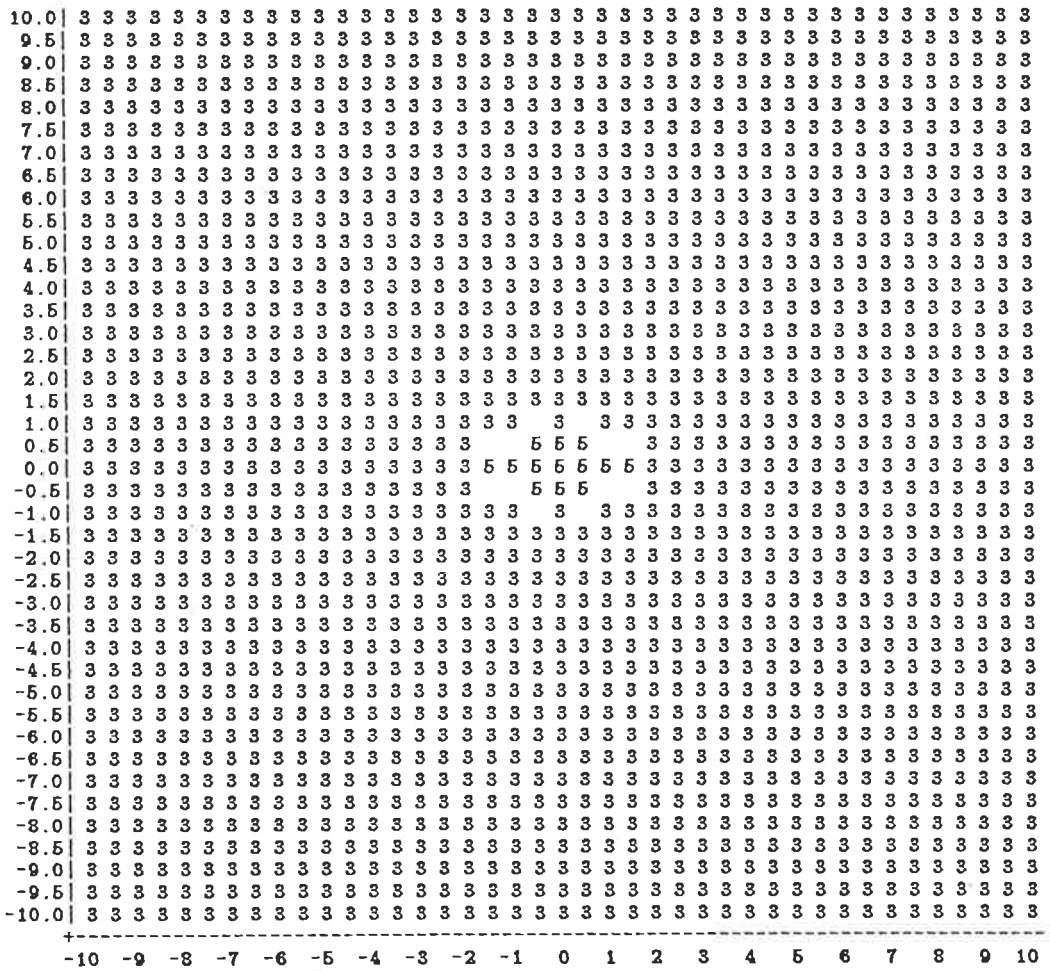

Figure 3.19 Multiplicity of images as a function of the position of the source, for a galaxy modelled by two concentric spheroids, each having a constant density distribution. The projected spheroid appears in fig.3.20.

We see that the region where the projected source has three images has increased dramatically, yet the five image region has not changed significantly. A

more interesting comparison between this model and the others is to compare the images due to a particular source. Thus, in fig 3.20 we present the images of the projected source $z = 0 + 0.5i$, as well as the projected spheroid itself.

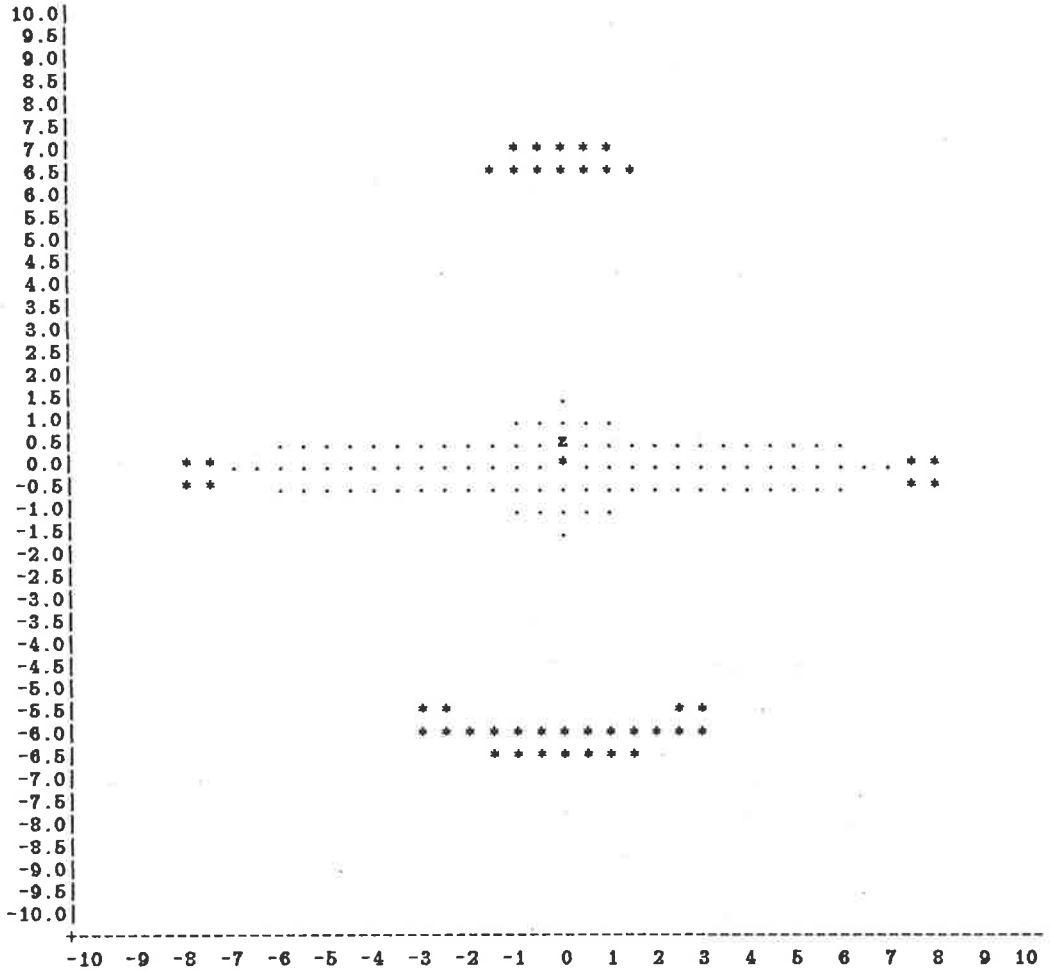

Figure 3.20 The projected spheroid, which generated fig.3.19. The stars are impact parameters for $z = 0 + 0.5i$. The central spheroid is a sphere with a radius of 1.5 kpc. The oblate spheroid has an axis ratio of 1/10 and a semi-major axis of 8 kpc. Both spheroids have the same mass of $2.5 \times 10^{11} M_{\odot}$.

Remarkably, the positions of the images are again very similar, with the main difference being the large magnification of the lower external image.

A summary of this the first part of the thesis is contained in the enclosed reprint of the same name.

So far we have considered only galaxies as gravitational lenses. In the next part of this thesis we consider what effect a black hole would have on the images of a single source, should it be interposed between the source and the observer.

PART II

Kerr Black Hole as a Gravitational Lens

§4 Introduction

§4.1 Geodesics in the Kerr Space-Time

In the first part of this thesis we considered curvature in space by introducing to the empty space many small masses in a regular pattern each of which had a spherically symmetric gravitational field. Furthermore, since the deviations of rays of light about these masses were small, we used the linearised approximation for the bending angles, and so we were able to find the total bending angle by summing the contributions of the individual masses. We found that the newly created symmetries had an effect on the multiplicity of images of a single source.

In this, the second part of the thesis, rather than use many spherically symmetric small masses, we start with a single large mass with an axially symmetric gravitational field, in particular, a Kerr black hole. We are interested in the same questions as before regarding the multiplicities of images of single sources, the shapes of the regions where multiplicity occurs, as well as intensifications and the orientations of the images.

A star whose mass is much greater than that of the sun cannot remain in equilibrium once its thermonuclear energy sources have been exhausted, and must undergo gravitational collapse. According to present theories, gravitational collapse can result in only one of three final states Harrison et. al. (1965): a white dwarf Landau (1934), Chandrasekhar (1935), a neutron star Oppenheimer and Volkoff (1939), Börner (1973), or a black hole. During the early stages of the collapse, the star might eject matter in such a way that each resulting component would have a sufficiently small mass to remain stable as a white dwarf or a neutron star. If this does not happen then further collapse is inevitable. Under fairly general conditions Pajerski and Newman (1972), it seems that black holes will form. If this occurs, then rigorous theorems Penrose (1965), Hawking and Penrose (1970), imply the existence of a space-time singularity and the existence of an absolute event horizon Penrose (1968), (1969), Hawking (1972).

As stars possess a considerable amount of angular momentum, it is the Kerr family of metrics, Kerr (1963), that are used to desribe the end state of the gravitational collapse. Motivation for studying the Kerr fields comes from astrophysical

arguments concerning the evolution of cores of galaxies. Lynden-Bell (1969) and Lynden-Bell and Rees (1971) argued that one might expect to find black holes with masses in the range $10^7 - 10^9$ solar masses commonly occuring in the centres of galaxies. Bardeen (1970) has pointed out that accretion onto these black holes from the orbiting matter in the galactic disk would produce a Kerr black hole, with angular momentum parameter only slightly less than the mass, after a finite amount of material had be accreted.

Unfortunately the existence of event horizons makes it impossible to observe the black holes directly, and so the only way to "observe" a black hole is by its effect upon its environment. Thus, it is necessary to know the equations of motion of particles in the vicinity of a Kerr black hole. Carter (1968) found the equations of motion for all geodesics, time-like, null and space-like. It had been known that if the angular momentum parameter a exceeded the mass m then a naked singularity would result, Carter also showed that there could be causality violations, in the form of closed time-like loops. Thus we assume that a is less than m, but by Bardeen, is close to m. A black hole with $a = m$ is known as an extreme Kerr black hole. The realistic geodesics, time-like and null can be divided into three groups: those which are captured by the black hole, those which are bound in orbit about the black hole, and those which escape to infinity. The observation of the two latter categories probably give us the best hope of detecting black holes. As a result, a great deal of work has been done in this area. General particle motion has been discussed by Wilkins (1972), Krivenko, Pyrogas and Zhuk (1976), de Felice and Calvani (1972), Stewart and Walker (1973). The optical appearance of a star orbiting an exterme Kerr black hole is discussed by Cunningham and Bardeen (1973) in an article of the same name. Null geodesics have been studied by Stewart and Walker, de Felice and Calvani, Helliwell and Mallinckrodt (1975), Godfrey (1970), Bardeen (1973).

Due to the complexity of the metric, the resultant equations of motion are also complicated. Thus, many people restricted their study to specific, simplified cases. Many writers work mainly with the extreme Kerr black hole where $a = m$, others restrict themselves to studying motion in the equatorial plane only. This was done by Boyer and Price (1965) as a means of interpreting the Kerr metric, and by

Boyer and Lindquist (1967) during their analytical extension of the metric. Others restricted themselves to a study of circular geodesics, with a view to compare them with those in the Schwarzschild metric. This was done by Ruffini and Wheeler (1970), Bardeen, Press and Teukolsky (1972), Bardeen (1974).

In the first part of this thesis, we developed interest in null geodesics that suffer small deviations due to an intervening massive body on their way to the observer. In this, the second part, we maintain that interest, but replace a galaxy as the intervening object by a Kerr black hole. It is evident that the complexity of the equations of motion, which can be expressed as elliptic integrals, caused a great deal of difficulty in the study of the geodesics in the Kerr metric. However, it occured to us that as in the Schwarzschild metric, where we used only the first order approximation to the bending angle, we may be able to extract most of the information required, when considering small deviations, from the lower order terms of the Kerr's equations of motion.

The essential difference between Schwarzschild and Kerr is that the latter contains spin, which destroys spherical symmetry. We saw earlier the shape of the region where a source has multiple images for the spherically symmetric case, now, amongst other things, we are interested in how the introduction of spin distorts this region.

In the Schwarzschild case whenever the first order term, m/r_{min}, was small, the resultant deviation was small. In the Kerr case this is still true, except that another factor plays an important role in determining the size of the deviation. Whether the ray takes a retrograde or a prograde path determines the size of the deviation for a particular ratio of m/r_{min}. We will see later how this manifests itself in the aproximate solutions to the equations of motion, which we will derive. It turns out that the effect of spin comes in only at the second order level i.e., the Kerr black hole is a Schwarzschild black hole to first order. This means that in order to see the effect of spin, the second order terms must give sufficiently large corrections to the first order terms. This implies that the deflection may not be too small.

Thus, the aim of this part of the thesis is to find approximations to the equations of motion of null geodesics in the Kerr metric that work well for rays whose

paths have been significantly effected by the spin of the black hole. The author believes that the approximations derived contain all the information necessary for most, if not all, astrophysical calculations, and are considerably simpler and easier to use than the original, exact equations.

§4.2 The Equations of Motion

As stated earlier, Carter derived the equations of motion in the Kerr space-time using the Kerr-Newmann form of the metric. These equations have been transformed to Boyer-Lindquist coordinates for simplicity by a number of people, including Bardeen (1973) and Wilkins (1972). For completeness of this work and to correct the value of a constant of the motion given by Bardeen (1973), we derive the equations of motion for null geodesics about a rotating black hole directly from Boyer-Lindquist coordinates.

The form of the Kerr metrics in the Boyer-Lindquist coordinates is given by:

$$ds^2 = \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + \frac{\sin^2\theta}{\rho^2}\left(a\,dt - (r^2+a^2)d\phi\right)^2 - \frac{\Delta}{\rho^2}\left(dt - a\sin^2\theta\,d\phi\right)^2, \quad (4.1)$$

where

$$\rho^2 = r^2 + a^2\cos^2\theta \quad (4.1.1)$$

and

$$\Delta = r^2 - 2mr + a^2. \quad (4.1.2)$$

In contravariant form, equation (4.1) may be expressed most conveniently as

$$\left(\frac{\partial}{\partial s}\right)^2 = \frac{\Delta}{\rho^2}\left(\frac{\partial}{\partial r}\right)^2 + \frac{1}{\rho^2}\left(\frac{\partial}{\partial \theta}\right)^2 + \frac{\Delta - a^2\sin^2\theta}{\rho^2\Delta\sin^2\theta}\left(\frac{\partial}{\partial \phi}\right)^2$$

$$-\frac{4mar}{\rho^2\Delta}\left(\frac{\partial}{\partial t}\right)\left(\frac{\partial}{\partial \phi}\right) + \frac{a^2\Delta\sin^2\theta - (r^2+a^2)^2}{\rho^2\Delta}\left(\frac{\partial}{\partial t}\right)^2. \quad (4.2)$$

The form of the Hamilton-Jacobi equation excluding electro-magnetic interaction is

$$\frac{\partial S}{\partial \tau} = \frac{1}{2}g^{ij}S_{,i}\,S_{,j} \quad (4.3)$$

where S is the Jacobi action, τ is the affine parameter and $S_{,i}$ means $\partial S/\partial x_i$,

where $x_i = x_i(\tau)$. Setting $(x_1, x_2, x_3, x_4) \equiv (r, \theta, \phi, t)$ we see from (4.2) that

$$g^{11} = \frac{\Delta}{\rho^2}$$

$$g^{22} = \frac{1}{\rho^2}$$

$$g^{33} = \frac{\Delta - a^2 \sin^2\theta}{\rho^2 \Delta \sin^2\theta}$$

$$g^{34} = -\frac{2mar}{\rho^2 \Delta}$$

$$g^{44} = \frac{a^2 \Delta \sin^2\theta - (r^2 + a^2)^2}{\rho^2 \Delta}.$$

If there is a separable solution, then in terms of the already known constants of the motion, energy E and angular momentum about the symmetry axis Φ, the Jacobi action must be of the form

$$S = -Et + \Phi\phi + S_\theta(\theta) + S_r(r), \tag{4.4}$$

where $S_\theta(\theta)$ and $S_r(r)$ are, respectively, functions of θ and r only. Substituting (4.4) into (4.3) gives

$$0 = \frac{\Delta}{\rho^2}\left(\frac{dS_r}{dr}\right)^2 + \frac{1}{\rho^2}\left(\frac{dS_\theta}{d\theta}\right)^2 + \frac{\Delta - a^2\sin^2\theta}{\rho^2 \Delta \sin^2\theta}\Phi^2$$

$$+ \frac{4mar}{\rho^2\Delta}\Phi E + \frac{a^2\Delta\sin^2\theta - (r^2+a^2)^2}{\rho^2\Delta}E^2.$$

Therefore, rearranging this equation to have functions of θ on the left hand side, and functions of r on the right hand side, we have

$$\left(\frac{dS_\theta}{d\theta}\right)^2 + \frac{\Phi^2}{\sin^2\theta} + a^2E^2\sin^2\theta$$

$$= -\Delta\left(\frac{dS_r}{dr}\right)^2 + \frac{E^2(r^2+a^2)^2}{\Delta} - \frac{4marE\Phi}{\Delta} + \frac{a^2\Phi^2}{\Delta}$$

$$= \mathcal{K}.$$

Note that $\mathcal{K} \geq 0$ as the equation involving θ is positive. Thus we write

$$\frac{dS_\theta}{d\theta} = \sqrt{\Theta(\theta)} \tag{4.5}$$

$$\frac{dS_r}{dr} = \sqrt{\frac{R(r)}{\Delta}}, \tag{4.6}$$

where the functions $\Theta(\theta)$ and $R(r)$ are defined by

$$\Theta(\theta) = \eta + a^2\cos^2\theta - \lambda^2\cot^2\theta, \tag{4.7}$$

$$R(r) = r[r(r^2+a^2) + 2a^2m] - 4amr\lambda - (r^2-2mr)\lambda^2 - \Delta\eta, \tag{4.8}$$

and where

$$\lambda = \frac{\Phi}{E} \quad \text{and} \quad \eta = \frac{\mathcal{K}}{E^2} - \lambda^2 - a^2. \tag{4.9}$$

Rewriting equations (4.5) and (4.6) in integral form

$$S_\theta = \int^\theta \sqrt{\Theta(\theta)}\, d\theta \tag{4.10}$$

and

$$S_r = \int^r \frac{\sqrt{R(r)}}{\Delta} dr. \tag{4.11}$$

Thus $S = -Et + \Phi\phi + S_\theta(\theta) + S_r(r)$ where S_θ and S_r are defined by equations (4.10) and (4.11) respectively.

Using the fact that derivatives of the Hamilton-Jacobi action S with respect to the constants of the motion are themselves constant, $\partial S/\partial E$ gives

$$t = \int^{r} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}}dr + \int^{\theta} \frac{a^2\cos^2\theta}{\pm[\Theta(\theta)]^{1/2}}d\theta, \tag{4.12}$$

$\partial S/\partial\Phi$ gives

$$\phi = \int^{r} \frac{[r^2\lambda+2mr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}}dr + \int^{\theta} \frac{\lambda\cot^2\theta}{\pm[\Theta(\theta)]^{1/2}}d\theta, \tag{4.13}$$

and $\partial S/\partial Q$ gives

$$\int^{r} \frac{dr}{\pm[R(r)]^{1/2}} = \int^{\theta} \frac{d\theta}{\pm[\Theta(\theta)]^{1/2}}. \tag{4.14}$$

These are the equations of motion of null geodesics in the Kerr metric.

§5 Solving the equations of motion

§5.1 Solution for θ in the case $m = a = 0$

In this section we set up the equations for the case where $m = a = 0$ and solve for θ.

By chapter 4 the equations of motion for a ray of light about a rotating black hole in the Boyer-Lindquist coordinates (r, θ, ϕ, t) are:

$$\int^{r} \frac{dr}{\pm[R(r)]^{1/2}} = \int^{\theta} \frac{d\theta}{\pm[\Theta(\theta)]^{1/2}}, \tag{5.1}$$

$$\phi = \int^{r} \frac{[r^2\lambda + 2mr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}} dr + \int^{\theta} \frac{\lambda \cot^2\theta}{\pm[\Theta(\theta)]^{1/2}} d\theta \tag{5.2}$$

and

$$t = \int^{r} \frac{[r^2(r^2+a^2) + 2amr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}} dr + \int^{\theta} \frac{a^2\cos^2\theta}{\pm[\Theta(\theta)]^{1/2}} d\theta, \tag{5.3}$$

where

$$R(r) = r[r(r^2+a^2) + 2a^2m] - 4amr\lambda - (r^2 - 2mr)\lambda^2 - \Delta\eta, \tag{5.4}$$

$$\Theta(\theta) = \eta + a^2\cos^2\theta - \lambda^2\cot^2\theta \tag{5.5}$$

and

$$\Delta = r^2 - 2mr + a^2. \tag{5.6}$$

The constants m and a are the mass and spin parameters of the Kerr metric with $m = |m| \geq |a|$, λ and η are constants of the motion, see §4.2, which will be discussed later. The integrals are along the path of motion.

Suppose a ray of light originates at the source with coordinates $(r_s, \theta_s, \phi_s, t_s)$, passes by the black hole and arrives at the observer's coordinates $(r_o, \theta_o, \phi_o, t_o)$. Along this path r cordinate ranges from r_s to r_{min}, to r_o, while theta ranges from θ_s to θ_{min} or θ_{max} through to θ_o. These are the relevant integral paths, see fig.5.1. The minimum value of r, r_{min}, is the largest of the real roots of $R(r)$, whilst θ_{min} and θ_{max} are the solutions of $\Theta(\theta) = 0$.

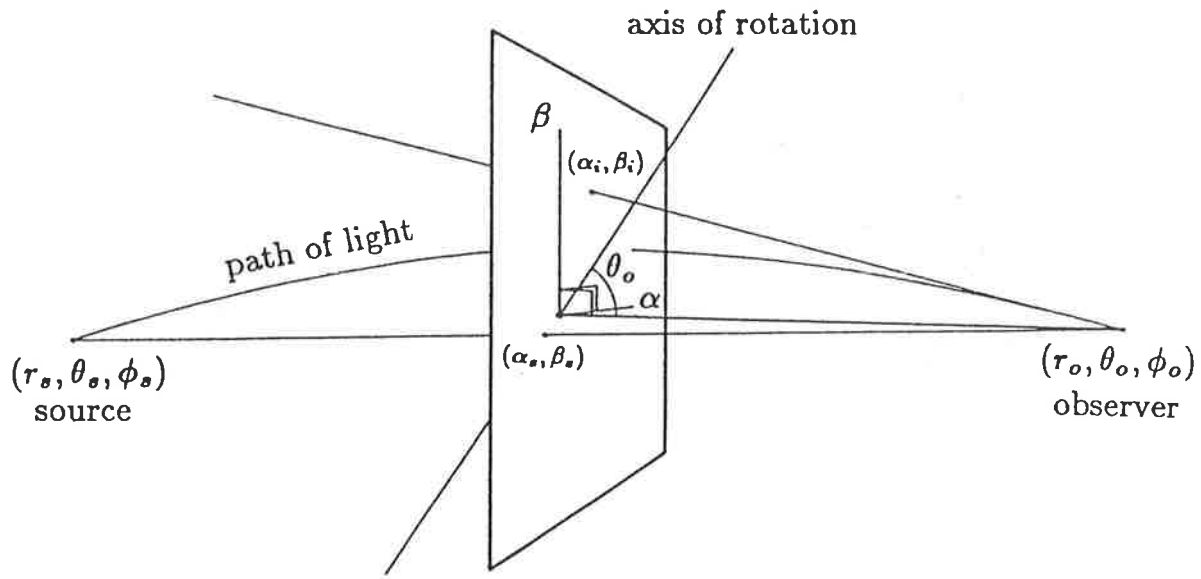

Figure 5.1 The polar coordinates (r, θ, ϕ) are defined by the axis of rotation. The α-β plane is perpendicular to the line passing through (0,0) - the position of the black hole, and the observer. Point (α_i, β_i) is the observed position of an image of the source, whilst (α_s, β_s) is the projected position of the source onto the α-β plane along the line through the source and the observer. It can be shown that $\lim_{m \to 0} (\alpha_i, \beta_i) = (\alpha_s, \beta_s)$.

The unlikely event where $\theta = \theta(r)$ does not have an extremum will be discussed later. The sign of $[R(r)]^{1/2}$ is positive if we are integrating in the direction of increasing values of r, and negative if otherwise. Similarly the sign of $[\Theta(\theta)]^{1/2}$ is taken to be positive if integrating in the direction of increasing values of θ and negative otherwise.

To get an understanding of what λ and η are, it proves to be most instructive to firstly solve the above equations of motion in the case where m and therefore a are both zero, and then relate λ and η to the position of an image of the source. These equations turn out to be useful since they are the solutions of tangents to the rays of light evaluated at the observer.

For $m = a = 0$ equation (5.1) becomes

$$\int^{r} \frac{dr}{r^2 \sqrt{1 - \frac{\lambda^2 + \eta}{r^2}}} = \int^{\theta} \frac{d\theta}{\sqrt{\eta - \lambda^2 \cot^2 \theta}}, \tag{5.7}$$

and (5.2) becomes

$$\phi = \int^{r} \frac{\lambda dr}{r^2\sqrt{1-\frac{\lambda^2+\eta}{r^2}}} + \int^{\theta} \frac{\lambda \cot^2\theta d\theta}{\sqrt{\eta - \lambda^2\cot^2\theta}}.$$

Using equation (5.7)

$$\phi = \lambda \int^{\theta} \frac{d\theta}{\sin^2\theta\sqrt{\eta - \lambda^2\cot^2\theta}}. \tag{5.8}$$

In order to find the integral paths we need to know r_{min} and θ_{min} or θ_{max}. To get r_{min} we set $R_{m=0}(r_{min}) = 0$. Since r_{min} is the largest of the real roots of $R(r)$, for the $m = 0$ case, $r_{min} = \sqrt{\lambda^2 + \eta}$. Similarly θ_{min} or θ_{max} are determined by setting $\Theta_{a=0}(\theta_{min/max}) = 0$. Thus,

$$\cot\theta_{min/max} = \pm\frac{\sqrt{\eta}}{\lambda} \quad \text{or} \quad \cos\theta_{min/max} = \pm\sqrt{\frac{\eta}{\lambda^2+\eta}}. \tag{5.9}$$

where the positive sign corresponds to $\cos\theta_{min}$ and the negative sign corresponds to $\cos\theta_{max}$. The way to find whether θ reaches θ_{max} or θ_{min} along the path will be discussed later.

Immediately from the results for r_{min} and $\cos\theta_{min/max}$, we can get an idea of what λ and η are. Suppose we are positioned in such a way that the ray of light, whose path is a straight line as the mass is zero, is perpendicular to the $\alpha - \beta$ plane of fig.5.1. Suppose further that the axis of rotation and the β axis are coincident, thus $\theta_o \approx \pi/2$. Then both r_{min} and $\theta_{min/max}$ occur in the $\alpha - \beta$ plane, with the resultant natural interpretation of λ and η, see fig.5.2. From the diagram we see that the position of the image of the source, (α_i, β_i), is directly related to λ and η. In particular, we have $\lambda^2 = \alpha_i^2$ and $\eta = \beta_i^2$, with the sign of equation (5.9) determining whether the θ extremum is θ_{min} or θ_{max}. Of the two assumptions, the former, namely that the ray path is perpendicular to the $\alpha - \beta$ plane, will be nearly always satisfied. This is due to the definition of the plane as

well as the fact that r_s and r_o are large, and that we are only interested in sources whose rays are significantly influenced by the black hole, thus implying that the (α_i, β_i) are near the black hole. The latter assumption, namely that $\theta_o \approx \pi/2$, will rarely be satisfied. Thus, we expect to be able to express λ^2 and η as α_i^2 and β_i^2 respectively, with both relations modified by some trigonometric functions of θ_o. The exact relations we find later. However for the moment it is useful to think of λ as the horizontal component of the position of an image, and of $\sqrt{\eta}$ as the vertical component of an image.

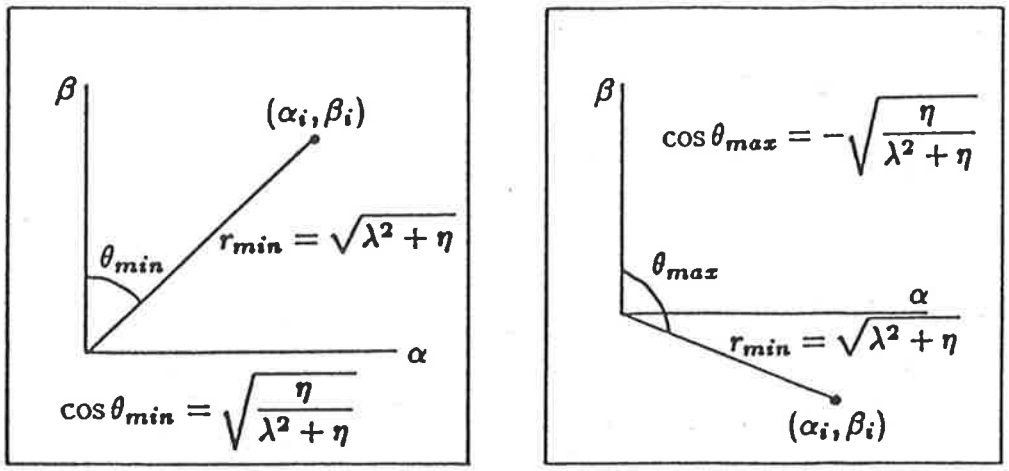

Figure 5.2 Interpretation of λ and η. Assuming that the ray's path is perpendicular to the plane and that the axis of rotation and the β axis are coincident, then the minimum value of r along the ray's path occurs in the $\alpha - \beta$ plane of fig.5.1, as does the extremum of θ. The two diagrams correspond to the two signs of equation (5.9). We see that by construction $\lambda^2 = \alpha_i^2$ and $\eta = \beta_i^2$.

Having found r_{min} and $\theta_{min/max}$ we are in a position to solve the equations of motion for the $m = 0$ case. Equation (5.7) becomes

$$\int_{\sqrt{\lambda^2+\eta}}^{r_s} \frac{dr}{r^2\sqrt{1-\frac{\lambda^2+\eta}{r^2}}} + \int_{\sqrt{\lambda^2+\eta}}^{r_o} \frac{dr}{r^2\sqrt{1-\frac{\lambda^2+\eta}{r^2}}}$$

$$= \int_{\theta_s}^{\theta_{max}} \frac{d\theta}{\sqrt{\eta - \lambda^2\cot^2\theta}} + \int_{\theta_o}^{\theta_{max}} \frac{d\theta}{\sqrt{\eta - \lambda^2\cot^2\theta}}$$

for the $\theta_s \to \theta_{max} \to \theta_o$ path, and

$$= \int_{\theta_{min}}^{\theta_s} \frac{d\theta}{\sqrt{\eta - \lambda^2 \cot^2\theta}} + \int_{\theta_{min}}^{\theta_o} \frac{d\theta}{\sqrt{\eta - \lambda^2 \cot^2\theta}} \quad (5.10)$$

for the $\theta_s \to \theta_{min} \to \theta_o$ path.

Consider

$$\int \frac{dr}{r^2\sqrt{1 - \frac{\lambda^2+\eta}{r^2}}}. \quad \text{Let } \cos\beta = \frac{\sqrt{\lambda^2+\eta}}{r} \quad \Rightarrow \quad -\sin\beta\, d\beta = -\frac{\sqrt{\lambda^2+\eta}}{r^2} dr$$

$$= \frac{1}{\sqrt{\lambda^2+\eta}} \int \frac{\sin\beta\, d\beta}{\sqrt{1-\cos^2\beta}}$$

$$= \frac{1}{\sqrt{\lambda^2+\eta}} \arccos\frac{\sqrt{\lambda^2+\eta}}{r}.$$

So the left hand side of (5.10) is

$$\frac{1}{\sqrt{\lambda^2+\eta}}\left(\arccos\frac{\sqrt{\lambda^2+\eta}}{r_s} + \arccos\frac{\sqrt{\lambda^2+\eta}}{r_o}\right). \quad (5.11)$$

Now consider

$$\int \frac{d\theta}{\sqrt{\eta - \lambda^2\cot^2\theta}}. \quad \text{Let } \cos\theta = \sqrt{\frac{\eta}{\lambda^2+\eta}}\cos\alpha \quad \Rightarrow \quad -\sin\theta\, d\theta = -\sqrt{\frac{\eta}{\lambda^2+\eta}}\sin\alpha\, d\alpha.$$

$$= \sqrt{\frac{\eta}{\lambda^2+\eta}} \int \frac{\sin\alpha\, d\alpha}{\sqrt{\eta\sin^2\theta - \lambda^2\cos^2\theta}}$$

$$= \sqrt{\frac{\eta}{\lambda^2+\eta}} \int \frac{\sin\alpha\, d\alpha}{\sqrt{\eta\left(1 - \frac{\eta}{\lambda^2+\eta}\cos^2\alpha\right) - \frac{\lambda^2\eta}{(\lambda^2+\eta)^2}\cos^2\alpha}}$$

$$= \sqrt{\frac{\eta}{\lambda^2+\eta}} \int \frac{\sin\alpha\, d\alpha}{\sqrt{\eta(1-\cos^2\alpha)}}$$

$$= \frac{1}{\sqrt{\lambda^2+\eta}} \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta\right). \quad (5.12)$$

So with the aid of (5.9) the right hand side of (5.10) becomes

$$\frac{1}{\sqrt{\lambda^2+\eta}}\left(2\pi - \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_s\right) - \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_o\right)\right)$$

for the $\theta_s \to \theta_{max} \to \theta_o$ path, and

$$\frac{1}{\sqrt{\lambda^2+\eta}}\left(\arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_s\right) + \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_o\right)\right)$$

for $\theta_s \to \theta_{min} \to \theta_o$ path. Combining the two paths we write

$$\frac{1}{\sqrt{\lambda^2+\eta}}\left(\begin{matrix}2\pi\\0\end{matrix} \mp \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_s\right) \mp \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_o\right)\right), \quad (5.13)$$

where 2π and the negative signs correspond to the $\theta_s \to \theta_{max} \to \theta_o$ path, whilst 0 and the positive signs correspond to the $\theta_s \to \theta_{min} \to \theta_o$ path. Note that we assumed that $\eta > 0$, for if $\eta = 0$ then $\theta(r) = \pi/2 \quad \forall \quad r$, is the trivial solution.

In equating (5.13) with (5.11) it turns out to be more convenient to use α_s and α_o than $\arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_s\right)$ and $\arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta_o\right)$ respectively. So we have

$$\frac{1}{\sqrt{\lambda^2+\eta}}\left(\begin{matrix}2\pi\\0\end{matrix} \mp \alpha_s \mp \alpha_o\right) = \frac{1}{\sqrt{\lambda^2+\eta}}\left(\arccos\frac{\sqrt{\lambda^2+\eta}}{r_s} + \arccos\frac{\sqrt{\lambda^2+\eta}}{r_o}\right)$$

$$= \frac{1}{\sqrt{\lambda^2+\eta}}\arccos\left(\frac{\lambda^2+\eta}{r_s r_o} - \sqrt{1-\frac{\lambda^2+\eta}{r_s^2}}\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}\right),$$

$$\Rightarrow \quad \begin{matrix}2\pi\\0\end{matrix} \mp \alpha_s \mp \alpha_o = \arccos\left(\frac{\lambda^2+\eta}{r_s r_o} - \sqrt{1-\frac{\lambda^2+\eta}{r_s^2}}\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}\right). \quad (5.14)$$

Consider

$$\frac{\lambda^2+\eta}{r_s r_o} - \sqrt{1-\frac{\lambda^2+\eta}{r_s^2}}\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}.$$

As both r_s and r_o are assumed to be very much greater than $r_{min} = \sqrt{\lambda^2+\eta}$,

$$\begin{aligned}
&\frac{\lambda^2+\eta}{r_s r_o} - \sqrt{1-\frac{\lambda^2+\eta}{r_s^2}}\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}} \\
&\qquad \approx \frac{\lambda^2+\eta}{r_s r_o} - \left(1-\frac{\lambda^2+\eta}{2r_s^2}\right)\left(1-\frac{\lambda^2+\eta}{2r_o^2}\right) \\
&\qquad \approx \frac{\lambda^2+\eta}{r_s r_o} - 1 + \frac{\lambda^2+\eta}{2r_s^2} + \frac{\lambda^2+\eta}{2r_o^2} \\
&\qquad = -1 + \frac{(\lambda^2+\eta)(r_s+r_o)^2}{2r_s^2 r_o^2} \\
&\qquad = -1\left(1 - \frac{\left(\frac{\sqrt{\lambda^2+\eta}(r_s+r_o)}{r_s r_o}\right)^2}{2}\right) \\
&\qquad \approx \cos\pi\cos\left(\sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}\right) \\
&\qquad = \cos\left(\pi \pm \sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}\right).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\arccos\left(\frac{\lambda^2+\eta}{r_s r_o} - \sqrt{1-\frac{\lambda^2+\eta}{r_s^2}}\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}\right) &\approx \arccos\left(\cos\left(\pi \pm \sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}\right)\right) \\
&= \pi - \sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}.
\end{aligned}$$

Note that the $\pm$ sign becomes a minus sign. This is due to the definition of arccos. Thus, equation (5.14) becomes

$$\begin{matrix}2\pi \\ 0\end{matrix} \mp \alpha_s \mp \alpha_o \approx \pi - \sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}$$

$$\Rightarrow \quad \pm\pi \mp \alpha_s \mp \alpha_o \approx -\sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}$$

$$\Rightarrow \quad -\pi + \alpha_s + \alpha_o \approx \pm\sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}. \tag{5.15}$$

Defining $\delta = -\sqrt{\lambda^2+\eta}\frac{r_s+r_o}{r_s r_o}$, we have

$$\alpha_s \approx \pi - \alpha_o \mp \delta$$

$$\cos\alpha_s \approx -\cos\alpha_o \cos\delta \pm \sin\delta \sin\alpha_o.$$

Using the definition of α, namely $\alpha = \arccos\left(\sqrt{\frac{\lambda^2+\eta}{\eta}}\cos\theta\right)$,

$$\cos\theta_s \approx -\cos\theta_o \cos\delta \pm \sin\delta\sqrt{\frac{\eta}{\lambda^2+\eta} - \cos^2\theta_o}. \tag{5.16}$$

Note that in the limit as r_s or r_o tend to infinity we have $\delta = 0$, thus by equation (5.16), $\cos\theta_s = -\cos\theta_o$, which implies that $\theta_s + \theta_o = \pi$ as expected.

§5.2 Solution for ϕ in the case $m = a = 0$

Having solved for θ_s we now solve for ϕ_s. Equation (5.8) states that

$$\phi = \phi_o - \phi_s = \lambda \int^{\theta} \frac{d\theta}{\sin^2\theta\sqrt{\eta - \lambda^2\cot^2\theta}}.$$

As before let $\cos\theta = \sqrt{\frac{\eta}{\lambda^2+\eta}}\cos\alpha \quad \Rightarrow \quad -\sin\theta\, d\theta = -\sqrt{\frac{\eta}{\lambda^2+\eta}}\sin\alpha\, d\alpha$, then

$$\begin{aligned}
\phi_o - \phi_s &= \frac{\lambda}{\sqrt{\lambda^2+\eta}} \int^{\alpha} \frac{d\alpha}{\sin^2\theta} \\
&= \frac{\lambda}{\sqrt{\lambda^2+\eta}} \int^{\alpha} \frac{d\alpha}{1 - \frac{\eta}{\lambda^2+\eta}\cos^2\alpha} \\
&= \frac{\lambda}{\sqrt{\lambda^2+\eta}} \left(\frac{-1}{\sqrt{1 - \frac{\eta}{\lambda^2+\eta}}} \arctan\left(\sqrt{1 - \frac{\eta}{\lambda^2+\eta}}\cot\alpha\right)\right) \\
&= -\frac{\lambda}{|\lambda|} \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha\right).
\end{aligned}$$

When $\theta = \theta_{max}$ i.e., $\cos\theta_{max} = -\sqrt{\frac{\eta}{\lambda^2+\eta}}$, $\alpha = \pi$. Whereas when $\theta = \theta_{min}$ i.e., $\cos\theta_{min} = \sqrt{\frac{\eta}{\lambda^2+\eta}}$, $\alpha = 0$. Therefore for the $\theta_s \to \theta_{max} \to \theta_o$ path

$$\phi_o - \phi_s = -\frac{\lambda}{|\lambda|}\left(-\pi - \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_s\right) - \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_o\right)\right),$$

whilst for thé $\theta_s \to \theta_{min} \to \theta_o$ path

$$= -\frac{\lambda}{|\lambda|}\left(-\pi + \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_s\right) + \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_o\right)\right).$$

Combining the two expressions,

$$\phi_o - \phi_s = \frac{\lambda}{|\lambda|}\left(\pi \pm \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_s\right) \pm \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\cot\alpha_o\right)\right).$$

In order to combine the two arctans we would like that $\cot\alpha_s \cot\alpha_o < 1$ since then

$$\phi_o - \phi_s = \frac{\lambda}{|\lambda|}\left(\pi \pm \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\frac{(\cot\alpha_s + \cot\alpha_o)}{(1 - \frac{\lambda^2}{\lambda^2+\eta}\cot\alpha_s\cot\alpha_o)}\right)\right),$$

otherwise one has to add or subtract π depending on other conditions. Fortunately we can show that $\cot\alpha_s \cot\alpha_o < 1$, although this proves to be rather involved.

Recall equation (5.15), which when slightly rearranged is

$$\pm\delta \approx \pi - \alpha_s - \alpha_o$$

$$= -(\alpha_s - \frac{\pi}{2} + \alpha_o - \frac{\pi}{2}),$$

where $\delta = -\sqrt{\lambda^2+\eta}\frac{r_s \pm r_o}{r_s r_o}$, which implies that $\delta < 0$ and $|\delta| \ll 1$.

Taking tan of both sides, and with $\tan\delta \approx \delta$

$$\pm\delta \approx -\tan(\alpha_s - \frac{\pi}{2} + \alpha_o - \frac{\pi}{2})$$

$$= -\frac{\tan(\alpha_s - \frac{\pi}{2}) + \tan(\alpha_o - \frac{\pi}{2})}{1 - \tan(\alpha_s - \frac{\pi}{2})\tan(\alpha_o - \frac{\pi}{2})}$$

$$= \frac{\cot\alpha_s + \cot\alpha_o}{1 - \cot\alpha_s\cot\alpha_o}$$

$$\Rightarrow \cot\alpha_s \approx \frac{\pm\delta - \cot\alpha_o}{1 \pm \delta\cot\alpha_o}. \qquad (5.17)$$

Therefore,

$$\cot\alpha_s\cot\alpha_o \approx \frac{\pm\delta - \cot\alpha_o}{1 \pm \delta\cot\alpha_o}\cot\alpha_o.$$

Suppose the ray of light takes the path $\theta_s \rightarrow \theta_{max} \rightarrow \theta_o$, then α ranges $\alpha_s \rightarrow \pi \rightarrow \alpha_o$. Since $\delta \approx \pi - \alpha_s - \alpha_o$ and $\delta < 0$ it follows that $\alpha_s \in [|\delta|, \pi]$ and $\alpha_o \in [|\delta|, \pi]$, where to get the minimum value of one angle we maximised the other. Similarly for the $\theta_s \rightarrow \theta_{min} \rightarrow \theta_o$ path, $-\delta \approx \pi - \alpha_s - \alpha_o$ from which it follows that $\alpha_s \in [0, \pi - |\delta|]$ and $\alpha_s \in [0, \pi - |\delta|]$.

Thus, for the $\theta_s \rightarrow \theta_{max} \rightarrow \theta_o$ path we wish to show that

$$\cot\alpha \frac{\delta - \cot\alpha}{1 + \delta\cot\alpha} = \cot\alpha \frac{-|\delta| - \cot\alpha}{1 - |\delta|\cot\alpha} < 1 \quad \forall \quad \alpha \in [|\delta|, \pi].$$

Since $1 - |\delta|\cot\alpha \geq 0$, $1 - |\delta|\cot\alpha \approx 0$ for $\alpha = |\delta|$, we need only consider the case when $-\cot\alpha(|\delta| + \cot\alpha) > 0$, otherwise the result is ≤ 0. Now,

$$\begin{aligned} & -\cot\alpha(|\delta| + \cot\alpha) > 0 \\ \Rightarrow \quad & \cot\alpha(|\delta| + \cot\alpha) < 0 \\ \Rightarrow \quad & \cot\alpha < 0 \quad \text{and} \quad |\delta| + \cot\alpha > 0 \\ \Rightarrow \quad & |\cot\alpha| < |\delta|. \end{aligned}$$

Therefore, in this case

$$\begin{aligned} \cot\alpha \frac{-|\delta| - \cot\alpha}{1 - |\delta|\cot\alpha} &= |\cot\alpha| \frac{|\delta| - |\cot\alpha|}{1 + |\delta||\cot\alpha|} \\ &< |\cot\alpha|(|\delta| - |\cot\alpha|) \\ &< \delta^2 \\ &< 1. \end{aligned}$$

Similarly, for the $\theta_s \rightarrow \theta_{min} \rightarrow \theta_o$ path

$$\cot\alpha \frac{-\delta - \cot\alpha}{1 - \delta\cot\alpha} = \cot\alpha \frac{|\delta| - \cot\alpha}{1 + |\delta|\cot\alpha} < 1 \quad \forall \quad \alpha \in [0, \pi - |\delta|],$$

since $1 + |\delta|\cot\alpha \geq 0$ and $\cot\alpha(|\delta| - \cot\alpha) > 0$ imply we need to consider the

case $\cot\alpha > 0$ and $|\delta| - \cot\alpha > 0$ only, which impies $|\delta| > |\cot\alpha|$. This

$$\Rightarrow \quad \cot\alpha \frac{|\delta| - \cot\alpha}{1 + |\delta|\cot\alpha} = |\cot\alpha| \frac{|\delta| - |\cot\alpha|}{1 + |\delta||\cot\alpha|} < 1 \quad \text{from above.}$$

Thus,

$$\phi_o - \phi_s = \frac{\lambda}{|\lambda|}\left(\pi \pm \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\frac{(\cot\alpha_s + \cot\alpha_o)}{(1 - \frac{\lambda^2}{\lambda^2+\eta}\cot\alpha_s\cot\alpha_o)}\right)\right).$$

Defining $k^2 = \frac{\eta}{\lambda^2+\eta} \quad \Rightarrow \quad 1 - k^2 = \frac{\lambda^2}{\lambda^2+\eta}$ and using (5.17) we have

$$\begin{aligned}
\frac{\cot\alpha_s + \cot\alpha_o}{1 - \frac{\lambda^2}{\lambda^2+\eta}\cot\alpha_s\cot\alpha_o} &= \frac{\cot\alpha_s + \cot\alpha_o}{1 - (1-k^2)\cot\alpha_s\cot\alpha_o} \\
&= \frac{\frac{\pm\delta - \cot\alpha_o}{1 \pm \delta\cot\alpha_o} + \cot\alpha_o}{1 - (1-k^2)\cot\alpha_o\frac{\pm\delta - \cot\alpha_o}{1 \pm \delta\cot\alpha_o}} \\
&= \frac{\pm\delta(1 + \cot^2\alpha_o)}{1 \pm \delta\cot\alpha_o - (1-k^2)\cot\alpha_o(\pm\delta - \cot\alpha_o)} \\
&= \frac{\pm\delta(1 + \cot^2\alpha_o)}{1 + (1-k^2)\cot^2\alpha_o \pm \delta k^2\cot\alpha_o} \\
&\approx \frac{\pm\delta(1 + \cot^2\alpha_o)}{1 + (1-k^2)\cot^2\alpha_o}\left(1 \mp \frac{\delta k^2\cot\alpha_o}{1 + (1-k^2)\cot^2\alpha_o}\right),
\end{aligned}$$

where we used the fact that $|\delta|$ is small. So

$$\begin{aligned}
\phi_o - \phi_s &\approx \frac{\lambda}{|\lambda|}\left(\pi \pm \arctan\left(\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\left(\frac{\pm\delta(1 + \cot^2\alpha_o)}{1 + (1-k^2)\cot^2\alpha_o}\right)\right.\right. \\
&\qquad \left.\left. \times\left(1 \mp \frac{\delta k^2\cot\alpha_o}{1 + (1-k^2)\cot^2\alpha_o}\right)\right)\right) \\
&\approx \frac{\lambda}{|\lambda|}\left(\pi + \delta\sqrt{\frac{\lambda^2}{\lambda^2+\eta}}\left(\frac{1 + \cot^2\alpha_o}{1 + (1-k^2)\cot^2\alpha_o}\right)\left(1 \mp \frac{\delta k^2\cot\alpha_o}{1 + (1-k^2)\cot^2\alpha_o}\right)\right).
\end{aligned}$$

Recall that

$$\cos\theta = k\cos\alpha \quad \Rightarrow \quad \cot^2\alpha = \frac{\cos^2\theta}{k^2-\cos^2\theta}, \quad \text{where } k = \sqrt{\tfrac{\eta}{\lambda^2+\eta}}.$$

Therefore,

$$\frac{1+\cot^2\alpha_o}{1+(1-k^2)\cot^2\alpha_o} = \frac{k^2}{(k^2-\cos^2\theta_o)\left(1+(1-k^2)\frac{\cos^2\theta_o}{k^2-\cos^2\theta_o}\right)}$$

$$= \csc^2\theta_o.$$

And finally we have

$$\phi_o - \phi_s \approx \pi\frac{\lambda}{|\lambda|} + \frac{\lambda}{\sqrt{\lambda^2+\eta}}\delta\csc^2\theta_o\left(1 \mp \delta\cot\theta_o\csc\theta_o\sqrt{\frac{\eta}{\lambda^2+\eta} - \cos^2\theta_o}\right), \quad (5.18)$$

where

$$\delta = -\sqrt{\lambda^2+\eta}\left(\frac{r_s+r_o}{r_s r_o}\right),$$

and where the negative sign corresponds to the $\theta_s \to \theta_{max} \to \theta_o$ path, whilst the positive sign corresponds to the $\theta_s \to \theta_{min} \to \theta_o$ path. It is at first surprising to see a term of the form $\lambda/|\lambda|$ as this does not have a limit as λ goes to zero. However this merely corresponds to the equivalence of $-\pi$ and π. In the limit as r_s or r_o go to infinity $\phi_o - \phi_s = \pi$ as expected.

Note that for the case where the observer lies in the equatorial plane of the black hole i.e., $\theta_o = \pi/2$, both equations (5.16) and (5.18) simplify considerably to give

$$\cos\theta_s \approx \pm\sin\delta\sqrt{\frac{\eta}{\lambda^2+\eta}} \quad \text{and} \quad \phi_o - \phi_s \approx \pi\frac{\lambda}{|\lambda|} + \delta\frac{\lambda}{\sqrt{\lambda^2+\eta}}. \quad (5.19)$$

Since the effect of spin is most evident for $\theta_o = \pi/2$ we will tend to concern ourselves mostly with this case.

Thus, we have found the position of the source given the position of the observer for the case $m = a = 0$. All that is left is to solve equation (5.3) i.e., to find t for $m = a = 0$. This happens to be trivial and so we will not derive it here, but rather leave it until it proves to be useful in the derivation of t for $m \neq 0$.

§5.3 Relating λ and η to the position of the image

Recall that in section 5.1 we tried to give an approximate interpretation of the meaning of the two constants of the motion λ and η. We found that they are directly related to the position of an image of the source. In particular, under special conditions the relations are extremely simple, namely $\lambda^2 = \alpha_i^2$ and $\eta = \beta_i^2$. In this section we derive the exact relationships between λ and η, and the position of an image (α_i, β_i).

Having found θ_s and ϕ_s, we need to relate them to the projected position of the source. Positions of sources and their images are projected onto the $\alpha - \beta$ plane, which is defined in fig.5.1. For convenience we reproduce it here.

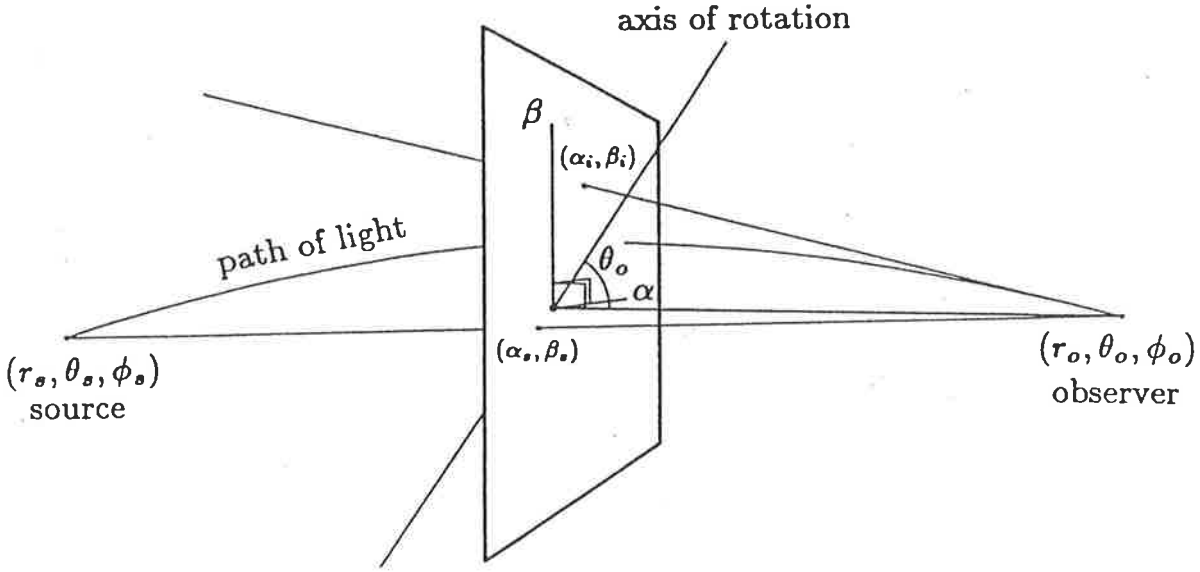

Figure 5.3 Reproduction of fig. 5.1. The polar coordinates (r, θ, ϕ) are defined by the axis of rotation. The α-β plane is perpendicular to the line connecting the observer and (0,0) - the position of the black hole. Point (α_i, β_i) is the observed position of an image of the source, whilst (α_s, β_s) is the projected position of the source onto the α-β plane along the line joining the source and the observer. It can be shown that $\lim_{m \to 0} (\alpha_i, \beta_i) = (\alpha_s, \beta_s)$.

The point (α_s, β_s) is the projection of the source at (r_s, θ_s, ϕ_s) onto the plane along the line connecting the source with the observer at (r_o, θ_o, ϕ_o). The point (α_i, β_i) is the position of an image of the source which is found by calculating the

intersection of the tangent to the ray at the observer with the plane. We first find (α_s, β_s) and then (α_i, β_i).

The $\alpha - \beta$ plane defined above can be written in rectilinear coordinates as $\alpha = y$ and $\beta = z \sin\theta_o$, where the rectilinear are defined in the usual way by the axis of rotation i.e., the z axis. So to find (α_s, β_s) we simply find the equation of the line connecting the points $(r_s \cos\phi_s \sin\theta_s, r_s \sin\phi_s \sin\theta_s, r_s \cos\theta_s)$ and $(r_o \cos\phi_o \sin\theta_o, r_o \sin\phi_o \sin\theta_o, r_o \cos\theta_o)$ as a function of x, and then set $x = 0$ to get the intersection of the line with the plane $x = 0$. Without any loss of generality we choose $\phi_o = 0$. Thus we have

$$\alpha_s = y_{x=0} = \frac{r_s r_o \sin\phi_s \sin\theta_s \sin\theta_o}{r_o \sin\theta_o - r_s \cos\phi_s \sin\theta_s} \tag{5.20}$$

and

$$\frac{\beta_s}{\sin\theta_o} = z_{x=0} = \frac{r_s r_o (\cos\theta_s \sin\theta_o - \cos\phi_s \sin\theta_s \cos\theta_o)}{r_o \sin\theta_o - r_s \cos\phi_s \sin\theta_s}. \tag{5.21}$$

Now, we have found that λ and η can be related to the observed position of an image of the source. Suppose the observer is at the coordinates $(r_o \sin\theta_o, 0, r_o \cos\theta_o)$ i.e., as before we set $\phi_o = 0$. Define α_i and β_i by making the tangent to the ray of light at the observer pass through the point $(0, \alpha_i, \beta_i / \sin\theta_o)$ i.e., intersect the $\alpha - \beta$ plane at (α_i, β_i). As the tangent passes through $(r_o \sin\theta_o, 0, r_o \cos\theta_o)$ as well as $(0, \alpha_i, \beta_i / \sin\theta_o)$ its equation may be written

$$(x_t, y_t, z_t) = (x_t, -\frac{\alpha_i}{r_o \sin\theta_o} x_t + \alpha_i, \frac{r_o \cos\theta_o \sin\theta_o - \beta_i}{r_o \sin^2\theta_o} x_t + \frac{\beta_i}{\sin\theta_o}).$$

Therefore, using $x_t = r \cos\phi \sin\theta$,

$$y_t = -\frac{\alpha_i}{r_o \sin\theta_o} r \cos\phi \sin\theta + \alpha_i,$$

$$\Rightarrow \qquad \frac{dy_t}{dr} = -\frac{\alpha_i}{r_o \sin\theta_o}(\cos\phi\sin\theta - r\sin\phi\dot{\phi}\sin\theta + r\cos\phi\cos\theta\dot{\theta}),$$

$$\Rightarrow \qquad \left.\frac{dy_t}{dr}\right|_{\substack{\phi=\phi_o \\ \theta=\theta_o \\ r=r_o}} = -\frac{\alpha_i}{r_o \sin\theta_o}(\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o).$$

Now consider the ray of light. Its spatial coordinates are

$$(x_l, y_l, z_l) = (r\cos\phi\sin\theta, r\sin\phi\sin\theta, r\sin\theta),$$

where all quantities are functions of time. Therefore,

$$\frac{dy_l}{dr} = \sin\phi\sin\theta + r\cos\phi\dot{\phi}\sin\theta + r\sin\phi\cos\theta\dot{\theta}$$

$$\Rightarrow \qquad \left.\frac{dy_l}{dr}\right|_{\substack{\phi=\phi_o \\ \theta=\theta_o \\ r=r_o}} = r_o\dot{\phi}_o\sin\theta_o.$$

Equating the two y derivatives evaluated at the observer's position we have

$$r_o\dot{\phi}_o\sin\theta_o = -\frac{\alpha_i}{r_o\sin\theta_o}(\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o)$$

$$\Rightarrow \qquad \alpha_i = -\frac{r_o^2\sin^2\theta_o\dot{\phi}_o}{\sin\theta_o + r_o\cos\theta\dot{\theta}_o}. \tag{5.22}$$

Now we find β_i. For the tangent,

$$z_t = \frac{r_o\cos\theta_o\sin\theta_o - \beta_i}{r_o\sin^2\theta_o}r\cos\phi\sin\theta_o + \frac{\beta_i}{\sin\theta_o},$$

$$\Rightarrow \qquad \frac{dz_t}{dr} = \frac{r_o\cos\theta_o\sin\theta_o - \beta_i}{r_o\sin^2\theta_o}(\cos\phi\sin\theta - r\sin\phi\dot{\phi}\sin\theta + r\cos\phi\cos\theta\dot{\theta}),$$

$$\Rightarrow \qquad \left.\frac{dz_t}{dr}\right|_{\substack{\phi=\phi_o \\ \theta=\theta_o \\ r=r_o}} = \frac{r_o\cos\theta_o\sin\theta_o - \beta_i}{r_o\sin^2\theta_o}(\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o).$$

Whilst for the ray of light

$$z_l = r\cos\theta$$

$$\Rightarrow \qquad \frac{dz_l}{dr} = \cos\theta - r\sin\theta\dot{\theta}$$

$$\Rightarrow \qquad \left.\frac{dz_l}{dr}\right|_{\substack{\phi=\phi_o \\ \theta=\theta_o \\ r=r_o}} = \cos\theta_o - r_o\sin\theta_o\dot{\theta}_o.$$

Equating the two z derivatives evaluated at the observer's position we have

$$\cos\theta_o - r_o\sin\theta_o\dot{\theta}_o = \frac{r_o\cos\theta_o\sin\theta_o - \beta_i}{r_o\sin^2\theta_o}(\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o)$$

$$\Rightarrow \quad \beta_i = r_o\cos\theta_o\sin\theta_o + \frac{r_o\sin^2\theta_o(r_o\sin\theta_o\dot{\theta}_o - \cos\theta_o)}{\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o}$$

$$= \frac{r_o\cos\theta_o\sin\theta_o(\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o) + r_o\sin^2\theta_o(r_o\sin\theta_o\dot{\theta}_o - \cos\theta_o)}{\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o}$$

$$= \frac{r_o^2\sin\theta_o\dot{\theta}_o}{\sin\theta_o + r_o\cos\theta_o\dot{\theta}_o}. \qquad (5.23)$$

Now $\dot{\theta}$ and $\dot{\phi}$ are determined by equations (5.1) and (5.2) respectively. Since r_o is assumed to be sufficiently large that at this r_o the space is flat i.e., $m = 0$, to get approximate expressions for $\dot{\theta}_o$ and $\dot{\phi}_o$ we use equations (5.7) and (5.8) respectively. Thus,

$$\dot{\theta}_o = \left.\frac{d\theta}{dr}\right|_{\substack{\theta=\theta_o \\ r=r_o}} \approx \mp\frac{\sqrt{\Theta(\theta_o)}}{r_o^2\sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}},$$

where the upper sign corresponds to the $\theta_s \to \theta_{max} \to \theta_o$ path, whilst the lower sign corresponds to the $\theta_s \to \theta_{min} \to \theta_o$ path, and $\Theta(\theta)$ is defined by equation (5.5). Also,

$$\dot{\phi} = \frac{d\phi}{dr} = \frac{\partial\phi}{\partial r} + \frac{\partial\phi}{\partial\theta}\frac{d\theta}{dr} \approx \frac{\lambda}{\sin^2\theta\sqrt{\Theta(\theta)}}\frac{\sqrt{\Theta(\theta)}}{r^2\sqrt{1-\frac{\lambda^2+\eta}{r^2}}}$$

$$\Rightarrow \quad \dot{\phi}_o \approx \frac{\lambda \csc^2 \theta_o}{r_o^2 \sqrt{1 - \frac{\lambda^2+\eta}{r_o^2}}}.$$

Substituting $\dot{\theta}_o$ and $\dot{\phi}_o$ back into equation (5.22) and (5.23) we find that

$$\begin{aligned}
\alpha_i &\approx -\frac{r_o^2 \sin^2 \theta_o \frac{\lambda \csc^2 \theta_o}{r_o^2 \sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}}}{\sin \theta_o \mp r_o \cos \theta_o \frac{\sqrt{\Theta(\theta_o)}}{r_o^2 \sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}}} \\
&= \frac{-\lambda}{\sin \theta_o \sqrt{1 - \frac{\lambda^2+\eta}{r_o^2}} \mp \cos \theta_o \frac{\sqrt{\Theta(\theta_o)}}{r_o}} \qquad (5.24)
\end{aligned}$$

and

$$\begin{aligned}
\beta_i &\approx \frac{\mp r_o^2 \sin \theta_o \frac{\sqrt{\Theta(\theta_o)}}{r_o^2 \sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}}}{\sin \theta_o \mp r_o \cos \theta_o \frac{\sqrt{\Theta(\theta_o)}}{r_o^2 \sqrt{1-\frac{\lambda^2+\eta}{r_o^2}}}} \\
&= \frac{\mp \sin \theta_o \sqrt{\Theta(\theta_o)}}{\sin \theta_o \sqrt{1 - \frac{\lambda^2+\eta}{r_o^2}} \mp \cos \theta_o \frac{\sqrt{\Theta(\theta_o)}}{r_o}}. \qquad (5.25)
\end{aligned}$$

Since α_i and β_i are the coordinates of an image i.e., they are astronomically determined, we need to invert equations (5.24) and (5.25) to express λ and η as functions of α_i and β_i. Dividing (5.25) by (5.24) we get

$$\frac{\beta_i}{\alpha_i} \approx \pm \frac{\sin \theta_o \sqrt{\Theta(\theta_o)}}{\lambda}. \qquad (5.26)$$

Use this equation to substitute for $\sqrt{\Theta(\theta_o)}$ in equation (5.24) to find that

$$\alpha_i \approx \frac{-\lambda \alpha_i}{\alpha_i \sin \theta_o \sqrt{1 - \frac{\lambda^2+\eta}{r_o^2}} - \frac{\lambda \beta_i}{r_o} \cot \theta_o}$$

$$\Rightarrow \quad \lambda \approx \frac{\lambda\beta_i \cot\theta_o}{r_o} - \alpha_i \sin\theta_o \sqrt{1 - \frac{\lambda^2+\eta}{r_o^2}}$$

$$\Rightarrow \quad \eta \approx r_o^2 - \lambda^2 - \frac{\lambda^2(r_o - \beta_i \cot\theta_o)^2}{\alpha_i^2 \sin^2\theta_o}.$$

Expanding $\sqrt{\Theta(\theta_o)}$ in (5.26) and squaring, we have

$$\frac{\beta_i^2}{\alpha_i^2} \approx \frac{\eta \sin^2\theta_o - \lambda^2\cos^2\theta_o}{\lambda^2}$$

$$= \frac{r_o^2 \sin^2\theta_o}{\lambda^2} - \sin^2\theta_o - \frac{(r_o - \beta_i\cot\theta_o)^2}{\alpha_i^2} - \cos^2\theta_o$$

$$\Rightarrow \quad \lambda^2 \approx \frac{r_o^2\alpha_i^2 \sin^2\theta_o}{\beta_i^2 + \alpha_i^2 + (r_o - \beta_i\cot\theta_o)^2},$$

with equation (5.24),

$$\Rightarrow \quad \lambda \approx \frac{-r_o\alpha_i \sin\theta_o}{\sqrt{\beta_i^2 + \alpha_i^2 + (r_o - \beta_i\cot\theta_o)^2}}. \tag{5.27}$$

Substituting this back into (5.26) and rearranging for η

$$\Rightarrow \quad \sqrt{\eta} \approx \frac{\mp r_o\sqrt{\beta_i^2 + \alpha_i^2\cos^2\theta_o}}{\sqrt{\beta_i^2 + \alpha_i^2 + (r_o - \beta_i\cot\theta_o)^2}}, \tag{5.28}$$

where the $\mp$ is the sign of $\dot{\theta}_o$. If it is negative at r_o then θ must be decreasing which implies that the ray followed the $\theta_s \rightarrow \theta_{max} \rightarrow \theta_o$ path, and if positive then the ray must have followed the $\theta_s \rightarrow \theta_{min} \rightarrow \theta_o$ path.

Note that in the above derivations for λ and η the $\approx$ sign means $=$ excluding first and higher order terms in mr_{min}^2/r_o^3 i.e., these equations are exact for $m = a = 0$.

Thus, as $r_o \rightarrow \infty$ for $\theta_o = \pi/2$, $\lambda = -\alpha_i$ and $\sqrt{\eta} = \mp\beta_i$. So a source with constants of the motion λ and η has an image seen at $(-\lambda, \mp\sqrt{\eta})$. Since $\sqrt{\eta}$ is

positive, $\sqrt{\eta} = \mp\beta_i$ tells us that the sign of β_i determines which path the ray of light takes. If β_i is negative then we must choose the upper negative sign which corresponds to the $\theta_s \rightarrow \theta_{max} \rightarrow \theta_o$ path. On the other hand if β_i is positive then we must choose the lower positive sign which corresponds to the $\theta_s \rightarrow \theta_{min} \rightarrow \theta_o$ path. This makes good physical sense, for it means that if the image is seen below the equatorial plane then θ must have reached θ_{max}, whilst if the image is seen above the equatorial plane then θ must have reached θ_{min}.

§5.4 Solution for θ

Now we solve the equations of motion (5.1) and (5.2) without assuming m or a to be zero. However, we do assume that the ratio m/r_{min} is small. As $0 < |a| < m$, a/r_{min} is small also. In subsequent calculations approximations are made which exclude third and higher orders in these ratios. The reason that we keep the second order terms is that the spin parameter comes in only at this order i.e., the ratio a/r_{min} appears as a^2/r_{min}^2 only. We have seen already that $r_{min} = \sqrt{\lambda^2 + \eta}$ for $m = 0$. Now that we suppose that m/r_{min} is small, it is reasonable to write $r_{min} = \sqrt{\lambda^2 + \eta}(1 + O(\frac{m}{\sqrt{\lambda^2+\eta}}) + ...)$. This we show later, but for now, it is useful to use the fact that m/r_{min} is of the same order as $m/\sqrt{\lambda^2 + \eta}$, and that a^2/r_{min}^2 is of the same order as $a^2/(\lambda^2 + \eta)$.

In order to evaluate the right hand side of equation (5.1) we need to determine θ_{min} and θ_{max}. These angles come from the solution of $\Theta(\theta) = 0$. So setting equation (5.5) to zero we have

$$0 = \eta + a^2 \cos^2 \theta_{min/max} - \lambda^2 \cot^2 \theta_{min/max},$$

let $t_m = \cos \theta_{min/max}$,

$$= \eta + a^2 t_m^2 - \frac{\lambda^2 t_m^2}{1 - t_m^2}$$

$$= \frac{\eta(1 - t_m^2) + a^2 t_m^2 (1 - t_m^2) - \lambda^2 t_m^2}{1 - t_m^2}$$

$$\Rightarrow \quad 0 = a^2 t_m^4 + (\lambda^2 + \eta - a^2) t_m^2 - \eta.$$

Therefore,

$$t_m^2 = \frac{-(\lambda^2 + \eta - a^2) \pm ((\lambda^2 + \eta - a^2)^2 + 4a^2\eta)^{1/2}}{2a^2}$$

$$= \frac{\lambda^2+\eta-a^2}{2a^2}\left(-1 \pm \left(1+\frac{4a^2\eta}{(\lambda^2+\eta-a^2)^2}\right)^{1/2}\right)$$

$$\approx \frac{\lambda^2+\eta-a^2}{2a^2}\left(-1 \pm \left(1+\frac{2a^2\eta}{(\lambda^2+\eta-a^2)^2}-\frac{2a^4\eta^2}{(\lambda^2+\eta-a^2)^4}\right)\right)$$

$$\approx \frac{\lambda^2+\eta-a^2}{2a^2}\left(-1 \pm \left(1+\frac{2a^2\eta}{(\lambda^2+\eta)^2}\left(1+\frac{2a^2}{\lambda^2+\eta}\right)-\frac{2a^4\eta^2}{(\lambda^2+\eta)^4}\right)\right),$$

as $t_m^2 > 0$, take the positive root,

$$= \frac{\eta(\lambda^2+\eta-a^2)}{(\lambda^2+\eta)^2}\left(1+\frac{2a^2}{\lambda^2+\eta}-\frac{a^2\eta}{(\lambda^2+\eta)^2}\right)$$

$$\approx \frac{\eta}{\lambda^2+\eta}\left(1+\frac{2a^2}{\lambda^2+\eta}-\frac{a^2\eta}{(\lambda^2+\eta)^2}-\frac{a^2}{\lambda^2+\eta}\right)$$

$$= \frac{\eta}{\lambda^2+\eta}\left(1+\frac{a^2(\lambda^2+\eta)}{(\lambda^2+\eta)^2}-\frac{a^2\eta}{(\lambda^2+\eta)^2}\right)$$

$$= \frac{\eta}{\lambda^2+\eta}\left(1+\frac{a^2\lambda^2}{(\lambda^2+\eta)^2}\right).$$

Thus, we have that

$$\cos\theta_{min/max} \approx \pm\sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right), \tag{5.29}$$

where the positive sign corresponds to $\cos\theta_{min}$ and the negative sign corresponds to $\cos\theta_{max}$.

Next we evaluate the right hand side of (5.1). As stated earlier, if the ray takes a path that reaches θ_{max} then

$$\int^{\theta} \frac{d\theta}{\pm[\Theta(\theta)]^{1/2}} = \int_{\theta_s}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}} + \int_{\theta_o}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}},$$

otherwise

$$= \int_{\theta_{min}}^{\theta_s} \frac{d\theta}{[\Theta(\theta)]^{1/2}} + \int_{\theta_{min}}^{\theta_o} \frac{d\theta}{[\Theta(\theta)]^{1/2}}.$$

Consider

$$\int_{\theta_s}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}} = \int_{\theta_s}^{\theta_{max}} \frac{d\theta}{(\eta + a^2\cos^2\theta - \lambda^2\cot^2\theta)^{1/2}}. \tag{5.30}$$

Let

$$\cos\theta = \sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1 + \frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right)\cos\alpha.$$

Therefore,

$$\cos^2\theta \approx \frac{\eta}{\lambda^2+\eta}\left(1 + \frac{a^2\lambda^2}{(\lambda^2+\eta)^2}\right)\cos^2\alpha$$

and,

$$-\sin\theta\, d\theta = -\sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1 + \frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right)\sin\alpha\, d\alpha.$$

For notational convenience define

$$k = \sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1 + \frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right). \tag{5.31}$$

Equation (5.30) becomes

$$k\int_{\alpha_s}^{\pi} \frac{\sin\alpha\, d\alpha}{\sin\theta(\eta + a^2\cos^2\theta - \lambda^2\cot^2\theta)^{1/2}}$$

$$= k\int_{\alpha_s}^{\pi} \frac{\sin\alpha\, d\alpha}{(\eta\sin^2\theta + a^2\cos^2\theta\sin^2\theta - \lambda^2\cos^2\theta)^{1/2}}$$

$$= k \int_{\alpha_s}^{\pi} \frac{\sin\alpha \, d\alpha}{(\eta(1-\cos^2\theta) + a^2\cos^2\theta(1-\cos^2\theta) - \lambda^2\cos^2\theta)^{1/2}}$$

$$= k \int_{\alpha_s}^{\pi} \frac{\sin\alpha \, d\alpha}{(\eta - (\lambda^2+\eta-a^2)\cos^2\theta - a^2\cos^4\theta)^{1/2}}$$

$$\approx k \int_{\alpha_s}^{\pi} \frac{\sin\alpha \, d\alpha}{(\eta - (\lambda^2+\eta-a^2)\frac{\eta}{\lambda^2+\eta}(1+\frac{a^2\lambda^2}{(\lambda^2+\eta)^2})\cos^2\alpha - \frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^4\alpha)^{1/2}}$$

$$\approx k \int_{\alpha_s}^{\pi} \frac{\sin\alpha \, d\alpha}{(\eta - \eta\cos^2\alpha - \frac{\eta a^2\lambda^2}{(\lambda^2+\eta)^2}\cos^2\alpha + \frac{a^2\eta}{\lambda^2+\eta} - \frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^4\alpha)^{1/2}}$$

$$= k \int_{\alpha_s}^{\pi} \frac{\sin\alpha \, d\alpha}{(\eta\sin^2\alpha + \frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^2\alpha(1-\cos^2\alpha))^{1/2}}$$

$$= k \int_{\alpha_s}^{\pi} \frac{d\alpha}{(\eta + \frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^2\alpha)^{1/2}}.$$

Now, this can be immediately integrated, however, the result mixes $\sin\alpha$ with α which proves to be inconvenient to work with. So we rearrange the equation further.

$$\approx \frac{k}{\sqrt{\eta}} \int_{\alpha_s}^{\pi} \frac{d\alpha}{1 + \frac{a^2\eta}{2(\lambda^2+\eta)^2}\cos^2\alpha}$$

$$= \frac{-k}{\sqrt{\eta}\sqrt{1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}}} \arctan\left(\sqrt{1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}}\cot\alpha\right)\Big]_{\alpha_s}^{\pi}$$

$$\approx \frac{k}{\sqrt{\eta}}\left(1 - \frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\left(\frac{\pi}{2} + \arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right)\right).$$

Now,

$$\frac{k}{\sqrt{\eta}}\left(1 - \frac{a^2\eta}{4(\lambda^2+\eta)^2}\right) = \frac{1}{\sqrt{\lambda^2+\eta}}\left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right)\left(1-\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta}}\left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}-\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta}}\left(1+\frac{a^2(\lambda^2+\eta-\eta-\eta/2)}{2(\lambda^2+\eta)^2}\right)$$

$$= \frac{1}{\sqrt{\lambda^2+\eta}}\left(1+\frac{a^2}{2(\lambda^2+\eta)}\left(1-\frac{3\eta}{2(\lambda^2+\eta)}\right)\right)$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right).$$

Thus,

$$\int_{\theta_s}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right)$$
$$\times\left(\frac{\pi}{2}+\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right)\right).$$

Similarly,

$$\int_{\theta_{min}}^{\theta_s} \frac{d\theta}{[\Theta(\theta)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right)$$
$$\times\left(\frac{\pi}{2}-\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right)\right).$$

So we will write

$$\int^{\theta} \frac{d\theta}{[\Theta(\theta)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right)$$

$$\times\left(\pi\pm\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right)\pm\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_o\right)\right), \quad (5.32)$$

where the plus sign corresponds to θ_s to θ_{max} path, whilst the minus sign corresponds to the θ_s to θ_{min} path.

Having evaluated the right hand side of equation (5.1) we now evaluate the left hand side. Firstly we evaluate r_{min}. We find it by solving $R(r) = 0$, where $R(r)$ was defined in equation (5.4). Now $R(r)$ is a quartic in r with r_{min} being the largest of the real roots. We find r_{min} by guessing its form, and then proving that $R(r) > 0$ for $r > r_{min}$, thus showing that it is the largest real root. Rewriting (5.4) in a more convenient form

$$R(r) = r^4 - (\lambda^2 + \eta - a^2)r^2 + 2mr((\lambda - a)^2 + \eta) - a^2\eta,$$

defining s by $r = s\sqrt{\lambda^2 + \eta}$, setting $R(r) = 0$, and dividing through by $(\lambda^2 + \eta)^2$ we have

$$0 = s^4 - \left(1 - \frac{a^2}{\lambda^2 + \eta}\right)s^2 + \frac{2m}{\sqrt{\lambda^2 + \eta}}s\left(1 - \frac{2a\lambda}{\lambda^2 + \eta} + \frac{a^2}{\lambda^2 + \eta}\right) - \frac{a^2\eta}{\lambda^2 + \eta}. \quad (5.33)$$

Guess s to have the form

$$s = c_1 + c_2\frac{m}{\sqrt{\lambda^2 + \eta}} + c_3\frac{m^2}{\lambda^2 + \eta} + c_4\frac{a^2}{\lambda^2 + \eta} + c_5\frac{a^2\eta}{(\lambda^2 + \eta)^2} + c_6\frac{am\lambda}{(\lambda^2 + \eta)^{3/2}}.$$

Therefore,

$$s^4 \approx c_1^4 + 4c_1^3\left(c_2\frac{m}{\sqrt{\lambda^2 + \eta}} + c_3\frac{m^2}{\lambda^2 + \eta} + c_4\frac{m}{\lambda^2 + \eta} + c_5\frac{a^2\eta}{(\lambda^2 + \eta)^2},\right.$$
$$\left. + c_6\frac{am\lambda}{(\lambda^2 + \eta)^{3/2}}\right) + 6c_1^2c_2^2\frac{m^2}{\lambda^2 + \eta}$$

$$s^2 \approx c_1^2 + 2c_1\left(c_2\frac{m}{\sqrt{\lambda^2 + \eta}} + c_3\frac{m^2}{\lambda^2 + \eta} + c_4\frac{m}{\lambda^2 + \eta} + c_5\frac{a^2\eta}{(\lambda^2 + \eta)^2}\right.$$
$$\left. + c_6\frac{am\lambda}{(\lambda^2 + \eta)^{3/2}}\right) + c_2^2\frac{m^2}{\lambda^2 + \eta}.$$

Substituting back into (5.33) we find that

$$0 \approx c_1^4 + 4c_1^3\left(c_2\frac{m}{\sqrt{\lambda^2+\eta}} + c_3\frac{m^2}{\lambda^2+\eta} + c_4\frac{m}{\lambda^2+\eta} + c_5\frac{a^2\eta}{(\lambda^2+\eta)^2}\right.$$

$$\left. + c_6\frac{am\lambda}{(\lambda^2+\eta)^{3/2}}\right) + 6c_1^2c_2^2\frac{m^2}{\lambda^2+\eta} - c_1^2 - 2c_1\left(c_2\frac{m}{\sqrt{\lambda^2+\eta}} + c_3\frac{m^2}{\lambda^2+\eta}\right.$$

$$\left. + c_4\frac{m}{\lambda^2+\eta} + c_5\frac{a^2\eta}{(\lambda^2+\eta)^2} + c_6\frac{am\lambda}{(\lambda^2+\eta)^{3/2}}\right) - c_2^2\frac{m^2}{\lambda^2+\eta} + c_1^2\frac{a^2}{\lambda^2+\eta}$$

$$+ \frac{2m}{\sqrt{\lambda^2+\eta}}\left(c_1 + c_2\frac{m}{\sqrt{\lambda^2+\eta}}\right)\left(1 - 2\frac{a\lambda}{\lambda^2+\eta}\right) - \frac{a^2\eta}{(\lambda^2+\eta)^2}.$$

Equating the coefficients we have

$$0 \approx c_1^4 - c_1^2 \quad \Rightarrow \quad c_1 = \pm 1.$$

Choose $c_1 = 1$ since we want the largest root.

$$0 \approx (4c_2 - 2c_2 + 2)\frac{m}{\sqrt{\lambda^2+\eta}} \quad \Rightarrow \quad c_2 = -1,$$

$$0 \approx (4c_3 + 6 - 2c_3 - 1 - 2)\frac{m^2}{\lambda^2+\eta} \quad \Rightarrow \quad c_3 = -\frac{3}{2},$$

$$0 \approx (4c_4 - 2c_4 + 1)\frac{a^2}{\lambda^2+\eta} \quad \Rightarrow \quad c_4 = -\frac{1}{2},$$

$$0 \approx (4c_5 - 2c_5 - 1)\frac{a^2\eta}{(\lambda^2+\eta)^2} \quad \Rightarrow \quad c_5 = \frac{1}{2},$$

$$0 \approx (4c_6 - 2c_6 - 4)\frac{am\lambda}{(\lambda^2+\eta)^{3/2}} \quad \Rightarrow \quad c_6 = 2.$$

Thus, we have found that

$$r_{min} \approx \sqrt{\lambda^2+\eta}\left(1-\frac{m}{\sqrt{\lambda^2+\eta}}-\frac{3m^2}{2(\lambda^2+\eta)}-\frac{a^2}{2(\lambda^2+\eta)}+\frac{a^2\eta}{2(\lambda^2+\eta)^2}\right.$$

$$\left.+\frac{2am\lambda}{(\lambda^2+\eta)^{3/2}}\right), \quad (5.34)$$

as we suggested earlier. It is easily seen that r_{min} is the largest of the real roots since $R(r)$ can now be written as

$$R(r) = (r-r_{min})((r+r_{min})(r^2+r_{min}^2+a^2-\lambda^2-\eta)+2m((a-\lambda)^2+\eta)),$$

which is greater than zero for r greater than r_{min}.

Something rather unexpected happens for the extreme Kerr black hole in the case where the ray stays in the equatorial plane. Substituting $a=m$ and $\eta=0$ into equation (5.34) we find that for $\lambda=|\lambda|$, $r_{min} \approx \lambda - m$. In fact $r_{min} = \lambda - m$ as it is easily verified that under these conditions $R(\lambda-m)=0$. This is surprising as in the simpler Schwarzschild case no such simple solution exists.

Now we are in a position to evaluate the left hand side of (5.1), which is

$$\int^{r} \frac{dr}{\pm[R(r)]^{1/2}} = \int_{r_{min}}^{r_s} \frac{dr}{[R(r)]^{1/2}} + \int_{r_{min}}^{r_o} \frac{dr}{[R(r)]^{1/2}}.$$

In most astronomical calculations one would expect that r_s and r_o are very much greater than r_{min} and so in the following calculations we only carry terms up to and including second order terms in r_{min}/r_s and r_{min}/r_o. Recall also that r_{min} is of the same order as $\sqrt{\lambda^2+\eta}$, which proves to be more convenient to use than r_{min} in the following calculations. Thus,

$$\int^{r} \frac{dr}{\pm[R(r)]^{1/2}} = 2\int_{r_{min}}^{\infty} \frac{dr}{[R(r)]^{1/2}} - \int_{r_s}^{\infty} \frac{dr}{[R(r)]^{1/2}} - \int_{r_o}^{\infty} \frac{dr}{[R(r)]^{1/2}}$$

$$\approx 2\int_{r_{min}}^{\infty} \frac{dr}{[R(r)]^{1/2}} - \frac{1}{r_s} - \frac{1}{r_o}. \quad (5.35)$$

Consider

$$R(r) = r^4 - (\lambda^2 + \eta - a^2)r^2 + 2mr((\lambda - a)^2 + \eta) - a^2\eta.$$

Let $r = 1/u$, then

$$R(u) = \frac{1}{u^4}\left(1 - (\lambda^2 + \eta - a^2)u^2 + 2mu^3((\lambda - a)^2 + \eta) - a^2\eta u^4\right)$$

$$= \frac{1}{u^4}\Bigg(- (\lambda^2 + \eta - a^2)(u^2 - u_{max}^2) + 2m(u^3 - u_{max}^3)((\lambda - a)^2 + \eta)$$

$$- a^2\eta(u^4 - u_{max}^4)\Bigg),$$

where $u_{max} = 1/r_{min}$,

$$= \frac{1}{u^4}(u^2 - u_{max}^2)\Bigg[- a^2\eta(u^2 + u_{max}^2) + a^2 - \lambda^2 - \eta$$

$$+ 2m((\lambda - a)^2 + \eta)\left(\frac{u^2 + uu_{max} + u_{max}^2}{u + u_{max}}\right)\Bigg]$$

$$= \frac{u_{max}^2}{u^4}(\lambda^2 + \eta - a^2)(1 - x^2)\Bigg[1 + \frac{a^2\eta u_{max}^2(1 + x^2)}{\lambda^2 + \eta - a^2}$$

$$- 2mu_{max}\left(\frac{(\lambda - a)^2 + \eta}{\lambda^2 + \eta - a^2}\right)\left(\frac{x^2 + x + 1}{x + 1}\right)\Bigg],$$

where we defined x by $u = u_{max}x$. Define $f(x)$ and $g(x)$ by

$$f(x) = \frac{\eta(1 + x^2)}{\lambda^2 + \eta - a^2} \quad \text{and} \quad g(x) = -2\left(\frac{(\lambda - a)^2 + \eta}{\lambda^2 + \eta - a^2}\right)\left(\frac{x^2 + x + 1}{x + 1}\right),$$

then since $r = 1/u \quad \Rightarrow \quad dr = -1/u^2 du$,

$$2\int_{r_{min}}^{\infty} \frac{dr}{[R(r)]^{1/2}} = \frac{2}{\sqrt{\lambda^2+\eta-a^2}} \int_0^1 \frac{dx}{((1-x^2)(1+a^2u_{max}^2 f(x)+mu_{max}g(x))^{1/2}},$$

$$\approx \frac{2}{\sqrt{\lambda^2+\eta-a^2}} \int_0^1 \frac{dx}{\sqrt{1-x^2}} \left(1 - \frac{1}{2}a^2u_{max}^2 f(x) - \frac{1}{2}mu_{max}g(x) + \frac{3}{8}m^2u_{max}^2 g^2(x)\right)$$

$$\approx \frac{2}{\sqrt{\lambda^2+\eta-a^2}} \int_0^1 \frac{dx}{\sqrt{1-x^2}} (1 - \frac{\eta}{2(\lambda^2+\eta)}a^2u_{max}^2(1+x^2) + mu_{max}\left(\frac{(\lambda-a)^2+\eta}{\lambda^2+\eta}\right)\left(\frac{x^2+x+1}{x+1}\right) + \frac{3}{2}m^2u_{max}^2\left(\frac{x^2+x+1}{x+1}\right)^2.$$

Now,

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}} = \frac{\pi}{2}, \quad \int_0^1 \frac{xdx}{\sqrt{1-x^2}} = 1, \quad \int_0^1 \frac{x^2dx}{\sqrt{1-x^2}} = \frac{\pi}{4},$$

$$\int_0^1 \frac{xdx}{\sqrt{1-x^2}(1+x)} = \frac{\pi}{2} - 1, \quad \int_0^1 \frac{x^2dx}{\sqrt{1-x^2}(1+x)} = 2 - \frac{\pi}{2},$$

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}(1+x)^2} = \frac{2}{3}, \quad \int_0^1 \frac{xdx}{\sqrt{1-x^2}(1+x)^2} = \frac{1}{3}.$$

Therefore,

$$\int_0^1 \frac{dx}{\sqrt{1-x^2}}\left(\frac{x^2+x+1}{x+1}\right) = 2 \quad \text{and} \quad \int_0^1 \frac{dx}{\sqrt{1-x^2}}\left(\frac{x^2+x+1}{x+1}\right)^2 = \frac{5\pi}{4} - \frac{4}{3}.$$

So,

$$2\int_{r_{min}}^{\infty} \frac{dr}{[R(r)]^{1/2}} \approx \frac{2}{\sqrt{\lambda^2+\eta-a^2}} \left(\frac{\pi}{2}(1 - \frac{\eta}{2(\lambda^2+\eta)}a^2u_{max}^2(1+\frac{1}{2})) + 2mu_{max}\right.$$

$$- 4mu_{max}\frac{a\lambda}{\lambda^2+\eta} + \frac{3}{2}m^2u_{max}^2\left(\frac{5\pi}{4} - \frac{4}{3}\right)\Bigg)$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\Bigg(\pi\left(1 - \frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right) + \frac{4m}{\sqrt{\lambda^2+\eta}}\left(1 + \frac{m}{\sqrt{\lambda^2+\eta}}\right)$$

$$- \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}} + 3\left(\frac{5\pi}{4} - \frac{4}{3}\right)\frac{m^2}{\lambda^2+\eta}\Bigg)$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\Bigg(\pi\left(1 - \frac{3a^2\eta}{4(\lambda^2+\eta)^2} + \frac{15m^2}{4(\lambda^2+\eta)}\right)$$

$$+ \frac{4m}{\sqrt{\lambda^2+\eta}} - \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}\Bigg).$$

Thus, with the assistance of equation (5.35) we have that

$$\int_{\infty}^{r}\frac{dr}{\pm[R(r)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\Bigg(\pi\left(1 - \frac{3a^2\eta}{4(\lambda^2+\eta)^2} + \frac{15m^2}{4(\lambda^2+\eta)}\right)$$

$$+ \frac{4m}{\sqrt{\lambda^2+\eta}} - \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}\Bigg) - \frac{1}{r_s} - \frac{1}{r_o}. \quad (5.36)$$

Equating equations (5.32) and (5.36) we have

$$\frac{1}{\sqrt{\lambda^2+\eta-a^2}}\Bigg(\pi\left(1 - \frac{3a^2\eta}{4(\lambda^2+\eta)^2} + \frac{15m^2}{4(\lambda^2+\eta)}\right)$$

$$+ \frac{4m}{\sqrt{\lambda^2+\eta}} - \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}\Bigg) - \frac{1}{r_s} - \frac{1}{r_o}$$

$$\approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1 - \frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right)$$

$$\times\left(\pi \pm \arctan\left(\left(1 + \frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right) \pm \arctan\left(\left(1 + \frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_o\right)\right).$$

Therefore,

$$\frac{15\pi}{4}\frac{m^2}{(\lambda^2+\eta)}+\frac{4m}{\sqrt{\lambda^2+\eta}}-\frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}-\sqrt{\lambda^2+\eta}\left(\frac{1}{r_s}+\frac{1}{r_o}\right)$$

$$\approx \pm\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right)\pm\arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_o\right),$$

taking the tangent of both sides,

$$\approx \pm\frac{\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)(\cot\alpha_s+\cot\alpha_o)}{1-\left(1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}\right)\cot\alpha_s\cot\alpha_o}$$

$$\approx \pm\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\left(\frac{\cot\alpha_s+\cot\alpha_o}{1-\cot\alpha_s\cot\alpha_o}\right)\left(1-\frac{a^2\eta}{2(\lambda^2+\eta)^2}\frac{\cot\alpha_s+\cot\alpha_o}{1-\cot\alpha_s\cot\alpha_o}\right),$$

since the left hand side is of first order then so is $\frac{\cot\alpha_s+\cot\alpha_o}{1-\cot\alpha_s\cot\alpha_o}$,

$$\approx \pm\frac{\cot\alpha_s+\cot\alpha_o}{1-\cot\alpha_s\cot\alpha_o}=\pm\tan(\pi-\alpha_s-\alpha_o),$$

taking the arctangent of both sides,

$$\approx \pi-\alpha_s-\alpha_o+\begin{cases}\pi\\0\\-\pi\end{cases},$$

where π is added if $\pi-\alpha_s-\alpha_o<-\pi/2$ and subtracted if $\pi-\alpha_s-\alpha_o>\pi/2$.
Defining δ by

$$\delta=\frac{15\pi}{4}\frac{m^2}{(\lambda^2+\eta)}+\frac{4m}{\sqrt{\lambda^2+\eta}}-\frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}-\sqrt{\lambda^2+\eta}\frac{(r_s+r_o)}{r_or_s},$$

from above we have

$$\pm\delta\approx\pi-\alpha_s-\alpha_o+\begin{cases}\pi\\0\\-\pi\end{cases}.$$

This equation is very similar to the one we had for $m = 0$ case. Then δ was negative which implied that $\pi - \alpha_s - \alpha_o$ was always in the range $[-\pi/2, \pi/2]$. Now however δ may be positive, negative or zero, the only restriction being that $|\delta| < 1$. At first it may seem surprising that δ may be zero, as then $\theta_o = \pi/2 \Rightarrow \theta_s = \pi/2$. This however does not mean that no deviation took place, but rather corresponds to the freak case where the ray of light starts at the source which is in the equatorial plane of the black hole, then leaves this plane to reach θ_{min} or θ_{max} i.e., goes above or below the plane respectively, and finally arrives back at the equatorial plane at the observer's coordinates.

The above equation essentially defines α_s as a function of α_o. We expect this function to be a continuous function of α_o, so it is surprising to find that π may suddenly be added or subtracted. This needs further investigation.

Recall that in the case $\delta < 0$ we proved that $\cot\alpha_s \cot\alpha_o < \delta^2$. This means that either $\alpha_s \in [0, \pi/2]$ and $\alpha_o \in [\pi/2, \pi]$ (or vice versa), or both α_s and α_o lie on the same side of $\pi/2$. In the former case it is clear that $-\pi/2 \leq \pi - \alpha_s - \alpha_o \leq \pi/2$ so no discontinuities can arise. In the latter case $0 \leq \cot\alpha_s \cot\alpha_o < \delta^2$ thus implying that either both α_s and α_o are just greater than $\pi/2$ or both are just less than $\pi/2$. In either case $-\pi/2 \leq \pi - \alpha_s - \alpha_o \leq \pi/2$ so again no discontinuities can arise. Now for the $\delta \geq 0$ case.

Without loss of generality we assume that the ray took the $\theta_s \to \theta_{max} \to \theta_o$ path, so α ranges $\alpha_s \to \pi \to \alpha_o$. So we have $\delta = |\delta| \approx \tan(\pi - \alpha_s - \alpha_o)$. Now, as along this path α reaches π, the maximum value that α_o can take is also π. Consider what happens to α_s as we increase α_o from say $\pi - 2\delta$ through to π. Now for $\alpha_o = \pi - 2\delta$ we have $\delta \approx \tan(2\delta - \alpha_s) \Rightarrow \alpha_s \approx \delta$. For $\alpha_o = \pi - \delta$ we have $\delta \approx \tan(\delta - \alpha_s) \Rightarrow \alpha_s \approx 0$. Whilst for $\alpha_o = \pi$ we have $\delta \approx \tan(-\alpha_s) \Rightarrow \alpha_s \approx \pi - \delta$. Clearly there is a discontinuity occuring here. However this is unphysical, and therefore some explanation is required.

Recall that as the ray of light travels $r_s \to r_{min} \to r_o$, we assumed that we have only two possible paths namely the $\theta_s \to \theta_{max} \to \theta_o$ path and the $\theta_s \to \theta_{min} \to \theta_o$ path, as opposed to say $\theta_s \to \theta_o$ i.e., θ_{max} or θ_{min} not attained, or the case where $\theta_s \to \theta_{max} \to \theta_{min} \to \theta_o$, etc. The correct path is determined by the value of the integral on the left hand side of (5.1). The above discontinuity

occurs when instead of the assumed $\theta_s \rightarrow \theta_{max} \rightarrow \theta_o$ path, the real path is $\theta_s \rightarrow \theta_{max} \rightarrow \theta_{min} \rightarrow \theta_o$. So what happens is that having evaluated the left hand side of (5.1) to get K, we have

$$K = \int_{\theta_s}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}} + \int_{\theta_o}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}},$$

where $\alpha_o \leq \pi - \delta$. As $\alpha_o \rightarrow \pi$ i.e., $\theta_o \rightarrow \theta_{max}$, the integral of $[\Theta(\theta)]^{-1/2}$ from θ_o to θ_{max} gets smaller and smaller which implies that θ_s also gets smaller until it reaches θ_{min}, and then it starts getting bigger again. It is at this point that the above equation becomes

$$K = \int_{\theta_{min}}^{\theta_s} \frac{d\theta}{[\Theta(\theta)]^{1/2}} + \int_{\theta_{min}}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}} + \int_{\theta_o}^{\theta_{max}} \frac{d\theta}{[\Theta(\theta)]^{1/2}}.$$

This new contribution is exactly π to the equation $\delta \approx \tan(\pi - \alpha_s - \alpha_o)$ i.e., for the case where $\alpha_o > \pi - \delta$ we have $\delta \approx \tan(2\pi - \alpha_s - \alpha_o)$. So now as $\alpha_o \rightarrow \pi - 2\delta \rightarrow \pi - \delta \rightarrow \pi$, $\alpha_s \rightarrow \delta \rightarrow 0 \rightarrow \delta$ respectively. Thus no discontinuities arise.

The argument for the $\theta_s \rightarrow \theta_{min} \rightarrow \theta_s$ path is similar. Since it would be fairly rare for θ_o to be extremely close to θ_{max} or θ_{min} we shall ignore this case and simply assume that $\alpha_o \in [|\delta|, \pi - |\delta|]$.

Thus, for $\alpha_o \in [|\delta|, \pi - |\delta|]$

$$\alpha_s \approx \pi - \alpha_o \mp \delta \tag{5.37}$$

$$\Rightarrow \cos\alpha_s \approx -\cos\alpha_o \cos\delta \pm \sin\alpha_o \sin\delta$$

$$\Rightarrow \cos\theta_s \approx -\cos\theta_o \cos\delta \pm \sin\delta \sqrt{\frac{\eta}{\lambda^2 + \eta} - \cos^2\theta_o}, \tag{5.38}$$

where

$$\delta = \frac{15\pi}{4}\frac{m^2}{(\lambda^2+\eta)} + \frac{4m}{\sqrt{\lambda^2+\eta}} - \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}} - \sqrt{\lambda^2+\eta}\frac{(r_s+r_o)}{r_o r_s}, \qquad (5.39)$$

and as before the plus sign corresponds to the path that reaches θ_{max} and the minus sign corresponds to the path that reaches θ_{min}. Note that other than for the definition of δ, this result is identical to that for the $m = 0$ case.

§5.5 Solution for ϕ

Having evaluated θ_s we are now in a position to evaluate ϕ_s. Equation (5.2) can be rewritten as

$$\phi = \phi_o - \phi_s = \int^{r} \frac{a(2mr - a\lambda)}{\pm\Delta[R(r)]^{1/2}} dr + \int^{\theta} \frac{\lambda d\theta}{\pm \sin^2\theta [\Theta(\theta)]^{1/2}}. \tag{5.40}$$

Consider the first integral of (5.40). As in the evaluation of $\int \frac{dr}{[R(r)]^{1/2}}$, let $u = 1/r$ and define x by $u = u_{max}x$.

$$\begin{aligned}
\int^{r} \frac{a(2mr - a\lambda)}{\pm\Delta[R(r)]^{1/2}} dr &= \int_{r_{min}}^{r_s} \frac{a(2mr - a\lambda)}{\Delta[R(r)]^{1/2}} dr + \int_{r_{min}}^{r_o} \frac{a(2mr - a\lambda)}{\Delta[R(r)]^{1/2}} dr \\
&\approx 2 \int_{r_{min}}^{\infty} \frac{a(2mr - a\lambda)}{\Delta[R(r)]^{1/2}} dr \\
&= \frac{2a}{\sqrt{\lambda^2 + \eta - a^2}} \\
&\times \int_0^1 \frac{(2mr - a\lambda)dx}{(r^2 - 2mr + a^2)((1 - x^2)(1 + \frac{a^2}{r_{min}^2} f(x) + \frac{m}{r_{min}} g(x))^{1/2}},
\end{aligned}$$

divide through by r^2, substitute $x = r_{min}/r$ and drop third and higher order terms,

$$\begin{aligned}
&\approx \frac{2\lambda}{\sqrt{\lambda^2 + \eta}} \int_0^1 \frac{dx}{\sqrt{1 - x^2}} \left(2x \frac{ma}{r_{min}\lambda} - x^2 \frac{a^2}{r_{min}^2} \right) \\
&= \frac{2\lambda}{\sqrt{\lambda^2 + \eta}} \left(\frac{2ma}{r_{min}\lambda} - \frac{\pi}{4} \frac{a^2}{r_{min}^2} \right) \\
&\approx \frac{4ma}{\lambda^2 + \eta} - \frac{\pi}{2} \frac{\lambda a^2}{(\lambda^2 + \eta)^{3/2}}.
\end{aligned} \tag{5.41}$$

Now consider the second integral of (5.40). As in the evaluation of $\int \frac{d\theta}{[\Theta(\theta)]^{1/2}}$

we make the substitution $\cos\theta = k\cos\alpha$, where $k = \sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right)$. Then

$$\int^{\theta} \frac{\lambda d\theta}{\sin^2\theta[\Theta(\theta)]^{1/2}} \approx k\lambda \int^{\alpha} \frac{d\alpha}{(1-k^2\cos^2\alpha)\sqrt{\eta+\frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^2\alpha}}$$

$$\approx \frac{k\lambda}{\sqrt{\eta}} \int^{\alpha} \frac{d\alpha}{(1-k^2\cos^2\alpha)(1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}\cos^2\alpha)}$$

$$= \frac{k\lambda}{\sqrt{\eta}(k^2+\epsilon^2)} \int^{\alpha} d\alpha\left(\frac{k^2}{1-k^2\cos^2\alpha}+\frac{\epsilon^2}{1+\epsilon^2\cos^2\alpha}\right),$$

where $\epsilon^2 = \frac{a^2\eta}{2(\lambda^2+\eta)^2}$,

$$= -\frac{k\lambda}{\sqrt{\eta}(k^2+\epsilon^2)}\left(\frac{k^2}{\sqrt{1-k^2}}\arctan(\sqrt{1-k^2}\cot\alpha)\right.$$

$$\left.+\frac{\epsilon^2}{\sqrt{1+\epsilon^2}}\arctan(\sqrt{1+\epsilon^2}\cot\alpha)\right)\Bigg]^{\alpha}$$

$$= \frac{k\lambda}{\sqrt{\eta}(k^2+\epsilon^2)}\left(\frac{k^2}{\sqrt{1-k^2}}(\pi \pm \arctan(\sqrt{1-k^2}\cot\alpha_s)\right.$$

$$\left.\pm\arctan(\sqrt{1-k^2}\cot\alpha_o)\right)$$

$$+\frac{\epsilon^2}{\sqrt{1+\epsilon^2}}(\pi \pm \arctan(\sqrt{1+\epsilon^2}\cot\alpha_s)$$

$$\left.\pm\arctan(\sqrt{1+\epsilon^2}\cot\alpha_o)\right)\Big).$$

Now consider $(1-k^2)\cot\alpha_s\cot\alpha_o$ and $(1+\epsilon^2)\cot\alpha_s\cot\alpha_o$. We want both of these to be less than one because then we can combine the corresponding arctans to get

$$\int^{\theta} \frac{\lambda d\theta}{\sin^2\theta[\Theta(\theta)]^{1/2}}$$

$$\approx \frac{k\lambda}{\sqrt{\eta}(k^2+\epsilon^2)}\left(\frac{k^2}{\sqrt{1-k^2}}(\pi \pm \arctan(\sqrt{1-k^2}\frac{\cot\alpha_s+\cot\alpha_o}{1-(1-k^2)\cot\alpha_s\cot\alpha_o}))\right.$$

$$\left. + \frac{\epsilon^2}{\sqrt{1+\epsilon^2}}(\pi \pm \arctan(\sqrt{1+\epsilon^2}\frac{\cot\alpha_s+\cot\alpha_o}{1-(1+\epsilon^2)\cot\alpha_s\cot\alpha_o}))\right).$$

To prove this we recall the results of §5.2. Namely that

$$\pm\delta \approx \pi - \alpha_s - \alpha_o \quad \text{with} \quad |\delta| < 1 \quad \Rightarrow \quad \cot\alpha_s \approx \frac{\pm\delta - \cot\alpha_o}{1 \pm \delta\cot\alpha_o},$$

and that therefore

$$\cot\alpha_s\cot\alpha_o < \delta^2 < 1 \quad \forall \quad \alpha_o \in [|\delta|, \pi - |\delta|].$$

Since by equation (5.39)

$$\delta = \frac{15\pi}{4}\frac{m^2}{(\lambda^2+\eta)} + \frac{4m}{\sqrt{\lambda^2+\eta}} - \frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}} - \sqrt{\lambda^2+\eta}\left(\frac{1}{r_s}+\frac{1}{r_o}\right),$$

$$1-k^2 = \frac{\lambda^2}{\lambda^2+\eta}\left(1-\frac{a^2\eta}{(\lambda^2+\eta)^2}\right)$$

and

$$1+\epsilon^2 = 1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}$$

it is clear that indeed $(1-k^2)\cot\alpha_s\cot\alpha_o < 1$ and $(1+\epsilon^2)\cot\alpha_s\cot\alpha_o < 1$. Hence the above expression holds. Also in the same section we showed that

$$\frac{\cot\alpha_s+\cot\alpha_o}{1-(1-k^2)\cot\alpha_s\cot\alpha_o} \approx \frac{\pm\delta(1+\cot^2\alpha_o)}{1+(1-k^2)\cot^2\alpha_o}\left(1 \mp \frac{\delta k^2\cot\alpha_o}{1+(1-k^2)\cot^2\alpha_o}\right).$$

This is of order δ and so the arctans can be removed. Also, since ϵ^2 is of second order already only the $\epsilon^2\pi$ term of the second line of the above expression gives a

significant contribution. Thus, we have

$$\int^{\theta} \frac{\lambda d\theta}{\sin^2\theta[\Theta(\theta)]^{1/2}} \approx \pi \frac{k\lambda}{\sqrt{\eta}} \frac{k^2\sqrt{1+\epsilon^2}+\epsilon^2\sqrt{1-k^2}}{\sqrt{1-k^2}(k^2+\epsilon^2)\sqrt{1+\epsilon^2}}$$

$$+ \frac{k\lambda}{\sqrt{\eta}} \frac{k^2}{(k^2+\epsilon^2)} \frac{\delta(1+\cot^2\alpha_o)}{(1+(1-k^2)\cot^2\alpha_o)} \left(1 \mp \frac{\delta k^2 \cot\alpha_o}{1+(1-k^2)\cot^2\alpha_o}\right).$$

Now,

$$\frac{k\lambda}{\sqrt{\eta}} \frac{k^2\sqrt{1+\epsilon^2}+\epsilon^2\sqrt{1-k^2}}{\sqrt{1-k^2}(k^2+\epsilon^2)\sqrt{1+\epsilon^2}}$$

$$\approx \frac{\lambda}{\sqrt{\lambda^2+\eta}} \left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right) \frac{(k^2(1+\frac{\epsilon^2}{2})+\epsilon^2\sqrt{1-k^2})(1-\frac{\epsilon^2}{k^2})(1-\frac{\epsilon^2}{2})}{k^2\sqrt{1-k^2}}$$

$$\approx \frac{\lambda}{\sqrt{\lambda^2+\eta}} \left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right) \frac{\sqrt{\lambda^2+\eta}}{|\lambda|} \left(1+\frac{a^2\eta}{2(\lambda^2+\eta)^2}\right) \frac{(k^2-\epsilon^2+\epsilon^2\sqrt{1-k^2})}{k^2}$$

$$\approx \frac{\lambda}{|\lambda|} \left(1+\frac{a^2}{2(\lambda^2+\eta)}\right) \left(1-\frac{\epsilon^2}{k^2}(1-\frac{|\lambda|}{\sqrt{\lambda^2+\eta}})\right)$$

$$\approx \frac{\lambda}{|\lambda|} \left(1+\frac{a^2}{2(\lambda^2+\eta)}\right) \left(1-\frac{a^2}{2(\lambda^2+\eta)}(1-\frac{|\lambda|}{\sqrt{\lambda^2+\eta}})\right)$$

$$\approx \frac{\lambda}{|\lambda|} \left(1+\frac{a^2|\lambda|}{2(\lambda^2+\eta)^{3/2}}\right).$$

So

$$\int^{\theta} \frac{\lambda d\theta}{\sin^2\theta[\Theta(\theta)]^{1/2}} \approx \pi \frac{\lambda}{|\lambda|} \left(1+\frac{a^2|\lambda|}{2(\lambda^2+\eta)^{3/2}}\right)$$

$$+ \frac{\lambda}{\sqrt{\lambda^2+\eta}} \frac{\delta(1+\cot^2\alpha_o)}{1+(1-k^2)\cot^2\alpha_o} \left(1 \mp \frac{\delta k^2 \cot\alpha_o}{1+(1-k^2)\cot^2\alpha_o}\right). \quad (5.42)$$

And therefore, adding equations (5.41) and (5.42) we have

$$\phi_o - \phi_s \approx \frac{4ma}{\lambda^2+\eta} - \frac{\pi}{2} \frac{\lambda a^2}{(\lambda^2+\eta)^{3/2}} + \pi \frac{\lambda}{|\lambda|} \left(1+\frac{a^2|\lambda|}{2(\lambda^2+\eta)^{3/2}}\right)$$

$$+\frac{\lambda}{\sqrt{\lambda^2+\eta}}\frac{\delta(1+\cot^2\alpha_o)}{1+(1-k^2)\cot^2\alpha_o}\left(1\mp\frac{\delta k^2\cot\alpha_o}{1+(1-k^2)\cot^2\alpha_o}\right)$$

$$=\pi\frac{\lambda}{|\lambda|}+\frac{4ma}{\lambda^2+\eta}+\frac{\lambda}{\sqrt{\lambda^2+\eta}}\frac{\delta(1+\cot^2\alpha_o)}{1+(1-k^2)\cot^2\alpha_o}\left(1\mp\frac{\delta k^2\cot\alpha_o}{1+(1-k^2)\cot^2\alpha_o}\right).$$

Recall that

$$\cos\theta=k\cos\alpha \quad\Rightarrow\quad \cot^2\alpha=\frac{\cos^2\theta}{k^2-\cos^2\theta}.$$

Therefore,

$$\frac{1+\cot^2\alpha_o}{1+(1-k^2)\cot^2\alpha_o}=\frac{k^2}{(k^2-\cos^2\theta_o)\left(1+(1-k^2)\frac{\cos^2\theta_o}{k^2-\cos^2\theta_o}\right)}$$

$$=\csc^2\theta_o.$$

So finally,

$$\phi_o-\phi_s\approx\pi\frac{\lambda}{|\lambda|}+\frac{4ma}{\lambda^2+\eta}+\frac{\lambda\delta\csc^2\theta_o}{\sqrt{\lambda^2+\eta}}\left(1\mp\delta\cot\theta_o\csc\theta_o\sqrt{\frac{\eta}{\lambda^2+\eta}-\cos^2\theta_o}\right), \quad (5.43)$$

where by equation (5.39)

$$\delta=\frac{15\pi}{4}\frac{m^2}{(\lambda^2+\eta)}+\frac{4m}{\sqrt{\lambda^2+\eta}}-\frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}-\sqrt{\lambda^2+\eta}\left(\frac{1}{r_s}+\frac{1}{r_o}\right),$$

and where the negative sign corresponds to the $\theta_s\rightarrow\theta_{max}\rightarrow\theta_o$ path, whilst the positive sign corresponds to the $\theta_s\rightarrow\theta_{min}\rightarrow\theta_o$ path.

§5.6 Solution for t

In this section we solve equation (5.3) with the aim of deriving an expression for the time delay in the arrivals of two images from the same source. Thus we think of t as $t(\lambda,\eta)$ and are interested in $\Delta t = t(\lambda_2,\eta_2) - t(\lambda_1,\eta_1)$.

Equation (5.3) states that

$$t = \int^{r} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}}dr + \int^{\theta} \frac{a^2\cos^2\theta}{\pm[\Theta(\theta)]^{1/2}}d\theta,$$

where

$$R(r) = r[r(r^2+a^2)+2a^2m] - 4amr\lambda - (r^2-2mr)\lambda^2 - \Delta\eta,$$

$$\Theta(\theta) = \eta + a^2\cos^2\theta - \lambda^2 cot^2\theta$$

and

$$\Delta = r^2 - 2mr + a^2.$$

Firstly we solve the θ integral part of t.

Using the results of §5.4 we have that

$$\int^{\theta} \frac{a^2\cos^2\theta}{\pm[\Theta(\theta)]^{1/2}}d\theta \approx k\int^{\alpha} \frac{a^2\cos^2\theta}{\sqrt{\eta + \frac{a^2\eta^2}{(\lambda^2+\eta)^2}\cos^2\alpha}}d\alpha,$$

where α is defined by $\cos\theta = k\cos\alpha$, and $k = \sqrt{\frac{\eta}{\lambda^2+\eta}}\left(1+\frac{a^2\lambda^2}{2(\lambda^2+\eta)^2}\right)$,

$$= \frac{k^3a^2}{\sqrt{\eta}}\int^{\alpha} \frac{\cos^2\alpha}{\sqrt{1+\frac{a^2\eta}{(\lambda^2+\eta)^2}\cos^2\alpha}}d\alpha$$

$$\approx \frac{\eta a^2}{2(\lambda^2+\eta)^{3/2}}(\sin\alpha\cos\alpha+\alpha)\Big]^{\alpha},$$

For $\theta_s \to \theta_{max} \to \theta_o$ path α ranges $\alpha_s \to \pi \to \alpha_o$, so this is

$$= \frac{\eta a^2}{2(\lambda^2+\eta)^{3/2}}(2\pi - \sin\alpha_s\cos\alpha_s - \alpha_s - \sin\alpha_o\cos\alpha_o - \alpha_o),$$

and for the $\theta_s \to \theta_{min} \to \theta_o$ path α ranges $\alpha_s \to 0 \to \alpha_o$, thus it is

$$= \frac{\eta a^2}{2(\lambda^2+\eta)^{3/2}}(\sin\alpha_s\cos\alpha_s + \alpha_s + \sin\alpha_o\cos\alpha_o + \alpha_o).$$

Now in §5.4 we found that α_s and α_o are related via $\alpha_s \approx \pi - \alpha_o \mp \delta$. However, since the above integral is already of second order this reduces simply to $\alpha_s \approx \pi - \alpha_o$ and in particular, $\sin\alpha_s \approx \sin\alpha_o$ and $\cos\alpha_s \approx -\cos\alpha_o$. Thus, the final result turns out to be independent of the path, and so

$$\int^{\theta} \frac{a^2\cos^2\theta}{\pm[\Theta(\theta)]^{1/2}}d\theta \approx \frac{\pi}{2}\frac{\eta a^2}{(\lambda^2+\eta)^{3/2}}. \tag{5.44}$$

Now we evaluate the r integral part of t. Suppose we are integrating from r_{min} to r_s. Then we have

$$\int^{r} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\pm\Delta[R(r)]^{1/2}}dr = \int_{r_{min}}^{r_s} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\Delta[R(r)]^{1/2}}dr.$$

As in the solution of the left hand side of equation (5.1) we firstly make the substitution $u = 1/r$ and then we let $u = u_{max}x$, thus we have $x = r_{min}/r$ and so

$$\int_{r_{min}}^{r_s} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\Delta[R(r)]^{1/2}}dr$$

$$= \frac{1}{\sqrt{\lambda^2+\eta-a^2}}$$

$$\int_{\frac{r_{min}}{r_s}}^{1} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]dx}{(r^2-2mr+a^2)((1-x^2)(1+a^2u_{max}^2f(x)+mu_{max}g(x))^{1/2}},$$

where $f(x) = \dfrac{\eta(1+x^2)}{\lambda^2+\eta-a^2}$ and $g(x) = -2\left(\dfrac{(\lambda-a)^2+\eta}{\lambda^2+\eta-a^2}\right)\left(\dfrac{x^2+x+1}{x+1}\right)$.

Dividing through by r^4, the above integral becomes

$$= \frac{1}{\sqrt{\lambda^2+\eta-a^2}}$$

$$\int_{\frac{r_{min}}{r_a}}^{1} \frac{[(1+a^2u^2)+2amu^3(a-\lambda)]dx}{u^2(1-2mu+a^2u^2)((1-x^2)(1+a^2u_{max}^2f(x)+mu_{max}g(x))^{1/2}}$$

$$= \frac{r_{min}^2}{\sqrt{\lambda^2+\eta-a^2}}$$

$$\int_{\frac{r_{min}}{r_a}}^{1} \frac{[(1+a^2u_{max}^2x^2)+2amu_{max}^3x^3(a-\lambda)]dx}{x^2(1-2mu_{max}x+a^2u_{max}^2x^2)((1-x^2)(1+a^2u_{max}^2f(x)+mu_{max}g(x))^{1/2}}$$

$$\approx \frac{r_{min}^2}{\sqrt{\lambda^2+\eta-a^2}} \int_{\frac{r_{min}}{r_a}}^{1} \frac{dx}{x^2\sqrt{1-x^2}}\left(1+x^2\frac{a^2}{r_{min}^2}-2x^3\frac{am\lambda}{r_{min}^3}\right)$$

$$\times\left(1+2x\frac{m}{r_{min}}+x^2\frac{4m^2-a^2}{r_{min}^2}\right)\left(1-\frac{f(x)}{2}\frac{a^2}{r_{min}^2}-\frac{g(x)}{2}\frac{m}{r_{min}}+\frac{3g^2(x)}{8}\frac{m^2}{r_{min}^2}\right)$$

$$\approx \frac{r_{min}^2}{\sqrt{\lambda^2+\eta}} \int_{\frac{r_{min}}{r_a}}^{1} \frac{dx}{x^2\sqrt{1-x^2}}\left(1+x^2\frac{a^2}{r_{min}^2}-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}+x^2\frac{4m^2-a^2}{r_{min}^2}\right.$$

$$\left.-\frac{f(x)}{2}\frac{a^2}{r_{min}^2}-\frac{g(x)}{2}\frac{m}{r_{min}}+\frac{3g^2(x)}{8}\frac{m^2}{r_{min}^2}-xg(x)\frac{m^2}{r_{min}^2}+\frac{a^2}{2(\lambda^2+\eta)}\right)$$

$$\approx \frac{r_{min}^2}{\sqrt{\lambda^2+\eta}} \int_{\frac{r_{min}}{r_a}}^{1} \frac{dx}{x^2\sqrt{1-x^2}}\left(1-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}+x^2\frac{4m^2}{r_{min}^2}-(1+x^2)\frac{a^2\eta}{2r_{min}^4}\right.$$

$$\left.+\frac{m}{r_{min}}\left(1+\frac{x^2}{1+x}\right)\left(1-\frac{2a\lambda}{r_{min}^2}\right)\left(1+2x\frac{m}{r_{min}}\right)+\frac{3}{2}\left(1+\frac{x^2}{1+x}\right)^2\frac{m^2}{r_{min}^2}+\frac{a^2}{2r_{min}^2}\right),$$

using equation (5.34), expand $r_{min}/\sqrt{\lambda^2+\eta}$,

$$\approx r_{min}\int_{\frac{r_{min}}{r_s}}^{1}\frac{dx}{x^2\sqrt{1-x^2}}\Bigg(1-\frac{m}{r_{min}+m}-\frac{3m^2}{2r_{min}^2}-\frac{a^2}{2r_{min}^2}+\frac{a^2\eta}{2r_{min}^4}+\frac{2am\lambda}{r_{min}^3}$$

$$-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}\left(1-\frac{m}{r_{min}}\right)-\frac{a^2\eta}{2r_{min}^4}+x^2\left(\frac{4m^2}{r_{min}^2}-\frac{a^2\eta}{2r_{min}^4}\right)$$

$$+\frac{m}{r_{min}}\left(1+\frac{x^2}{1+x}\right)\left(1-\frac{m}{r_{min}}\right)\left(1-\frac{2a\lambda}{r_{min}^2}\right)\left(1+2x\frac{m}{r_{min}}\right)$$

$$+\frac{3}{2}\left(1+\frac{2x^2}{1+x}+\frac{x^4}{(1+x)^2}\right)\frac{m^2}{r_{min}^2}+\frac{a^2}{2r_{min}^2}\Bigg)$$

$$\approx r_{min}\int_{\frac{r_{min}}{r_s}}^{1}\frac{dx}{x^2\sqrt{1-x^2}}\left(1-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}+x^2\left(\frac{4m^2}{r_{min}^2}-\frac{a^2\eta}{2r_{min}^4}\right)\right.$$

$$\left.+\frac{m}{r_{min}}\frac{x^2}{(1+x)}\left(1-\frac{2a\lambda}{r_{min}^2}\right)\left(1+2x\frac{m}{r_{min}}\right)+\frac{2x^2}{(1+x)}\frac{m^2}{r_{min}^2}+\frac{3}{2}\frac{x^4}{(1+x)^2}\frac{m^2}{r_{min}^2}\right)$$

$$\approx r_{min}\int_{\frac{r_{min}}{r_s}}^{1}\frac{dx}{x^2\sqrt{1-x^2}}\Bigg(1-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}+x^2\left(\frac{6m^2}{r_{min}^2}-\frac{a^2\eta}{2r_{min}^4}\right)$$

$$+\frac{m}{r_{min}}\frac{x^2}{(1+x)}\left(1-\frac{2a\lambda}{r_{min}^2}\right)+\frac{3}{2}\frac{x^4}{(1+x)^2}\frac{m^2}{r_{min}^2}\Bigg)$$

$$\approx r_{min}\int_{\frac{r_{min}}{r_s}}^{1}\frac{dx}{x^2\sqrt{1-x^2}}\Bigg(1-2x^3\frac{am\lambda}{r_{min}^3}+2x\frac{m}{r_{min}}+x^2\left(\frac{15m^2}{2r_{min}^2}-\frac{a^2\eta}{2r_{min}^4}\right)$$

$$+\frac{m}{r_{min}}\frac{x^2}{(1+x)}\left(1-\frac{2a\lambda}{r_{min}^2}-\frac{3m}{r_{min}}\right)+\frac{3}{2}\frac{x^2}{(1+x)^2}\frac{m^2}{r_{min}^2}\Bigg).$$

In order to solve this integral we need the following results.

$$\int\frac{dx}{x^2\sqrt{1-x^2}}=-\frac{\sqrt{1-x^2}}{x},\quad\int\frac{xdx}{\sqrt{1-x^2}}=-\sqrt{1-x^2},$$

$$\int\frac{dx}{x\sqrt{1-x^2}}=-\ln\left(\frac{2+2\sqrt{1-x^2}}{x}\right),$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x), \quad \int \frac{dx}{(1+x)\sqrt{1-x^2}} = -\sqrt{\frac{1-x}{1+x}},$$

$$\int \frac{dx}{(1+x)^2\sqrt{1-x^2}} = -\frac{2+x}{3(1+x)}\sqrt{\frac{1-x}{1+x}}.$$

Therefore,

$$\int_{r_{min}}^{r_s} \frac{[r^2(r^2+a^2)+2amr(a-\lambda)]}{\Delta[R(r)]^{1/2}}dr$$

$$\approx \sqrt{r_s^2 - r_{min}^2} - \frac{2am\lambda}{r_{min}^2}\sqrt{1-\frac{r_{min}^2}{r_s^2}} + 2m\ln\left(\frac{r_s}{r_{min}} + \sqrt{\frac{r_s^2}{r_{min}^2}-1}\right)$$

$$+\left(\frac{\pi}{2} - \arcsin\left(\frac{r_{min}}{r_s}\right)\right)\left(\frac{15m^2}{2r_{min}} - \frac{a^2\eta}{2r_{min}^3}\right) + \frac{m^2}{r_{min}}\frac{(2r_s+r_{min})}{2(r_s+r_{min})}\sqrt{\frac{r_s-r_{min}}{r_s+r_{min}}}$$

$$+m\left(1 - \frac{2a\lambda}{r_{min}^2} - \frac{3m}{r_{min}}\right)\sqrt{\frac{r_s-r_{min}}{r_s+r_{min}}}$$

$$\approx r_s - \frac{r_{min}^2}{2r_s} - \frac{4am\lambda}{r_{min}^2} + 2m\ln\left(\frac{2r_s}{r_{min}}\right) + \frac{\pi}{2}\left(\frac{15m^2}{2r_{min}} - \frac{a^2\eta}{2r_{min}^3}\right)$$

$$+m\left(1 - \frac{2m}{r_{min}} - \frac{r_{min}}{r_s}\right). \quad (5.45)$$

So finally with the assistance of equation (5.44), and adding the integral from r_{min} to r_o we have

$$t \approx r_s + r_o - \frac{r_{min}^2}{2}\left(\frac{1}{r_s}+\frac{1}{r_o}\right) - \frac{8am\lambda}{r_{min}^2} + 2m\ln\left(\frac{4r_s r_o}{r_{min}^2}\right)\frac{15\pi m^2}{2r_{min}}$$

$$+2m\left(1 - \frac{2m}{r_{min}} - \frac{r_{min}}{2}\left(\frac{1}{r_s}+\frac{1}{r_o}\right)\right), \quad (5.46)$$

and therefore,

$$\Delta t = t(\lambda_2, \eta_2) - t(\lambda_1, \eta_1)$$

$$\approx \frac{r_2^2 - r_1^2}{2}\left(\frac{1}{r_s} + \frac{1}{r_o}\right) + 8am\left(\frac{\lambda_2}{r_2^2} - \frac{\lambda_1}{r_1^2}\right) + 4m\ln\left(\frac{r_2}{r_1}\right) - \frac{15\pi m^2}{2}\left(\frac{1}{r_2} - \frac{1}{r_1}\right)$$

$$+4m^2\left(\frac{1}{r_2} - \frac{1}{r_1}\right) + m(r_2 - r_1)\left(\frac{1}{r_s} + \frac{1}{r_o}\right), \quad (5.47)$$

where r_1 is r_{min} for the ray of light with constants of the motion λ_1 and η_1, whilst r_2 is r_{min} for the ray with λ_2 and η_2 as its constants of the motion.

§5.7 Quality of the approximations

In this section we discuss the quality of the approximations made in the previous chapters and indicate how the derived expressions can be combined with numerical integration to give excellent results for relatively high deflections.

From past experience with Schwarzschild black holes we know that the deviation of a ray of light is small whenever the ratio m/r_{min} is small. In the Kerr case this is still true, however a new factor plays an important part in the size of the deviation. A particular ratio m/r_{min} determines the deviation up to two possibilities which in turn are determined by the sign of λ. This is most evident in the equatorial plane as the effect of spin is most apparent there, so we will restrict this discussion to the equatorial plane. Another advantage of this is that the deflection is simply $|\pi - \phi|$ i.e., θ is not involved.

A positive λ corresponds to a ray travelling with the spin, a prograde orbit, and which is therefore not as greatly deviated as the ray with negative λ which corresponds to a ray travelling against the spin, a retrograde orbit. As a result we find that the accuracy of our approximations of the previous section depends on the sign of λ. This is easily seen when substituting the approximate value for r_{min} into $R(r)$. We find that $R(r_{min})$ is much closer to 0 for positive λ than for negative λ, for any given ratio of $m/\sqrt{\lambda^2+\eta}$. Similarly, any of the integrals approximated above which have terms proportional to λ are more accurate for positive λ than for negative λ. Fortunately the integrals involving θ have λ appear as λ^2 only, and furthermore, have higher order terms of integral powers of $a^2/(\lambda^2+\eta)$. Thus, the approximations of integrals which involve θ remain good for relatively high ratios of $m/\sqrt{\lambda^2+\eta}$ and are independant of the sign of λ.

There is another more significant source of error which comes in the evaluation of δ. The value δ was obtained by equating $\int \frac{dr}{[R(r)]^{1/2}}$ and $\int \frac{d\theta}{[\Theta(\theta)]^{1/2}}$ We found that

$$\int \frac{dr}{[R(r)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(\pi\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}+\frac{15m^2}{4(\lambda^2+\eta)}\right)\right.$$

$$\left.+\frac{4m}{\sqrt{\lambda^2+\eta}}-\frac{8ma\lambda}{(\lambda^2+\eta)^{3/2}}\right)-\frac{1}{r_s}-\frac{1}{r_o}$$

and

$$\int \frac{d\theta}{[\Theta(\theta)]^{1/2}} \approx \frac{1}{\sqrt{\lambda^2+\eta-a^2}}\left(1-\frac{3a^2\eta}{4(\lambda^2+\eta)^2}\right)$$

$$\times\left(\pi \pm \arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_s\right) \pm \arctan\left(\left(1+\frac{a^2\eta}{4(\lambda^2+\eta)^2}\right)\cot\alpha_o\right)\right).$$

When these equations are equated the "large" $\pi/\sqrt{\lambda^2+\eta-a^2}$ terms cancel on both sides leaving us with a small δ. Recall that δ was set to be equal to the sum of the arctans. Since the absolute error of the former equation involving r is of third order and is particularly large for negative λ the relative error in δ can easily become unacceptable. If this occurs then this integral should be evaluated numerically, something which can be done quickly using Gauss-Jacobi quadratures and even more quickly using the method of Elhay and Kautsky (1984), where succesive approximations use nodes already used in previous approximations. Thus the new δ is the numerical integral minus the term proportional to π of the latter equation involving θ. Then the equation

$$\cos\theta_s \approx -\cos\theta_o\cos\delta \pm \sin\delta\sqrt{\frac{\eta}{\lambda^2+\eta}-\cos^2\theta_o}$$

may still be used to give remarkably good results.

In the evaluation of ϕ_s the integral containing r may give poor results for negative λ. In this case evaluate it numerically as above, add the result to equation (5.43), and replace the $4ma/(\lambda^2+\eta)$ term by $\frac{\pi}{2}a^2\lambda/(\lambda^2+\eta)^{3/2}$. Thus, we have

$$\phi_o-\phi_s \approx \int^r \frac{a(2mr-a\lambda)}{\pm\Delta[R(r)]^{1/2}}dr + \pi\frac{\lambda}{|\lambda|}+\frac{\pi}{2}\frac{a^2\lambda}{(\lambda^2+\eta)^{3/2}}$$

$$+\frac{\lambda\delta\csc^2\theta_o}{\sqrt{\lambda^2+\eta}}\left(1\mp\delta\cot\theta_o\csc\theta_o\sqrt{\frac{\eta}{\lambda^2+\eta}-\cos^2\theta_o}\right).$$

§6 Numerical Investigation

Recall that the motivation behind the work in this the second part of the thesis, is to produce diagrams which show the multiplicity of images as a function of the position of the source for black holes. Firstly we have to demonstrate that this indeed makes sense.

Black holes can have null geodesics which have circular orbits. This means that a single source will generate infinitely many images at the observer should a black hole exist anywhere in the universe. These images are due to rays which arrive at the position of the observer directly, and those which first orbit the black hole any countable number of times. Clearly, not all of these images are observed. Thus, it becomes necessary to talk not of images in the mathematical sense, but of images in the observable sense. We differentiate between images on the grounds that some have a greater probability of being observed than others. A criterion for an image to be an observable image must be established. This is almost arbitrary, as our only information is that the more orbits a ray has about the black hole before reaching the observer the less likely it is of being observed, Cunningham and Bardeen (1973). The criterion we adopt, without loss of much generality, is that in order for an image to be an observable image, we only consider rays with direct paths i.e., those which do not orbit the black hole, but arrive directly at the position of the observer suffering a relatively small deviation. Furthermore, we stipulate that the source is in the background of the black hole as seen by the observer. The latter condition is imposed as the former would allow rays which suffer deviations near 180 degrees.

The equations we derived in chapter 5 are applicable to these rays i.e., the direct rays due to a source in the background of the black hole as seen by the observer.

Under these conditions the Schwarzschild black hole should give similar results for the multiplicity of images as a function of the position of the source, as did the spherical gravitational lens models in part one of the thesis. The major difference is that unlike the earlier models, black holes are not transparent. Thus, for the Schwarzschild black hole we expect to see at most two images, with the table of

image multiplicities to have spherical symmetry. This is fairly clear. What is more interesting is to find the effect of introducing spin to the black hole. Would we get more than two images, and what would be the effect on the symmetry of the table of image multiplicities, are immediate questions that arrise.

The numerical techniques used in answering these questions are identical to those used in the first part of this thesis. The $\alpha - \beta$ plane of fig.5.1 is considered as a complex plane with $z_i = \alpha_i + i\beta_i$ being a projected image position which is mapped onto a projected source position $z_s = \alpha_s + i\beta_s$. The mapping is done by firstly finding λ and η corresponding to a particular α_i and β_i by equations (6.1) and (6.2), evaluating the equations of motion (6.3) and (6.4) to find θ_s and ϕ_s, and finally finding α_s and β_s by equations (6.6) and (6.7).

We assume that our black hole is an extreme Kerr black hole i.e., $|a| = m$. Without any loss of generality we choose $a = m$ and $\phi_o = 0$. The choice of θ_o is not arbitrary. For simplicity and maximal effect of spin we take $\theta_o = \pi/2$. Equations (5.27), (5.28), (5.38), (5.43), (5.39), (5.20) and (5.21) now become, respectively

$$\lambda \approx \frac{-r_o \alpha_i}{\sqrt{\beta_i^2 + \alpha_i^2 + r_o^2}} \tag{6.1}$$

$$\sqrt{\eta} \approx \frac{r_o \beta_i}{\sqrt{\beta_i^2 + \alpha_i^2 + r_o^2}} \tag{6.2}$$

$$\cos\theta_s \approx \pm \sin\delta \sqrt{\frac{\eta}{\lambda^2 + \eta}} \tag{6.3}$$

$$-\phi_s \approx \pi \frac{\lambda}{|\lambda|} + \frac{4m^2}{\lambda^2 + \eta} + \frac{\lambda\delta}{\sqrt{\lambda^2 + \eta}} \tag{6.4}$$

$$\delta = \frac{15\pi}{4} \frac{m^2}{(\lambda^2 + \eta)} + \frac{4m}{\sqrt{\lambda^2 + \eta}} - \frac{8m^2\lambda}{(\lambda^2 + \eta)^{3/2}} - \sqrt{\lambda^2 + \eta} \frac{(r_s + r_o)}{r_s r_o}. \tag{6.5}$$

$$\alpha_s = \frac{r_s r_o \sin\theta_s \sin\phi_s}{r_o - r_s \cos\phi_s \sin\theta_s} \tag{6.6}$$

and

$$\beta_s = \frac{r_s r_o \cos\theta_s}{r_o - r_s \cos\phi_s \sin\theta_s}. \tag{6.7}$$

We next need to choose r_s, r_o and m.

As stated earlier we are mainly interested in the effect of spin. Since spin appears only at second order we require relatively large deflections. However, we project the position of the source and the image onto the same plane. So that the sources and their images will not be too far apart we choose a relatively small value for r_o. We also take $r_s = r_o$, as the deflection is maximal if the deflecting object is between the source and the observer. The following diagrams were generated by choosing $m = a = \frac{1}{2}$ and $r_o = r_s = 14$.

Having chosen all the necessary parameters we proceed as stated above. The calculation of mapping z_i to z_s is done for every point of the grid which is not "too far away" from the central point $(0,0)$. In particular, the following diagrams were generated with $||(\alpha_i,\beta_i)|| \leq 10$.

Points (α_i,β_i) in the neighbourhood of $(0,0)$ correspond to rays which suffer "large" deflections. These calculations were done using the numerical integration technique described in §5.7. In the diagrams we defined a ray to have a "large" deflection if either $||(\alpha_i,\beta_i) - (\alpha_s,\beta_s)|| \geq 8$, or if the angle θ does not vary as $\theta_s \to \theta_{min/max} \to \theta_o$. The latter condition assigns "large" deflections to rays which orbit the black hole or are captured by the black hole. Fig.6.1 shows the regions in which the rays suffer "large" deviations for both the Kerr and Schwarzschild black holes. The Kerr region is comparable with the region due to the black hole being illuminated by an extended light source whose angular size is larger than that of the black hole, Bardeen(1973).

Fig.6.2 shows the distribution of the number of images as a function of the position of the source. For contrast we compare the results for the Kerr black hole with those for a Schwarzschild black hole. We see that introducing spin does not increase the multiplicity of images, but merely extends the region where a source generates two images, to the right.

Recall that a point (α_i,β_i) is considered to be an observable image of a source if the ray did not suffer a "large" deflection. This definition of an image is borne out by Fig.6.3, where the image corresponding to rays with "large" deflections is

quite small, and therefore is less likely to be observed.

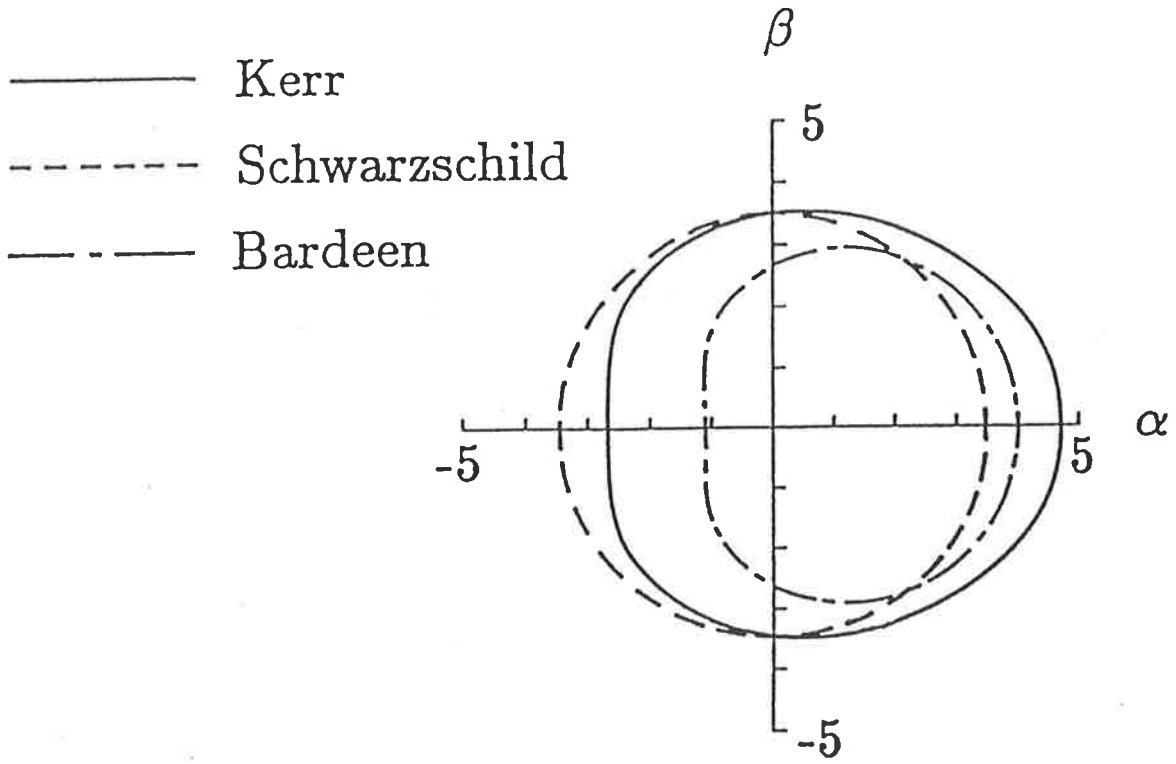

Figure 6.1 Regions in which (α_i,β_i) correspond to sources whose rays suffer "large" deviations. Bardeen's region is calculated by setting $R(r) = \dot{R}(r) = 0$.

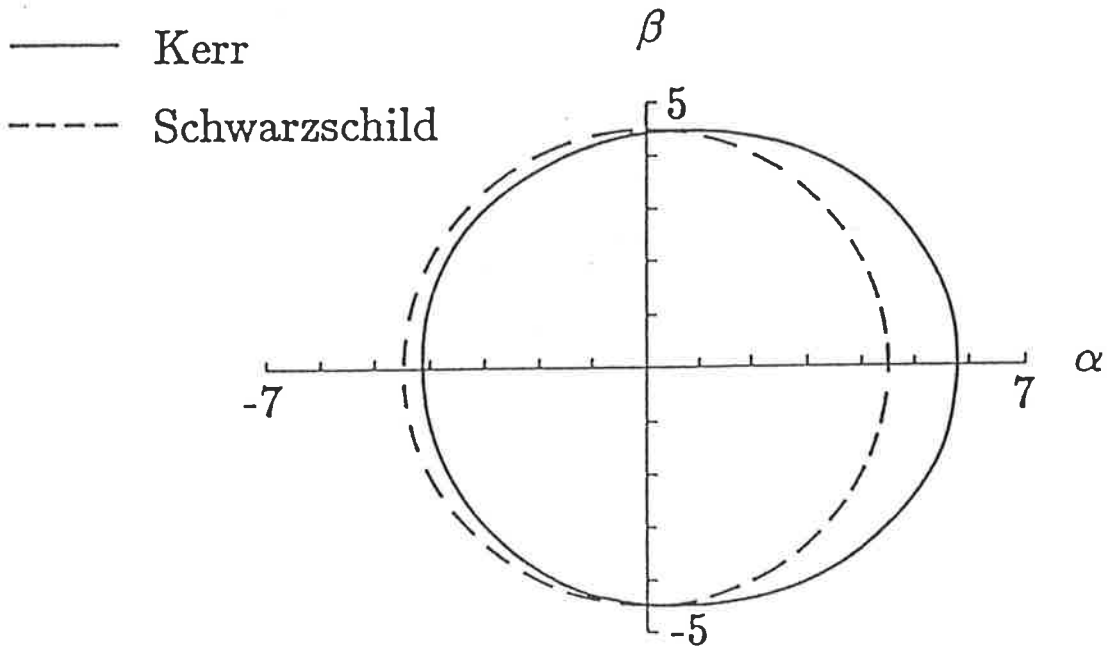

Figure 6.2 Sources within these regions have two images, those not within these regions have only one.

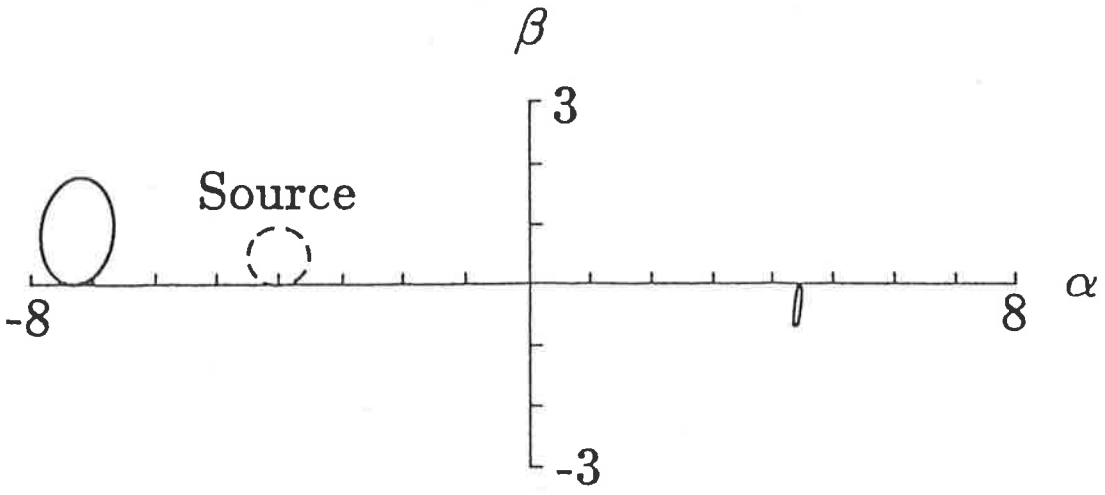

Figure 6.3 Images of a source near the boundary of the Kerr region of fig.6.2. Note that the image on the left, which corresponds to rays with small deviations, is amplified. On the other hand, the image on the right hand side, which corresponds to rays with large deviations, is diminished, and therefore is less likely to be observed.

A summary of this the second part of the thesis is contained in the enclosed preprint of the same name. It is to be published in The Physical Review D15 (1986).

References

Bardeen, J. M. 1970, *Ap. J.*, **161**, 103.

Bardeen, J. M. 1973. In: *Black Holes*, eds. DeWitt, C. & DeWitt, B. S.

Bardeen, J. M. 1974. In: *Gravitational Radiation and Gravitational Collapse*, IAU Symposium 64, ed. deWitt-Morette, C. (Reidel, Dordrecht).

Bardeen, J. M., Press, W. H. & Teukolsky, S. A. 1972, *Ap. J.*, **178**, 347.

Barnothy, J. M. & Barnothy, M. F. 1972, *Ap. J.*, **174**, 477.

Börner, 1973, Springer Tracts in Modern Physics, Vol. 69 (Springer, Berlin).

Bourassa, R. R., Kantowski, R. & Norton, T. D. 1973, *Ap. J.*, **185**, 747.

Bourassa, R. R. & Kantowski, R. 1975 *Ap. J.*, **195**, 13.

Boyer, R. H. & Lindquist, R. W. 1967, *J. Math. Phys.*, **8**, 265.

Boyer, R. H. & Price, T. G. 1965, *Proc. Cambridge Philos. Soc.*, **61**, 531.

Bray, I. 1984, *M. N. R. A. S.*, **208**, 511

Bray, I. 1986, *Phys. Rev.* in press.

Burke, W. L. 1981, *Astrophysical Journal*, **244**:L1.

Carter, B. 1968, *Physical Review* D**174**, 1559.

Chaffee, F. H., Jr. 1980, *Sc. A.*, **11**, 60.

Chandrasekhar, S. 1935, *M. N. R. A. S.*, **95**, 207.

Cunningham, C. T. & Bardeen, J. M. 1973, *Asrophysical Journal*, **183**, 237.

Einstein, A. 1936, *Science*, **84**, 506.

Elhay, S. & Kautsky, J. 1984, *Aust. Comp. Sc. Comm.*, **6**, 15-1.

Etherington, I. M. H. 1933, *Phil. Mag.*, **15**, 761.

Felice, F. de & Calvani, M. 1972, *Nuovo Cimento Cer II*, **10**B, 447.

Godfrey, B. B. 1970, *Phys. Rev. D*, **1**, 2721.

Greenfield, P. E., Roberts, D. H. & Burke, B. F. 1980, *Science*, **208**, 495.

Harrison, B. K., Thorne, K. S., Wakano, M., Wheeler, J. A. 1965, Gravitational Theory and Gravitational Collapse. Chicago: Univ. of Chicago.

Hawking, S. W. & Penrose, R. 1970, *Proc. Roy. Soc.*, A**314**, 529.

Hawking, S. W. 1972, *Comm. Math. Phys.*, **25**, 152.

Helliwell, T. M. & Mallinckrodt, A. J. 1975, *Phys. Rev. D*, **12**, 2993.

Kerr, R. P. 1963, *Phys. Rev. Letters*, **11**, 237.

King, I. R. 1966, *Astronomical Journal*, **71**, 64.

Klimov, Yu. G. 1963, *Soviet Phys. - Doclady*, **8**, 119.

Krivenko, O. P., Pyragas, K. A. & Zhuk, I. T. 1976, *Ap. Space Sci.*, **40**, 39.

Landau, L. 1934, *Phys. Z. Sowjetunion*, **1**, 285.

Liebes, S. 1964, *Physical Review*, **133**, B835.

Lynden-Bell, D. 1969, *Nature*, **223**, 690.

Lynden-Bell, D. & Rees, M. J. 1971, *M. N. R. A. S.*, **152**, 461.

Metzner, R. W. K. 1963, *J. Math. Phys.*, **4**,1194.

Noonan, T. W. 1971, *A. J.*, **76**, 9.

Oppenheimer, J. R. & Volkoff, G. 1939, *Physical Review*, **55**, 374.

Pajerski, D. W. & Newman, E. T. 1972, *J. Math. Phys.*, **12**, 1929.

Penrose, R. 1965, *Proc. Roy. Soc.*, A**284**, 159.

Penrose, R. 1968, Contempory Physics: Trieste Symposium 1968, Vol I, p. 545.

Penrose, R. 1969, *Rivista Nuovo Cimento*, Serie I, **1**, Numero Speciale, 252.

Perek, L. 1962, *Advances in Astronomy and Astrophysics*
(London and New York: Academic Press), **1**, 165.

Refsdal, S. 1964a, *M.N.R.A.S.*, **128**, 290.

Refsdal, S. 1964b, *ibid.*, p. 307.

Refsdal, S. 1966, *ibid.*, **132**, 101.

Roberts, D. H., Greenfield, P. E., Burke, B. F. 1979, *Science*, **205**, 894.

Ruffini, R. & Wheeler, J. A. 1970, in : Moore, A. F. & Hardy, V. 1970.
The Significance of Space Research for Fundamental Physics (ESRO, Paris).

Schmidt, M. 1965, in *Galactic Structure* (Chicago: University of Chicago Press), p.513.

Stewart, J. M. & Walker, M. 1973, *Black Holes: The Outside Story*,
Springer Tracts in Modern Physics, Vol. 69 (Springer, Berlin).

Stockton, A. 1980, *Ap. J. (Letters)*

Tikhov, G. A. 1937, *Docl. Akad. Nauk.*, **16**,199.

Walsh, D., Carswell, R. F. and Weymann, R. J. 1979, *Nature*, **279**, 381.

Wilkins D.C. 1972, *Physical Review D*, **5**, 814.

Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. 1980,
Ap. J., **240**, 507.

Young, P., Gunn, J. E., Kristian, J., Oke, J. B. & Westphal, J. A. 1981,
Ap. J., **244**, 736.

Zwicky, F. 1937, *Physical Review Letters*, **51**, 290.

Zwicky, F. 1957, Morphological Astronomy (Berlin: Springer-Verlag), p.215.