

GRAVITATIONAL LENSING IN AN EXACT LOCALLY INHOMOGENEOUS
COSMOLOGY

by

Allen Attard

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Graduate Department of Astronomy and Astrophysics
University of Toronto

Copyright © 2005 by Allen Attard

Abstract

Gravitational Lensing in an Exact Locally Inhomogeneous Cosmology

Allen Attard

Doctor of Philosophy

Graduate Department of Astronomy and Astrophysics

University of Toronto

2005

A Recursive Swiss-Cheese (RSC) cosmological model is an exact solution to Einstein's general relativistic field equations allowing for dramatic local density inhomogeneities while maintaining global homogeneity and isotropy. It is constructed by replacing spherical regions of an FRW background with higher density cores placed at the centre of a Schwarzschild vacuum, with each core itself potentially being given the same treatment and the process repeated to generate a range of multifractal structures.

Code was developed to tightly pack spheres into spaces of constant curvature in an efficient manner, and was used to develop libraries of packings with positive, negative, and zero curvature. Various projections are used to illustrate their structure, and means of measuring its dimensionality are discussed. A method by which these packings can be used as building blocks of an RSC model, along with a way of selecting parameters to define the model, is described, and a coordinate system allowing a relativistically consistent means of synchronizing its various components is developed.

Formulations of the optical scalar equations for the expansion and shear rates of a beam are considered, and a set suitable for numerical integration selected. The forms of the null geodesic beam trajectories in each region of the model are computed, and a parallel propagated shadow plane basis that can be consistently followed between the various model sections is established. This allowed the development of code using a fourth order, variable step size Runge-Kutta integration routine to compute the gravitational

lensing effect within an RSC model by tracking the amplification and distortion of a series of beams that are propagated through it. The output generated allows the redshift evolution of these quantities to be plotted for each beam, and enables maps to be made of the “observed sky”. The amplification signature produced by a single lens in the model is examined, and the form shown to be generally consistent with that found using a thin lens approximation, particularly when the lensing is weak. Distortion values are likewise shown to be reasonable, and results derived from propagating beams through a full RSC model are also presented.

Acknowledgements

First and foremost, I would like to thank my supervisor and mentor, Professor Charles Dyer. He introduced me to gravitational lensing and the RSC model, and has proven to be an endless source of ideas and enthusiasm for my research. I truly appreciate the support, encouragement, and welcoming manner he has shown, and have learned a great deal from the fascinating discussions, both academic and otherwise, that we have had.

Over the years that I have been in the Department of Astronomy and Astrophysics, I have had the pleasure of interacting with a large number of interesting and inspiring individuals, and have gained many friends. While I will not attempt to name them all, there are a few who in one way or another have had a significant impact on this work. I would like to recognize Chris Burns for assisting me from the very beginning of my post-graduate work, leading to a friendship that was cemented over the many discussions and various journeys we have had. Whether pondering scientific problems, backpacking through driving rain, or hanging by our fingertips, I found a kindred spirit in Arno Dirks, with his shared passion for the wilderness adventures that have helped me keep a healthy perspective on the daily grind. I express my sincere appreciation to Patrick Carey, Mark Brodwin, and Vjera Miović, whose friendship and understanding, particularly at the beginning, middle, and end of this thesis have meant so much. I would also like to thank Megan McClure for helping me generate a number of the packings. To all the rest, and there are many of you, thank you for helping me get so much more out of the experience.

I gratefully acknowledge the funding for this work that was provided in large part by the Natural Science Engineering and Research Council of Canada, the Government of Ontario, and the University of Toronto.

Finally, I would like to express my deepest gratitude to my parents, for their unwavering support and belief in me which has been a tremendous source of inspiration throughout the years and over the course of this work. Fl-aħħar spicćajt!

Contents

1	Introduction	1
1.1	Notation and Conventions Adopted	2
1.2	Swiss-Cheese Cosmology	4
1.2.1	The Schwarzschild Solution	4
1.2.2	The FRW Model	5
1.2.3	Combining the Solutions	8
1.3	Gravitational Lensing	10
1.3.1	The Thin Lens Approximation	10
1.3.2	The Optical Scalar Approach	13
1.4	Previous Work	20
2	The RSC Model	23
2.1	The SC Packings	23
2.1.1	Generating the Packings	24
2.1.2	Packing Structure	27
2.2	Generating an RSC Model	42
2.2.1	Truncating the Model	44
2.2.2	Storing the Model	45
3	Propagating a Beam Through RSC	47
3.1	The FRW Regions	48

3.1.1	Orienting the System	49
3.1.2	Finding the Target	53
3.2	The Schwarzschild Regions	60
3.2.1	Determining the Potential Closest Approach	64
3.3	Connecting the Regions	65
3.3.1	FRW Background to Schwarzschild Vacuum	65
3.3.2	Schwarzschild Vacuum to FRW Background	69
3.3.3	Core Dynamics	69
3.3.4	Defining the Core	79
3.3.5	Moving Between the Vacuum and Core	85
3.3.6	Tracking the Redshift	87
3.4	Ending the Propagation	90
4	The Optical Scalars	93
4.1	The Driving Terms	93
4.2	Analyzing the Beam Cross-Section	94
4.2.1	The (A, ξ, φ) Form	95
4.2.2	The (C_{\pm}, α_{\pm}) Form	98
4.2.3	The (ψ, η) Form	99
4.3	Spinor Properties	100
4.4	A Parallel Propagated Shadow Plane Basis	102
4.4.1	The FRW Regions	102
4.4.2	The Schwarzschild Regions	112
4.5	Tracking the Beam Distortion	117
4.5.1	Initializing the Orientation	117
4.5.2	Following the Orientation Over Multiple Encounters	118

5	Computational Considerations	123
5.1	The Packing Code	123
5.2	Using the Lensing Code	124
5.2.1	Initialization	126
5.2.2	Processing the Output	128
5.3	The Integrator	131
5.3.1	Setting Step Sizes	132
5.3.2	Passing Closest Approach	134
5.4	Detecting Caustics	136
5.5	Consistency Checks	138
6	Sample Results and Discussion	143
6.1	Trajectory Analysis	143
6.2	The Optical Scalars	146
6.2.1	Propagation Through the FRW Background	146
6.2.2	Single Lens Without Core Encounter	150
6.2.3	Single Lens With Core Encounter	160
6.2.4	Multiple Lenses	167
6.3	Comments On Thick Versus Thin Lenses	174
6.4	Directions For Future Work	176
6.5	Concluding Remarks	180
A	RW Volume Computation	183
B	Computing the Cylindrical Mass	185
C	Coordinate Systems	187
C.1	Coordinate Conversions	187
C.1.1	Angular and Cartesian Coordinates	187

C.1.2 RW and Schwarzschild Coordinates	191
C.2 Alternate Coordinates for Vacuum Regions	193
D Random Point Generation	197
Bibliography	201

List of Figures

1.1	Tide-free surface in Swiss-Cheese	10
1.2	Thin lens setup	12
1.3	Expansion and shear of a circular beam	15
1.4	Definition of β and φ	16
2.1	Packing code flow diagram	24
2.2	Sample θ vs ω sphere distribution with $K=-1$, $\omega_{\max} = 0.49\pi$	29
2.3	Sample ϕ vs ω sphere distribution with $K=-1$, $\omega_{\max} = 0.49\pi$	29
2.4	Sample θ vs ϕ sphere distribution with $K=-1$, $\omega_{\max} = 0.49\pi$	30
2.5	θ vs ϕ sphere distributions for other packings	32
2.6	Cartesian projections of three packings	34
2.7	Sphere radius distribution for a packing	35
2.8	The first 23 and 100 spheres in a packing	36
2.9	A full packing depicted with opaque spheres	37
2.10	Sample packing correlation function	39
3.1	Quantities for FRW propagation	49
3.2	Space-time diagram of beam propagation through a hole	62
3.3	Propagation through a hole – top view	63
3.4	Beam grazing a core	63
3.5	Kruskal diagram	72

3.6	Kruskal diagram with infalling coordinates	75
4.1	The flag and shadow planes	103
4.2	Comparing shadow plane vectors	122
5.1	Determining ϕ adjustment at closest approach in FRW	135
5.2	Determining the location of a caustic	138
5.3	Vacuum propagation trajectories	140
6.1	Beam trajectories with varying core size	144
6.2	Beam trajectories with up to two lenses	145
6.3	Geometric amplification (\mathcal{A}) scan in FRW	147
6.4	\mathcal{A} vs z in FRW	148
6.5	Cross-sectional area vs z in FRW	149
6.6	\mathcal{A} across a lens with no core hit	152
6.7	\mathcal{A} scans with offset angle and redshift variation	153
6.8	\mathcal{A} vs z for a set of beams	155
6.9	Multiple scans of \mathcal{A} across a lens with no core hit	156
6.10	Multiple scans of beam distortion (\mathcal{D}) across a lens with no core hit	157
6.11	\mathcal{D} vs z for a set of beams	158
6.12	Distortion orientation (α_e) across a lens with no core hit	159
6.13	\mathcal{A} across a lens whose core is hit	161
6.14	Multiple scans of \mathcal{D} across a lens core	163
6.15	\mathcal{A} vs z with a core encounter	164
6.16	\mathcal{A} and \mathcal{D} vs z for a core encounter	165
6.17	α_e vs z with a core encounter	166
6.18	\mathcal{A} distribution from an RSC model	168
6.19	\mathcal{D} distribution from an RSC model	168
6.20	\mathcal{A} and cross-sectional area vs z with multiple lens encounters	170

6.21 \mathcal{D} vs z with multiple lens encounters 171

6.22 α_e contours from an RSC model 172

6.23 α_e variation with two lens encounters 173

B.1 Computing cylindrical mass 185

D.1 Defective random point and packed sphere distributions 200

Chapter 1

Introduction

The light we observe when we look out into the universe has travelled a great distance to reach us and during this time, it is influenced by the space through which it travels. Einstein's general theory of relativity tells us that the shape of spacetime is determined by the mass (or energy) it contains. One of the most interesting consequences of this spectacular insight is that light propagating in this warped space appears to be bent and distorted by the matter distribution it encounters. This gravitational lensing effect has been verified experimentally and is now providing us with a powerful tool in our quest to probe the structure of the universe, yielding data on matter that we can see directly, as well as on that which is dark and whose existence can only be inferred through the gravitational effect it produces.

If the mass in the universe were distributed smoothly at all scales, then the universe would have a constant curvature. It would be homogeneous and isotropic, so each point in it would be the same as any other point, and from a given vantage point, the appearance in any direction would be the same as it is in any other direction. These two properties define the Cosmological Principle which underlies much of modern cosmology. Observational evidence, such as that obtained through the investigation of the cosmic microwave background, does indeed indicate that when examined on large enough scales,

the universe is largely homogeneous and isotropic. However, this is clearly not the case on smaller scales. Filamentary structures and large voids become apparent in observational data on scales of about 10 to 100 Mpc, and there is an abundance of structure at even smaller scales, with galaxies grouping into clusters and clusters into super-clusters.

While the behaviour of a photon travelling through a smooth universe that adheres to the Cosmological Principle at all scales can be determined fairly easily, the presence of all of the structure in the real universe makes this task more daunting. With the increasingly important role that gravitational lensing is playing in astrophysics, it is useful to have a relativistically consistent model that can be used to investigate the lensing effect of this structure. The Recursive Swiss-Cheese (RSC) model is one such model that lends itself to the investigation of the potentially complicated photon trajectories that can arise due to gravitational lensing in a universe with interesting matter distributions.

1.1 Notation and Conventions Adopted

Signature: A timelike signature is used so that temporal spacetime intervals are positive while spatial ones are negative.

Tensor Indices: Unless otherwise indicated, lower case Latin letters will be used for tensor index values running from 0 to 3, with 0 being used for the timelike component. Lower case Greek letters will be used to denote the spacelike subset running from 1 to 3.

Spinors: Upper case Latin letters will be used for spinor indices, and can have the values of $\textcircled{1}$ or $\textcircled{2}$, where the numbers are circled as customary in order to distinguish them from values of tensor indices. The complex conjugate of a spinor will be noted by a bar over it and a dot over its index.

Derivatives: Unless otherwise noted, a quantity with a dot over it will represent the full derivative of that quantity with respect to an affine parameter. The covariant derivative will be indicated using a comma in the indices, so for instance, $g_{ab,c}$ would be used to

denote the covariant derivative of the metric tensor.

Summation Convention: The Einstein summation convention is assumed for repeated indices so

$$x^a y_a = \sum_{a=0}^3 x^a y_a$$

Coordinate Variables: In FRW regions, spacetime positions will be described using $x^a = (cT, \omega, \theta, \phi)$, where ω is a radial value, while θ and ϕ are the “latitude” and “longitude” respectively of conventional spherical polar coordinates. These are all angular except for ω which is linear when $K = 0$. Schwarzschild regions will use (ct, r, θ, ϕ) . The usual Cartesian coordinates (x, y, z) will be used for Euclidian spaces, and extended to (x, y, z, w) when discussing spatial positions in an embedding space for curved 3-manifolds.

Variables: Aside from the above, there are a number of variables that appear throughout the thesis. These are summarized below for reference.

- $K = 0, \pm 1$ is the curvature indicator.
- R is used for the FRW scale factor.
- k^a is the tangent vector to the null geodesic followed by a beam.
- β is the phase of the tidal force.
- φ is the phase of the shear rate.
- λ and τ are used as affine parameters along null and timelike geodesics respectively. They will be chosen to have units of length.
- A subscripted value of α denotes an angle.

1.2 Swiss-Cheese Cosmology

As a result of the strong connection between the geometry of spacetime and the matter that it contains, general relativity does not allow an arbitrary matter distribution to be superimposed on a given spacetime structure. Einstein's general relativistic field equations can be written as:

$$G_{mn} = R_{mn} - \frac{1}{2}R_b^b g_{mn} + \Lambda g_{mn} = -\kappa T_{mn} , \quad (1.1)$$

where the Einstein tensor G_{mn} describing the shape of the universe is determined by the energy momentum tensor T_{mn} . Here, g_{mn} is the metric tensor, R_{mn} is the Ricci tensor found via contraction of the Riemann Curvature tensor, $R_{ab} = R^c_{acb}$, Λ is the cosmological constant, and κ is the constant

$$\kappa = \frac{8\pi G}{c^4} \approx 2.07 \times 10^{-43} \text{ kg}^{-1}\text{m}^{-1}\text{s}^2 .$$

Clearly, adding matter to a universe described by some G_{mn} will change T_{mn} and hence, G_{mn} itself. Thus, to have a relativistically consistent model of the universe to work with, the two sides must be specified simultaneously.

1.2.1 The Schwarzschild Solution

One of the simplest solutions to the field equations was found by Karl Schwarzschild in 1916, shortly after Einstein published his field equations. It is known as the exterior Schwarzschild solution and describes the vacuum that lies outside a spherically symmetric distribution of matter. Generalized to include the cosmological constant, it has the line element:

$$ds^2 = f(r) d(ct)^2 - \frac{dr^2}{f(r)} - r^2 d\Omega^2, \quad (1.2)$$

where

$$f(r) = 1 - \frac{2m}{r} - \frac{\Lambda r^2}{3} \quad (1.3)$$

and $d\Omega^2$ describes the line element on the surface of a sphere,

$$d\Omega^2 = d\theta^2 + \sin^2 \theta d\phi^2 . \quad (1.4)$$

The quantity m is the *geometrized* mass, which has units of length and is given by

$$m = \frac{GM}{c^2} , \quad (1.5)$$

where M is the total mass enclosed by the vacuum, a relationship found by considering the Newtonian limit of the solution to the field equations (see for instance, [Stephani, 1994](#)).

This metric is manifestly static, in agreement with Birkhoff's theorem which states that every spherically symmetric vacuum solution to the field equations must be independent of time. Note that the coordinate time t does not have any deep physical meaning. It is simply the proper time for a hypothetical set of observers sitting in the vacuum at a constant distance from the central mass. The spacelike r coordinate is defined such that the surface area of a sphere specified by $r = \text{constant}$, $t = \text{constant}$ is $4\pi r^2$.

Given the form of $f(r)$ in equation (1.3), it is clear that the metric has a real singularity at $r = 0$. However, this point is contained inside the central mass and so the singularity will not present a problem when working inside the vacuum. The metric is also apparently singular when $f(r) = 0$. In the case where $\Lambda = 0$, this occurs at $r = 2m$, known as the *gravitational* or *Schwarzschild radius* of the source mass. Unlike the one at $r = 0$ though, this singularity is not fundamental, but only appears as a result of the choice of coordinates and can be removed by a coordinate transformation.

1.2.2 The FRW Model

Another solution to the field equations that also satisfies the Cosmological Principal is described by the Robertson-Walker (RW) line element:

$$ds^2 = d(ct)^2 - R^2(ct) \left(\frac{dr^2}{1 - Kr^2} + r^2 d\Omega^2 \right) , \quad (1.6)$$

where R is a *scale factor*, $d\Omega^2$ is the spherical line element given in (1.4), and $K = 0, \pm 1$ is a parameter often referred to as the *curvature indicator*. When $ct = \text{constant}$, the spatial hypersurface is topologically flat in the case of zero curvature when $K = 0$, spherical in shape and topologically closed in the case of positive curvature when $K = 1$, and hyperbolic in shape and topologically open in the case of negative curvature when $K = -1$. A coordinate transformation allows this metric to be written in a more symmetric and convenient form:

$$ds^2 = d(cT)^2 - R^2(cT)[d\omega^2 + \mathcal{S}_K^2(\omega) d\Omega^2] , \quad (1.7)$$

where

$$\mathcal{S}_K(\omega) = \frac{\sin(\sqrt{K}\omega)}{\sqrt{K}} = \begin{cases} \sinh(\omega) & : K = -1 \\ \omega & : K = 0 \\ \sin(\omega) & : K = +1 \end{cases} . \quad (1.8)$$

The dynamics of the universe described by this line element are completely contained in the evolution of the scale factor $R(cT)$. Any observers comoving with the universe remain at fixed spatial coordinates (ω, θ, ϕ) and have a proper time measured by cT .

For RW metrics, the energy momentum tensor must be that of a perfect fluid (see for instance, [Stephani, 1994](#)) so

$$T_{mn} = pg_{mn} + \left(\rho - u^c u_c \frac{p}{c^2} \right) u_m u_n , \quad (1.9)$$

where p and ρ are the pressure and density of the matter respectively, and u^a is the 4-velocity. The expression for G_{00} in the field equations (1.1) is given by

$$R_{00} - \frac{1}{2}R_a^a g_{00} + \Lambda g_{00} = -\kappa T_{00} . \quad (1.10)$$

When considering a matter dominated universe, pressure is typically negligible (see for instance, [Peebles, 1993](#)). The only non-zero component of the energy momentum tensor of the pressureless “dust” at rest in the coordinate system (1.7) is thus $T^{00} = \rho c^2$, so (1.10) gives

$$R_{00} - \frac{1}{2}R_a^a + \Lambda = -\kappa \rho c^2 \quad (1.11)$$

in this system. The 00 component of the Ricci tensor is given by

$$R_{00} = \frac{3\ddot{R}}{R},$$

where for this derivation, $\dot{R} \equiv \frac{dR(cT)}{d(cT)}$, and the Ricci scalar is

$$R_a^a = \frac{6(\ddot{R}R + \dot{R}^2 + K)}{R^2}.$$

Using these in (1.11),

$$\begin{aligned} \frac{3\ddot{R}}{R} - \frac{3(\ddot{R}R + \dot{R}^2 + K)}{R^2} + \Lambda &= -\kappa\rho c^2 \\ \frac{3}{R^2} \left[\left(\frac{dR}{d(cT)} \right)^2 + K \right] &= \frac{8\pi G\rho}{c^2} + \Lambda. \end{aligned} \quad (1.12)$$

This equation is typically referred to as the Friedmann Equation, and cosmological models that have a RW metric and satisfy the Friedmann Equation are collectively known as Friedmann Robertson Walker (FRW) models.¹

If $\Lambda = 0$, the dynamical behaviour of the universe is set by the value of K . In the case that $K = -1$, the universe expands monotonically and in an unbounded fashion. The situation where $K = 0$ is the critical case which also has the universe expanding monotonically, though the expansion is bound, with the expansion velocity approaching zero as the radius of the universe approaches infinity. When $K = 1$, the universe begins by expanding, but reaches a maximum extent and recollapses again. If $\Lambda > 0$, then it is possible to define a critical value,

$$\Lambda_c = \frac{1}{R_0^2} = \frac{\rho_0 c^2}{4\pi G},$$

which produces an unstable static model, the so-called *Einstein universe*. In general though, the presence of a non-zero Λ term does lead to some interesting behaviour in the dynamics of the universe when $K = 1$. For values of $\Lambda < \Lambda_c$, the universe still reaches

¹Models with non-zero Λ terms should technically be referred to as Friedmann Lemaître Robertson Walker (FLRW) models, though this practice seems far from universal.

a maximum extent before turning around and recollapsing. In the case when $\Lambda > \Lambda_c$ though, the topologically closed universe experiences unbound expansion. Interestingly, if Λ is only slightly greater than Λ_c , the universe initially experiences a slowing of its expansion, reaching a quasi-stationary phase before accelerating to expand to infinity. Both the $K = 0$ and $K = -1$ cases continue to expand to infinity with a positive Λ term. Recollapse is possible if this constant were negative, but there is no observational evidence for such an attractive force.

1.2.3 Combining the Solutions

A Swiss-Cheese (SC) universe has as its starting point, an FRW universe uniformly filled with pressureless dust. This universe may be expanding, contracting, or static as desired. Structure is added to the model by inscribing a sphere in this space and replacing the matter it contains with a smaller, spherically-symmetric mass distribution with a higher average density, placed at the centre of the hole created. This results in a model universe that contains three distinct phases – the initial FRW region, a spherically-symmetric core mass distribution, and a vacuum phase in the spherical shell separating the two (which must be the exterior Schwarzschild solution according to Birkhoff’s theorem). Each of these phases is still an exact solution and in fact, can be matched at their boundaries so that the total solution still satisfies the field equations exactly (Einstein and Straus, 1945; Weinberg, 1972; Dyer, 1973). The gravitational field outside the vacuum is not affected by the change made within the hole, and as a result, the solution will remain exact under a uniform expansion (or contraction) of the background universe. Other spheres can then be inscribed in the background FRW region and the mass contained replaced in a similar fashion until some desired fraction of this region has been replaced by a set of vacuum holes, each containing denser spheres at their centre. Though the inscribed spheres may be tangent to one another, as long as they do not overlap, the vacuum regions will remain distinct and the field equations will continue to be satisfied

as the boundaries of the holes evolve with the background universe.

Notice that other than requiring the core to be spherically symmetric, the model does not place any restriction on how the mass inside must be distributed. One may choose to describe it using the Schwarzschild interior solution so that each hole contains the complete Schwarzschild solution. Another option is to use a Tolman solution (Tolman, 1934), filling the core with the pressureless dust found in the background, but at a higher average density. In fact, a uniform distribution of this dust will allow the core to be described by a RW metric, and provided a suitable choice of dynamical parameters is made, the core can itself be described by an FRW model that has been embedded in the background one. This self-similarity gives rise to a natural way of creating substructure in the holes that have already been defined and will indeed be applied later to generate the RSC model used in this thesis.

While the construction as described thus far may seem contrived, with its sharp transitions between matter and vacuum, one should bear in mind that this is only a simplification to make the model easier to work with. Theoretically, there is nothing to prevent one from smoothing the transitions by using thin spherical shells of gradually varying dust density at the boundary of each vacuum region. These can be chosen so that, for instance, the central cores have density profiles that approximate the King models often used to describe galactic mass distributions.

From a general relativistic perspective, the SC model is quite unique in that it is an exact solution to the field equations that can contain substantial inhomogeneities and interesting matter distributions. It is a self-consistent theoretical construct that provides a powerful way of gaining insight into the gravitational lensing described by general relativity. The high degree of symmetry and the conceptual simplicity of its construction make it very appealing from a geometrical standpoint too. As a cosmological model, SC (and, as shall be seen, RSC) adheres to the cosmological principle. It is clearly homogeneous and isotropic on large scales in that it contains no preferred point

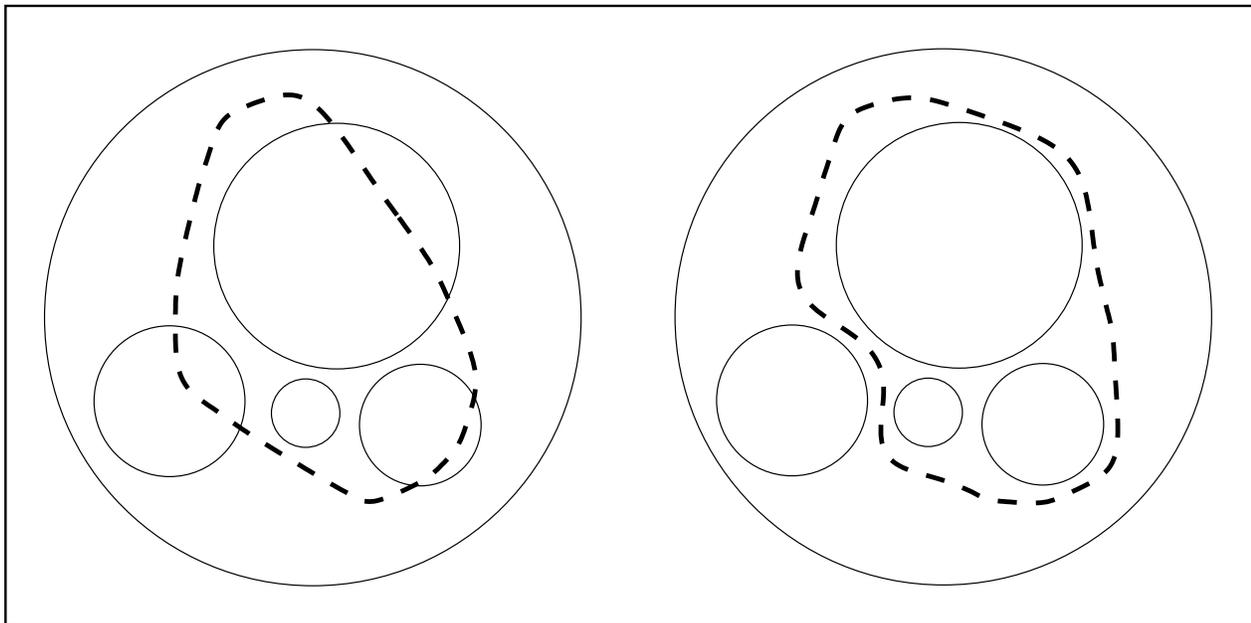


Figure 1.1: With a little deformation, an arbitrary surface placed in a simple SC model (left side) can be made into a surface on which the Weyl tensor is zero (right side).

or direction. From a gravitational perspective, this is equivalent to requiring that one be unable to orient oneself using a gravitational compass. That is to say there exists a surface over which the components of the Weyl tensor vanish. Physically, this means that there exist no tidal forces on this surface. As seen in Figure 1.1, with a little deformation, the model allows one to define such a surface over the entire range of scales which is a remarkable property for a model that can contain a high degree of local inhomogeneity.

1.3 Gravitational Lensing

1.3.1 The Thin Lens Approximation

To investigate the effect that some distribution of matter has on beams of light, one ideally wants to follow these beams through the distribution, keeping track of their cross-sectional area and shape in order to be able to determine what an observer of the beams would measure. Linearized ray-tracing techniques are often used in this regard, com-

monly projecting the mass distribution down into a plane and assuming all bending takes place instantaneously when a ray hits the plane. While relatively straightforward to compute, such a technique has limitations, which is not surprising since the propagation of light is an inherently non-linear process. Problems can arise trying to ensure that flux is conserved, and one generally needs to assume an angular size distance relationship a priori, even though this relationship is dependent on the structure encountered by each individual beam. It also is not clear how best to deal with multiple lensing events with this scheme. Nonetheless, it is worth briefly going over this method as it provides a baseline to which other methods can be compared.

A ray of light passing outside a static, spherically symmetric mass M (a *Schwarzschild lens*), with an impact parameter h , experiences a deflection by the angle

$$\alpha_E = \frac{4GM}{c^2 h} = \frac{4m}{h} ,$$

which is referred to as the *Einstein angle*. This is valid in the weak field limit, where $m/h \ll 1$ and hence, deflection angles are small. If the ray passes through the mass, then $M \rightarrow M(h)$, representing the mass contained within a cylinder of radius h , with an axis parallel to the ray and passing through the centre of the lens (see, for instance [Dyer and Roeder, 1981](#); [Schneider et al., 1992](#)). While the light would actually experience bending due to the mass all along its trajectory, the entire bending is applied at the instant the beam hits the plane onto which the lens has been collapsed (the *lens plane*).

Using simple trigonometry (refer to [Figure 1.2](#) for a schematic), the angular separation of the source and the lens, α_s , can be related to the angular separation of the observed source and the lens, $\alpha_{s'}$, via the relation

$$\alpha_s = \alpha_{s'} - \frac{\alpha_0}{\alpha_{s'}} , \tag{1.13}$$

which is often referred to as the *lens equation*, where

$$\alpha_0 \equiv \sqrt{\frac{4mD_{LS}}{D_{OL}D_{OS}}} \tag{1.14}$$

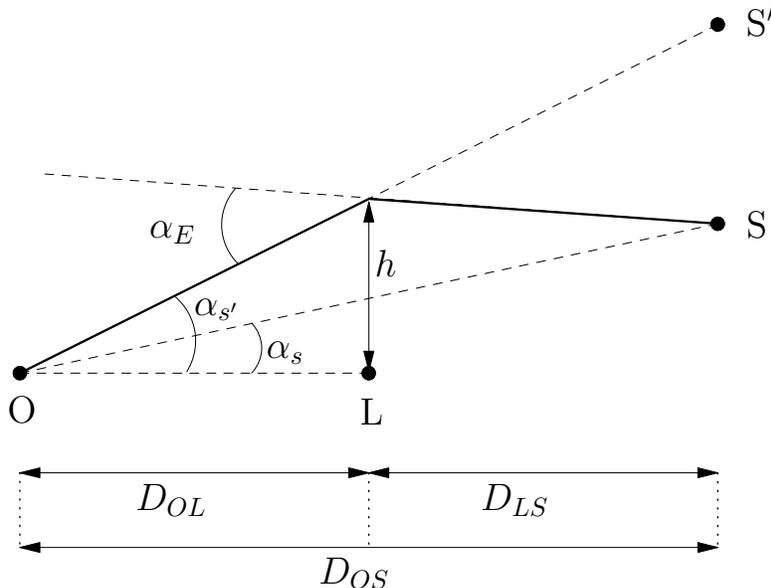


Figure 1.2: The various quantities used in the ray tracing approach. The observer is at O, the lens at L, and the actual source at S, with the apparent source lying at S'. The D values are angular size distances, while α_E is the Einstein angle, and α_s and $\alpha_{s'}$ are the respective angular offsets of the source and apparent source from the lens. The impact parameter h is the distance from the lens where all the bending is assumed to happen.

is a “characteristic angle” of the system (Schneider et al., 1992). The quantities D_{LS} , D_{OL} , and D_{OS} are angular size distances, and can be computed via the Dyer-Roeder relations (see section 5.2.2).

One can consider the lens equation to be a mapping that takes each point within some source of photons (often referred to as the *source plane*) and relates it to one or more points in the observed sky (the *image plane*). The number of images produced from a given source depend on the symmetry and opacity of the lens, along with the impact parameter of the beam under consideration. As a gravitational lens re-directs rays of light and can allow rays that would otherwise never meet to converge, some of these images may be brighter than they would have appeared were there no lens between the source and observer.² Considering the primary (brightest) image produced by the

²Of course the lenses cannot change the amount of radiation leaving a source, so the amplification of a particular source point would only be experienced by a particular observer or set of observers – others would experience a deamplification of the source, or even have photons from it diverted entirely.

lens, the amplification can be written as

$$\mathcal{A} = \frac{1}{4} \left(\frac{\tilde{\alpha}_s}{\sqrt{\tilde{\alpha}_s^2 + 4}} + \frac{\sqrt{\tilde{\alpha}_s^2 + 4}}{\tilde{\alpha}_s} + 2 \right), \quad (1.15)$$

where $\tilde{\alpha}_s \equiv \alpha_s/\alpha_0$ (Schneider et al., 1992). This will be used in Chapter 6 to compare to the amplification computed by following the cross-sectional area of a beam directly.

1.3.2 The Optical Scalar Approach

An alternative approach to investigating gravitational lensing (see, for instance, Dyer, 1973; Harper, 1991), which is the one that shall be adopted for this thesis, is to use the optical scalar equations developed by Sachs (1961). Besides enabling a proper non-linear treatment of the problem, this approach does not require one to specify an angular size distance relation a priori, and is also automatically flux conserving and self consistent. Further, as shall be seen below, when following the optical scalars, the geodesic deviation along the path of the beam is explicitly considered. Hence, for beams that travel along paths in regions of space where the geodesic divergence does not vary too radically from neighbouring paths (which is most of the time), there are fewer beams that need to be followed to get the same information out compared to a ray-tracing approach. The optical scalar approach is also not restricted to small bending angles, and when combined with the RSC model, allows for the treatment of dynamic lenses and beams that experience encounters with multiple lenses during their propagation.

In conventional two or three-dimensional space, a geodesic is easily visualized as the shortest possible curve connecting two points. In a four-dimensional pseudo-Riemannian space things become more complicated as spacetime intervals can be positive (spacelike), zero (null), or negative (timelike). However, while the intuitive notion of a geodesic must be generalized through the use of the variational principle such that

$$\int L d\lambda = \text{extremum} \quad (1.16)$$

for the Lagrangian $L = (ds/d\lambda)^2$ of a curve parameterized by λ , the curves still remain well-defined as the extremal paths between two points and are fundamentally connected to the symmetry of the spacetime under consideration. Using a geometrical optics description, light propagates along null geodesics and can be described as a set of wavefronts or contours of constant phase, everywhere normal to these geodesics. The optical scalar approach involves following the path of a beam which can be thought of as an infinitesimal but finite, bundle of light rays. Though in general there are three optical scalars that can be used to describe the bundle, namely its rate of expansion, shear, and rotation, the geometrical optics approach requires the use of an *irrotational null geodesic congruence*. Physically, this means that the bundle (congruence) of rays is not twisted. If rotation were allowed, then the axis of rotation would generate a caustic on the wavefront where it would not be possible to uniquely specify a normal to the wave and consequently, one would be unable to follow its propagation through space. A SC model does not induce any twisting of the beam so this limitation will not restrict the analysis.

Hence, there are two optical scalars that need to be followed as a beam of light is propagated through a SC model. The rate of expansion, ϑ , simply describes the change in the scale of the beam and is thus a real quantity. The shear rate σ on the other hand, is a complex quantity as it describes the rate of distortion of the beam which has both a magnitude and direction. These two quantities propagate through space according to two coupled, non-linear differential equations, the optical scalar equations (OSE):

$$\dot{\vartheta} + \vartheta^2 + \sigma\bar{\sigma} = \mathcal{R} \tag{1.17a}$$

$$\dot{\sigma} + 2\vartheta\sigma = \mathcal{F}e^{i\beta}. \tag{1.17b}$$

Here, the dot denotes differentiation with respect to an affine parameter along the path of the beam, while the bar indicates complex conjugation. \mathcal{R} is the Ricci driving term which describes the influence of matter that lies within the beam. It results in a uniform expansion (or contraction) of the beam area and is depicted on the left side of Figure 1.3.

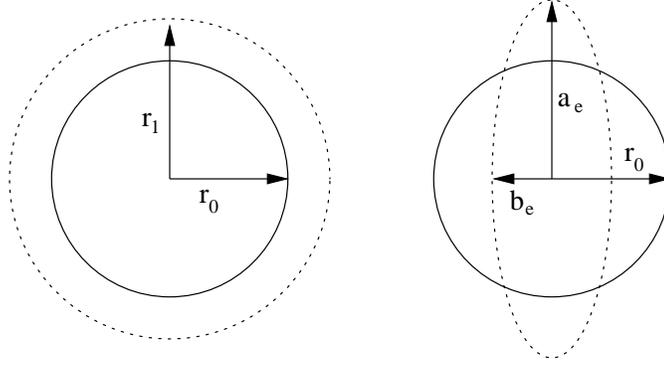


Figure 1.3: In the diagram on the left, the effect of the expansion rate on an initially circular beam is depicted. Here, $r_1 = r_0 (1 + \vartheta d\lambda)$ for some affine parameter λ . The diagram on the right demonstrates the effect of pure shear on an initially circular beam. The semi-axes of the resultant ellipse are given by $a_e = r_0 (1 + |\sigma| d\lambda)$ and $b_e = r_0 (1 - |\sigma| d\lambda)$.

\mathcal{F} is the Weyl driving term which describes the tidal force due to matter lying outside of the beam. Its effect is shown on the right side of Figure 1.3. Notice that to first order, shear alone does not change the area of the beam. To this order, the initially circular beam is distorted to an ellipse and does not become banana-shaped, which is reasonable for a beam with a small cross-section. When considered from a coordinate system that is parallel propagated along the beam, the vector sum of all the tidal fields felt at a given point has an orientation of $\beta/2$ (refer to Figure 1.4).

The terms \mathcal{R} and \mathcal{F} are defined as follows:

$$\mathcal{R} \equiv \frac{1}{2} R_{ab} k^a k^b \quad (1.18)$$

$$\mathcal{F} e^{i\beta} \equiv R_{abcd} k^a k^b \bar{t}^c t^d, \quad (1.19)$$

where R_{abcd} is the Riemann curvature tensor, R_{ab} is the Ricci tensor, k^a is a null tangent vector to the beam path, and t^a is a complex null vector (with \bar{t}^a its conjugate) that is parallel propagated along with the beam such that for the metric signature of -2 chosen,

$$t^a k_a = 0 \quad \text{and} \quad t^a \bar{t}_a = -1 \quad (1.20)$$

(Dyer, 1973). The quantities t^a and \bar{t}^a span the *shadow plane*, which is the plane in which the cross-section of the beam lies and the angle β is defined. The factor of a half in the

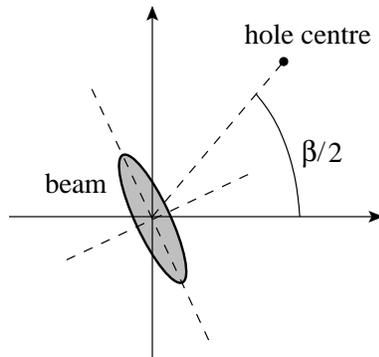


Figure 1.4: A graphical depiction of β as defined in the shadow plane of the beam. The spot labelled “hole centre” is the perpendicular projection of the centre of a lens (the core of a hole) onto this plane. In the RSC model, there is at most one mass distribution providing a tidal force at any given time so the meaning of the angle β is unambiguous. While the axes can have an arbitrary orientation, they must be consistent between holes and can be written in terms of t^a and \bar{t}^a .

orientation of the tidal force arises because the definition of $\mathcal{F}e^{i\beta}$ contains the square of \bar{t}^a . Since \mathcal{F} , R_{abcd} , and k^a are real quantities, it follows that the angle describing the orientation of the complex vector \bar{t}^a must be half the value of β .

Equations (1.18) and (1.19) can be rearranged to make them easier to work with. Since k^a is tangent to the null geodesic followed by the beam, $k^a k_a = 0$. Using this, along with the field equations (1.1) inside (1.18), and replacing the Riemann tensor in (1.19) with the relation

$$R^ab_{cd} = C^ab_{cd} + \frac{1}{2}(g_c^a R_d^b + g_d^b R_c^a - g_c^b R_d^a - g_d^a R_c^b) - \frac{1}{6}(g_c^a g_d^b - g_d^a g_c^b) R_n^n, \quad (1.21)$$

where C^ab_{cd} is the Weyl tensor, we get:

$$\mathcal{R} = -\frac{\kappa}{2} T_{ab} k^a k^b \quad (1.22)$$

$$\mathcal{F} = C_{abcd} k^a k^b \bar{t}^c \bar{t}^d. \quad (1.23)$$

In order to make the OSE suitable for numerical integration, some manipulation is required. There are various approaches that one can take, and three of these are presented below.

The (A, ξ, φ) Form

Perhaps the clearest way to restate the equations is that used by [Dyer \(1973\)](#) and [Harper \(1991\)](#), which relates ϑ and σ to more intuitive physical quantities. Expressing ϑ in terms of A , the cross-sectional area of the beam, and considering only the effect of the Ricci driving term,

$$\delta A = A_1 - A_0 = \pi[r_0(1 + \vartheta d\lambda)]^2 - \pi r_0^2 \quad (1.24)$$

$$\delta A = 2\pi r_0^2 \vartheta d\lambda + \pi r_0^2 \vartheta^2 d\lambda^2, \quad (1.25)$$

where λ is the affine parameter along the beam and r_0 is the initial radius of the circular beam cross-section. Dropping the higher order term in the limit of an infinitesimal beam and rearranging,

$$\vartheta = \frac{\dot{A}}{2A}. \quad (1.26)$$

Next, write σ as

$$\sigma = |\sigma|e^{i\varphi} = \frac{\xi}{A}e^{i\varphi}, \quad (1.27)$$

where ξ is the “scaled” shear rate, a real, positive quantity, and φ is the phase of the shear rate. Using these values in [\(1.17\)](#), and separating the real and imaginary components of the shear rate results in the following system of differential equations:

$$\ddot{A} = \frac{\dot{A}^2}{2A} - \frac{2\xi^2}{A} + 2\mathcal{R}A \quad (1.28a)$$

$$\dot{\xi} = A\mathcal{F} \cos(\beta - \varphi) \quad (1.28b)$$

$$\dot{\varphi} = \frac{A}{\xi}\mathcal{F} \sin(\beta - \varphi). \quad (1.28c)$$

For an unsheared beam with a circular cross-section, $\xi = 0$ and φ is undefined. At the first encounter with a lens, the beam starts to shear with an orientation imposed by the lens so $\varphi/2$ becomes defined with an initial value of $\beta/2$.

The quantities A , ξ , and φ can be used for the numerical integration of the optical scalars as they are well-behaved at caustics where the beam is reduced to a point or

line of zero area (Harper, 1991). From the form of the OSE, it is clear that physically, the behaviour of the system is the same over a caustic regardless of whether A is kept positive and φ changes by π (to describe the rotation of the growing and shrinking axes of the ellipse by $\pi/2$), or if A is allowed to change sign while φ remains continuous. This latter description is preferable from a numerical standpoint since it means the derivative \dot{A} changes smoothly across the caustic. A negative area should not be of any concern. In a sense, the beam is turned inside-out as it passes through a line caustic and one need simply take the absolute value of A whenever a physical area is desired. Note that in this case, equation (1.27) would need a slight modification as the ratio ξ/A would no longer always be positive, with the change in sign for σ across a caustic now coming from the area instead of the phase term.

The (C_{\pm}, α_{\pm}) Form

Another way of expressing the OSE that is suited to situations containing a high degree of symmetry is due to Kantowski (1968), and has been used by Dyer and Roeder (1981) and Dyer (1986) among others. He defines four real parameters, C_{\pm} and α_{\pm} , which are related to the *principal curvatures*³ of the wavefront, and their directions respectively. They can be written in terms of the optical scalars ϑ and σ via

$$\vartheta \pm \sigma = \frac{\dot{C}_{\pm}}{C_{\pm}} + i\dot{\alpha}_{\pm} \quad (1.29)$$

Using these in (1.17), with the restriction that the expansion rate ϑ be real, yields the expressions

$$\frac{\ddot{C}_{\pm}}{C_{\pm}} + \dot{\alpha}_{\pm}^2 = \mathcal{R} \pm \mathcal{F} \cos(\beta) \quad (1.30a)$$

$$\ddot{\alpha}_{\pm} + 2\dot{\alpha}_{\pm} \frac{\dot{C}_{\pm}}{C_{\pm}} - 2\sigma_r \sigma_i = \pm \mathcal{F} \sin(\beta) \quad (1.30b)$$

where σ_r and σ_i are the real and imaginary components of σ respectively. For a particular choice of orientation of the system in which the directions are expressed, $\dot{\alpha}_+ = -\dot{\alpha}_-$.

³The principal curvatures are the extrema in curvature of the wavefront.

Writing this quantity as simply $\dot{\alpha}$, equations (1.30) become

$$\frac{\ddot{C}_{\pm}}{C_{\pm}} + \dot{\alpha}^2 = \mathcal{R} \pm \mathcal{F} \cos(\beta) \quad (1.31a)$$

$$\ddot{\alpha} + \dot{\alpha} \left(\frac{\dot{C}_+}{C_+} + \frac{\dot{C}_-}{C_-} \right) = \mathcal{F} \sin(\beta) \quad (1.31b)$$

as found by Dyer (1977). Further, for an initial value of $\dot{\alpha} = 0$ as would be the case for a non-shearing beam, if the orientation of the driving term can be set such that $\beta = 0$, then the system (1.31) reduces to the very simple form:

$$\ddot{C}_{\pm} = (\mathcal{R} \pm \mathcal{F})C_{\pm}, \quad (1.32)$$

which is valid for a single lens encounter.

The (ψ, η) Form⁴

This formulation, which is the one ultimately used to track the distortion of a beam through the RSC model, follows the procedure outlined by Pineault (1975). As viewed in the shadow plane, the location of a point on the wavefront of the beam being followed (i.e. indicating a particular geodesic) can be described by a complex vector ζ . As shown by Penrose (1968), when the basis used to span the shadow plane is parallel propagated along with the beam, the evolution in ζ due to expansion and shear is given by

$$\dot{\zeta} = \vartheta\zeta + \sigma\bar{\zeta}, \quad (1.33)$$

where the terms on the right hand side are positive instead of negative as Penrose had them in order to be consistent with the sign conventions being used in this thesis.⁵

In order to describe the generally elliptical nature of the beam cross-section as opposed to a single point on the wavefront, ζ can be written as

$$\zeta = \psi e^{i\varpi} + \eta e^{-i\varpi}, \quad (1.34)$$

⁴This is commonly referred to as the (ξ, η) form, but a change from ξ to ψ has been made to avoid confusion with the previous use of ξ for the OSE in this thesis.

⁵Strictly, Penrose was considering the convergence of a congruence of geodesics, as opposed to the expansion which is of interest here, but the sign difference accounts for this.

where ψ and η are complex values, and ϖ is a dummy variable ranging from 0 to 2π to cover each point on the boundary of the ellipse. Substituting (1.34) into (1.33) and equating the real and imaginary parts separately,

$$\dot{\psi} = \vartheta\psi + \sigma\bar{\eta} \quad (1.35a)$$

$$\dot{\eta} = \vartheta\eta + \sigma\bar{\psi} \quad (1.35b)$$

Taking the derivatives of these again and using the OSE (1.17) to simplify the results,

$$\ddot{\psi} = \mathcal{R}\psi + \mathcal{F}e^{i\beta}\bar{\eta} \quad (1.36a)$$

$$\ddot{\eta} = \mathcal{R}\eta + \mathcal{F}e^{i\beta}\bar{\psi} \quad (1.36b)$$

Writing ψ and η in terms of real and imaginary components,

$$\psi = x + iy \quad \text{and} \quad \eta = z + iw . \quad (1.37)$$

Substituting them into (1.36) and separating the real and imaginary parts, we get

$$\ddot{x} = \mathcal{R}x + \mathcal{F}(z \cos \beta + w \sin \beta) \quad (1.38a)$$

$$\ddot{y} = \mathcal{R}y + \mathcal{F}(z \sin \beta - w \cos \beta) \quad (1.38b)$$

$$\ddot{z} = \mathcal{R}z + \mathcal{F}(x \cos \beta + y \sin \beta) \quad (1.38c)$$

$$\ddot{w} = \mathcal{R}w + \mathcal{F}(x \sin \beta - y \cos \beta) . \quad (1.38d)$$

These equations are suitable for numerical integration and in agreement with the expressions derived by Pineault (1975), though along with the vacuum regions he considered, they are also applicable in regions with a non-zero Ricci driving term.

1.4 Previous Work

The Swiss-Cheese model of the universe was initially developed by Einstein and Straus (1945) as a means of investigating the gravitational field of a mass well described by a

Schwarzschild solution, but that exists in background universe that is not Minkowskian. [Kantowski \(1969\)](#) used the optical scalar equations and an SC model to consider variations in the luminosity-redshift relationship caused by the introduction of clumps of matter into FRW models. [Dyer \(1973, 1976, 1977\)](#), and [Dyer and Roeder \(1974, and references therein\)](#) also used the optical scalars to consider the gravitational lensing effect of an opaque object at the centre of an SC hole, and analyzed the implications on photon propagation by considering interactions with holes in a stochastic fashion. [Oattes \(1987\)](#), and [Dyer and Oattes \(1988\)](#), performed similar calculations, but allowed the size of the opaque region to vary in an arbitrary way, leading to the concept of a “fuzzy” past null cone. As a beam of light propagates past a hole, a deformation is introduced into the null cone it defines. The more encounters with inhomogeneities that there are, the greater the net deformation becomes. Consequently, the null geodesics that describe the boundary of the null cone are no longer simple straight lines, but can curve in an irregular fashion. If the exact geometry of the previous encounters is uncertain, one cannot give a precise description of the null cone but can only compute a probability that these geodesics will remain bound within a given vicinity of the unperturbed geodesics. The region so defined can be interpreted as giving the past null cone a finite thickness (or fuzziness). [Harper \(1991\)](#) did a proper thick-lens calculation to compute the propagation of light through a hole in more detail. An analysis was done using both stochastic and limited predetermined distributions of holes, but used a single SC level with non-evolving cores.

This work describes the development of a Recursive Swiss-Cheese cosmological model which has been used, along with the optical scalars, to develop software that allows for the investigation of the gravitational lensing effects produced by evolving lenses in a relativistically consistent manner. [Chapter 2](#) describes SC packings and how they are combined to generate the full RSC model. A description of the process used to follow the beam trajectories through the various regions of the model is given in [Chapter 3](#).

Particular care is taken to describe the procedure used to define the substructure in the model, and to explain the matching at the interface between the FRW and Schwarzschild sections. In Chapter 4, the relevant quantities necessary to propagate the optical scalars through the RSC model are derived, and the means of extracting the observable quantities is discussed. Chapter 5 describes some practical aspects of the implementation of the lensing code, while Chapter 6 presents various results generated by the code to demonstrate its capabilities and to establish that the output it produces is reasonable and in agreement with what can be expected.

Chapter 2

The RSC Model

2.1 The SC Packings

As indicated previously, at its most basic level, generating an SC model involves taking a region of space and inscribing spheres inside it in order to create structure. The model itself does not specify where these spheres must be placed. The locations may be chosen purely at random, or they may correspond to some predetermined mapping, to emulate structure that is observed in the universe or found via processes such as N-body simulations, for instance. For the current work, it was decided to create sphere distributions based on an efficient volume-filling scheme that packed spheres into a spherical region of space. By choosing the optimal sphere size to place in a given region, the packing algorithm determines how much of the available mass in that region will be put into the structure inside the sphere. While the efficient packing of spheres is an interesting problem in and of itself (with Apollonian packing being an active field of mathematical study that dates back to the time of the ancient Greeks), and though the structure eventually created has some astrophysical relevance, one should bear in mind that there is nothing special about such a choice and the SC model is flexible enough to accommodate many different possibilities.

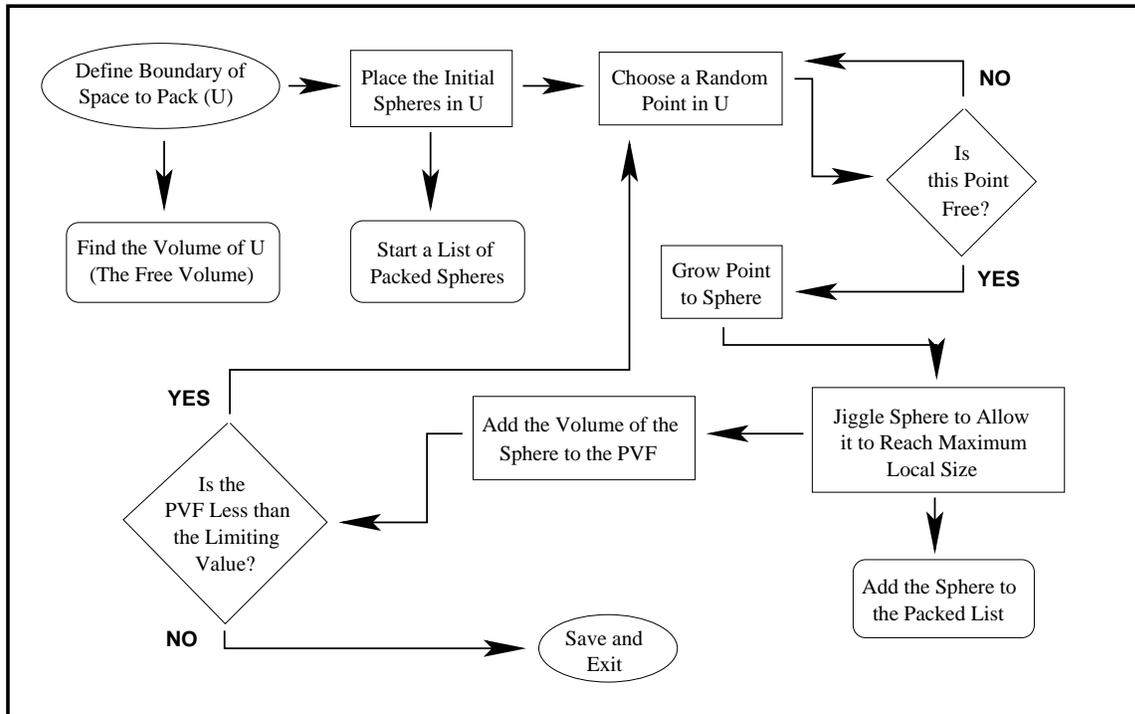


Figure 2.1: The general flow of the code used to generate a single-level Swiss-Cheese model.

2.1.1 Generating the Packings

The development of code to pack spheres into a three-dimensional region of space in a random, yet efficient manner has been described in [Attard \(1997\)](#). A general outline of the procedure is given in [Figure 2.1](#). The program first defines a spherical region of FRW space (U) extending out to some size ω_{\max} from the origin. This space can be of positive, negative, or zero curvature. In the cases of non-zero curvature, ω is an angular measure, while in the flat case it is a dimensionless linear quantity. When $K = +1$, this coordinate is cyclic and the space reaches a maximum extent at $\omega = \pi/2$. Regardless of the curvature, the size is given a physical dimension via the scale factor R .

The volume of U , which is the space available for packing (the free volume), is computed and stored. Next, a sphere of random size is thrown down at an arbitrary point inside U and placed at the first node of a “packed” list used to store all the spheres that the program can fit into U . The volume of this sphere is computed and subtracted from

the free volume. More spheres are then thrown down at random, subtracted from the free volume, and added to the packed list, until one of them overlaps with another. At this point, the overlapping sphere is removed and the actual packing begins.

The above procedure establishes random initial conditions inside U for the efficient packing routine to work with. This routine first selects a random point within U . If this point lies inside a sphere on the packed list, it is discarded and another is chosen. Once a suitable “free” point has been found, it is “inflated” until it makes contact with the boundary of U , or the boundary of one or more spheres on the packed list. In this way, the new sphere is given the largest possible radius that it can have without overlapping any of the other spheres present. This new sphere is then given a “push” away from its point(s) of contact, and the maximum radius that it can have at the new location is determined. If this turns out to be less than the value in the previous location, the old location and radius are added to the next node on the packed list and the volume of the sphere is subtracted from the free volume. Otherwise, the new location and larger radius are adopted, and the process of inflating and pushing the sphere is repeated. Pushing the sphere around in this way maximizes the volume used around every valid random point, making the packing more efficient. The process terminates when the total volume occupied by the packed spheres becomes larger than some predefined fraction of the total volume of U . The use of a limiting packed volume fraction (PVF) is necessary to keep the program running time reasonable since in practice, it is not possible to fill U completely.

The code developed was used to generate a library of sphere *packings* for each of the spatial curvatures using a few different values for ω_{\max} . Each packing is a simple list of spheres, each of which is described by four numbers, the triple (ω, θ, ϕ) specifying the location of the centre of the sphere, and a radius. These are all comoving quantities in the FRW model and are thus valid regardless of the value of the scale factor. The library generated is summarized in Table 2.1. The limiting PVF was chosen to be 80% for most of the packings as beyond this the free space in U was found to be sufficiently fractured

ω_{\max}	K	Packings	Spheres Contained	(Mean)	Radius Range	PVF
$\pi/10$	-1	50	42 007 – 132 476	(68 973)	2.7e-8 – 0.21	80.0
	0	69	34 654 – 140 564	(65 898)	9.4e-9 – 0.20	
	+1	51	14 026 – 146 172	(62 019)	4.0e-8 – 0.22	
$\pi/3$	-1	6	77 144 – 131 874	(105 421)	8.3e-7 – 0.54	80.0
	0	7	23 918 – 108 169	(70 485)	2.3e-7 – 0.70	
	+1	6	31 625 – 63 042	(42 805)	4.1e-6 – 0.51	
0.49π	-1	20	146 646 – 354 012	(227 109)	8.0e-9 – 0.87	80.0
	0	20	8762 – 121 953	(67 786)	1.3e-7 – 1.18	
	+1	20	12 504 – 44 506	(25 112)	3.7e-7 – 0.84	
2π	-1	20	189 938 – 242 822	(229 479)	3.3e-5 – 3.51	33.3

Table 2.1: A summary of the packing library generated. The first column indicates the radial extent of the space filled. Next is the curvature indicator, which is followed by the number of packings generated, then a range indicating the number of spheres contained in the packings (with the average number in brackets), the size range of the spheres packed, and finally the volume fraction of the space filled by the spheres is given in the last column.

that the packing process continued very slowly. In the case of $\omega_{\max} = 2\pi$, the limit was lowered to a third of U in order to maintain reasonable running times. While there is a significant variation in sphere counts for individual packings with a particular curvature and extent, the average varies in the expected way since a given ω_{\max} implies the largest volume when $K = -1$ and the smallest when $K = +1$ (refer to Appendix A).

While the packing algorithm discussed is effective, as the number of spheres placed in U increases, so does the time required to place each additional sphere. This is particularly noticeable in spaces with negative curvature, where a given increase in the radius of U leads to a proportionally larger increase in volume than in the open and closed cases. To address this issue, a modification was made to the way that the code tries to maximize the volume of the spheres it is trying to pack for those packings with $\omega_{\max} = 2\pi$. It still chooses a random free point for the centre of the new sphere and goes through all the spheres that have already been packed to determine the non-overlapping contact distance as was done before. In the process though, it also stores the distance required

to contact the next closest sphere. The new sphere is then moved radially away from the sphere it contacts by an amount equal to half the difference between these two distances, thus ensuring that no overlap results. This move is very quick computationally as it requires little more than placing the contacted sphere at the centre of the coordinate system. Provided the centre of the new sphere is not in between, and collinear with, the two closest spheres, the new sphere can be grown by more than the minimum amount described so the process can be repeated to fill the space more effectively. Once the gain in radius with each move falls below some threshold, the radius of the new sphere is fixed and the process is repeated with a new free random point.

This simplified scheme is significantly quicker than the full one described previously, with the speed increasing by about an order of magnitude initially. However, the difference decreases as the space becomes more filled and the determination of the contact distance to each of the packed spheres becomes increasingly expensive. Furthermore, the packings produced by the quicker method are less efficient than those created in the original way, requiring 10-20 percent more spheres to fill a third of the space when $\omega_{\max} = 2\pi$. As such, when working with large regions of negatively curved space, it may be most effective to use the original packing scheme, but modified so that the code only checks distances to all the packed spheres once, then considers just the subset of potential contacts when moving the new sphere around. This subset might be determined by comparing the distance of each sphere from the chosen random point with the total distance that the new sphere has been moved in the maximization process.

2.1.2 Packing Structure

Projected Distributions

The easiest way to look at the distribution of the spheres in a packing is by representing each one with a point placed at the angular coordinates corresponding to its centre.

Figure 2.2 shows one such distribution, plotting the centres of 197 944 spheres packed into a space of negative curvature with a maximum angular extent of $\omega_{\max} = 0.49\pi$. This limiting angular extent explains the sharp cut-off in the distribution at this value of ω . While the spheres span the range in θ at the cut-off, this is not the case at the other end for small values of ω . The spheres do not extend down to $\omega = 0$, and there are two noticeable gaps appearing in the distribution, one a roughly triangular-shaped region in the lower part of the graph, and the other an arc-shape at the top. This empty space is due to a packed sphere that encompasses the origin of the coordinate system. It is represented by the dot at $(\omega, \theta) \approx (0.27, 0.52)$, and has a radius of about 0.39 radians. Recall that the region of space specified by each value of ω is spherical in shape since θ and ϕ can vary over their full ranges there. The smaller the value of ω , the smaller the region, which is why any sphere encompassing the origin would create such a significant region of exclusion in a graph of this nature. This is also the reason that the density of points increases on average as one moves to larger ω values – there is simply more space available there. Similarly, as one scans across the range of θ values, one finds that the density of points generally decreases towards the extrema since these also represent poles in the coordinate system.¹

The rapid increase in the density of points right near ω_{\max} means that there have been more smaller spheres packed into the region near the boundary of U. This is not unexpected as larger spheres are forced away from the boundary in order to remain within the space under consideration. While this edge effect breaks some of the complex symmetry of the distribution, it does not create a preferential radial direction in the packings and does not alter the validity of the model.

Figure 2.3 shows the same packing used in Figure 2.2, but plots the distribution in ϕ over the angular extent of the packing. As happened with the previous figure, the point

¹Though in this case, only the ϕ coordinate is degenerate as opposed to both θ and ϕ being degenerate when $\omega = 0$, thus explaining why spheres placed near these poles do not have such a dramatic impact on the distribution.

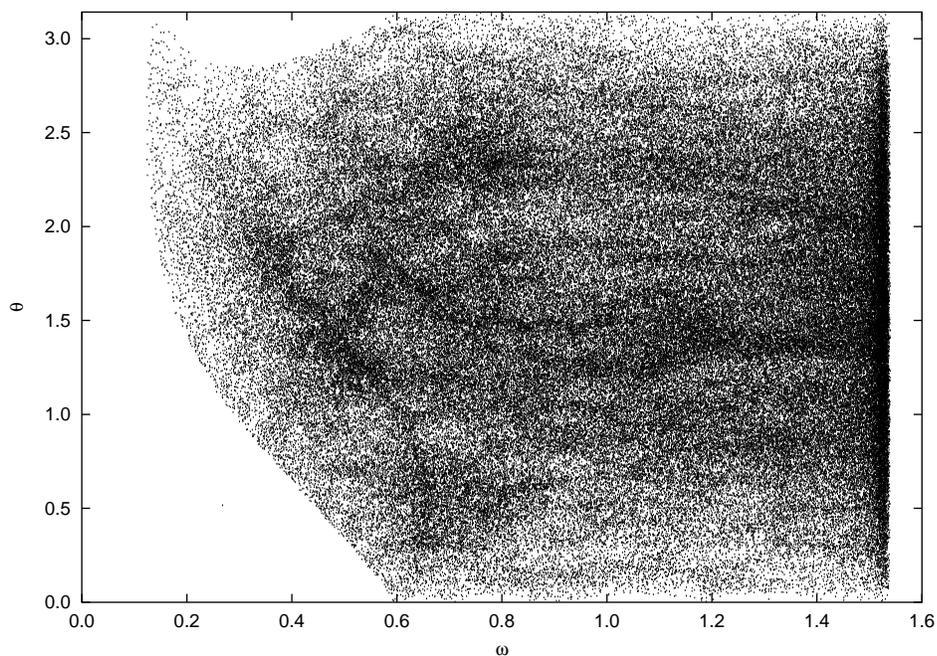


Figure 2.2: The distribution of sphere centres in ω and θ coordinates for a packing containing 197 944 spheres in a negatively curved space with a maximum angular extent of $\omega_{\max} = 0.49\pi$.

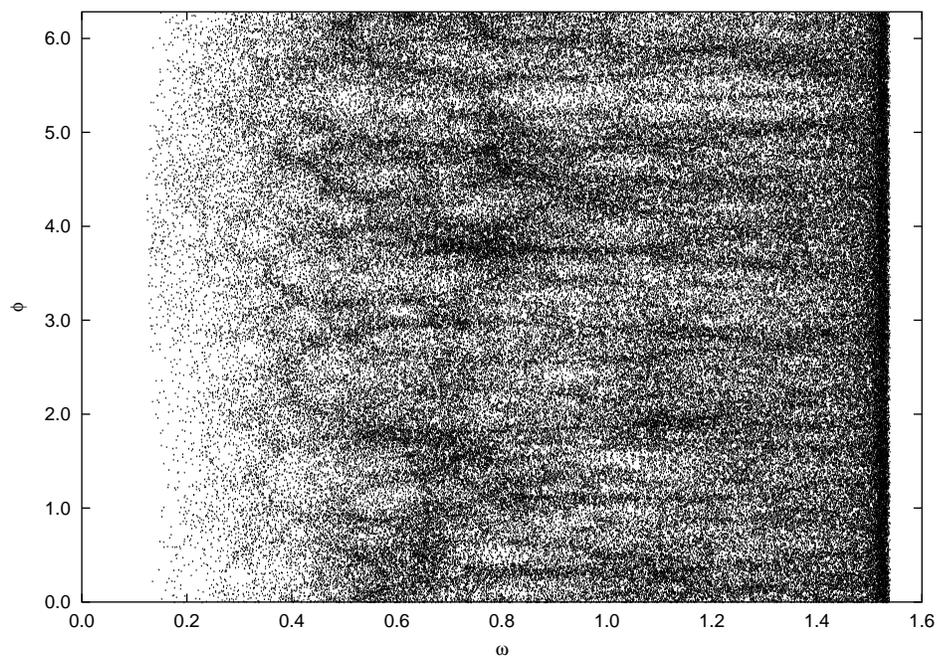


Figure 2.3: The distribution of sphere centres in ω and ϕ coordinates for the same packing displayed in Figure 2.2, with 197 944 spheres, $K = -1$, and $\omega_{\max} = 0.49\pi$.

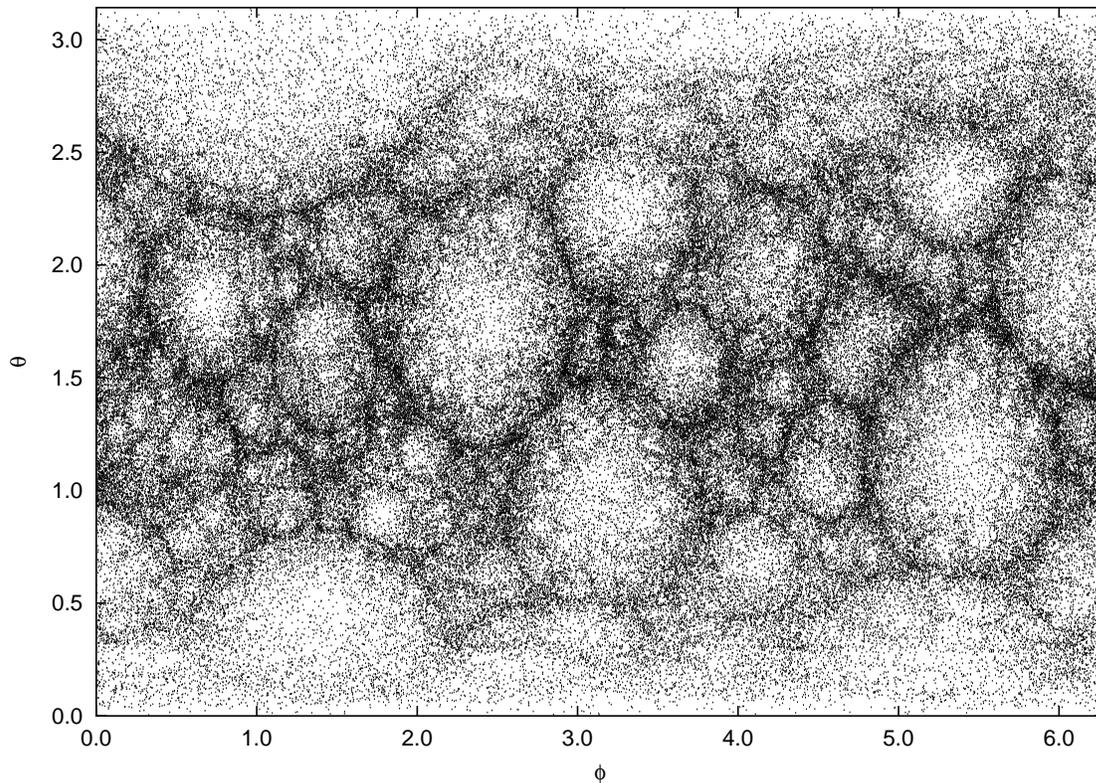


Figure 2.4: The distribution of sphere centres in θ and ϕ coordinates for the same packing displayed in Figures 2.2 and 2.3, with 197 944 spheres, $K = -1$, and $\omega_{\max} = 0.49\pi$.

density increases with ω . However, unlike the case with θ , there is no general trend over the range in ϕ since this coordinate does not define any poles in the system. The sphere that had such a large effect on the $\omega - \theta$ distribution is located at $(\omega, \phi) \approx (0.27, 0.87)$ in this graph. It still causes a gap in the distribution for $\omega \lesssim 0.12$, the extent to which it completely surrounds the pole, but it is clear that smaller spheres were packed around it where they could be fit in.

Perhaps the most interesting way of displaying the distribution in angular coordinates is shown in Figure 2.4, which plots θ against ϕ to produce a “sky projection”. The most striking feature apparent is that one now sees a number of spherical regions in the image corresponding to the locations of some of the larger spheres in the packing. In fact, the largest five spheres in the packing, all with a radius above 0.5 radians and with coordinates given in Table 2.2, are clearly evident in the image. Fewer points appear in these regions

ω	θ	ϕ	Radius
0.9669866	1.7661134	2.3997707	0.5720132
0.9721713	1.1779552	5.3974663	0.5667388
1.0033324	0.9948801	3.1626609	0.5358855
1.0221059	2.6853885	1.0160040	0.5170304
1.0226814	1.9399560	6.0845352	0.5166990

Table 2.2: The radius and centre location of the five largest spheres in the packing used to generate Figures 2.2 to 2.4.

because the large spheres create significant exclusion zones in that area of the “sky”. They are well-defined in this projection because each point in the diagram² represents a unique line of sight, while previous figures provided only a rotationally-collapsed radial profile, thus distorting and obscuring the distributions. The smaller spheres that are squeezed into the spaces between the larger spheres creates a filamentary structure in projection. There are fewer, more distorted spheres towards the poles in θ as can be expected, and it is clear that the structure at $\phi = 0$ is connected to that at $\phi = 2\pi$ as must be the case.

Figure 2.5 shows the θ versus ϕ distribution for four other packings. Both depicted in the top row were also made in a negatively curved space, though the one on the left had a smaller angular extent than the one used previously, with $\omega_{\max} = \pi/10$, while the other was larger with $\omega_{\max} = 2\pi$. The one on the left bears a greater resemblance to Figure 2.4, showing much the same structure, though not as well defined on account of the fact that it has only about a third the number of spheres of the other. On the other hand, though the packing on the right contains almost as many spheres as that in Figure 2.4, there is almost no structure apparent. There are a few small patches of lower-density present, along with the expected general density gradient in θ , but no voids have been clearly defined. There are two effects at work here. First of all, as is pointed

²Except for the poles at $\theta = 0, \pi$ of course.

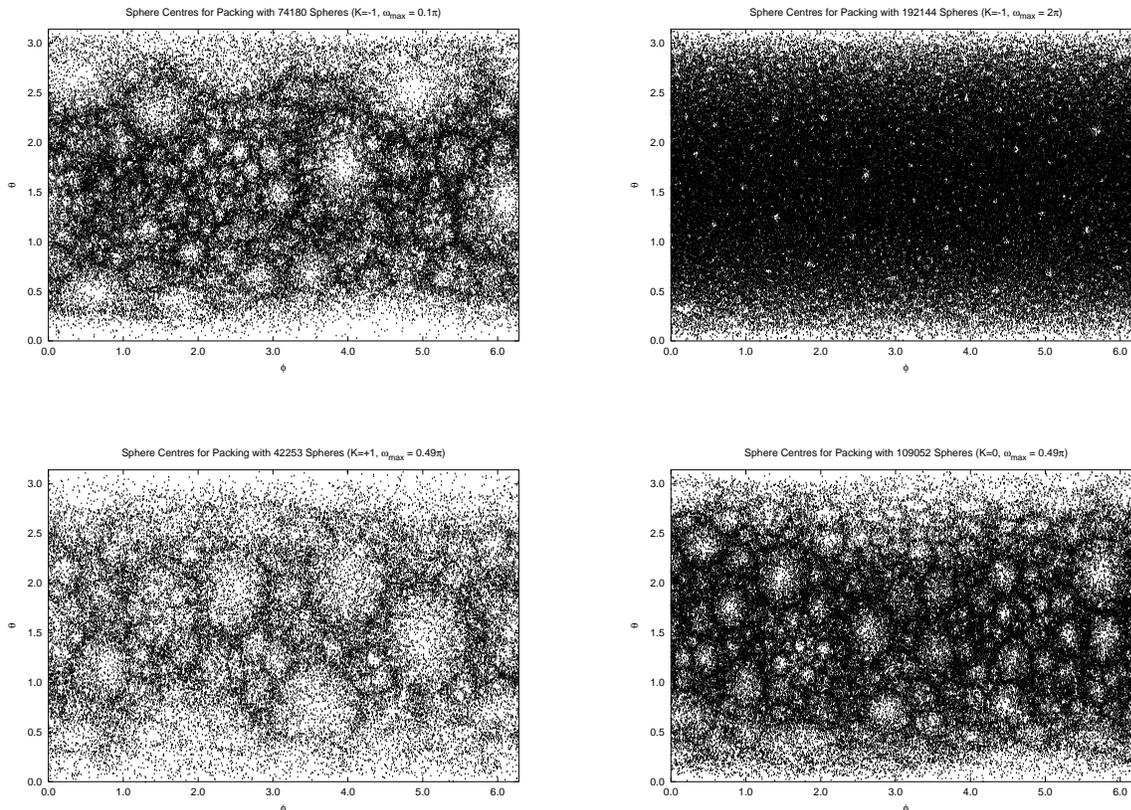


Figure 2.5: The distribution of sphere centres in θ and ϕ coordinates for four different packings. The top row still deals with negatively curved spaces, but varies their angular extent, while the bottom row uses $\omega_{\max} = 0.49\pi$ as was done previously, but shows the distribution for a packing in a positively curved space on the left side and a flat one on the right.

out in Table 2.1, this packing only occupies 33.3% of the region of space it is contained in, unlike all the other packings depicted which occupy 80% of the volume. As explained in Section 2.1.1, the packing algorithm attempts to fill the given region of space with the fewest spheres by maximizing the size of the sphere it can place in any section of free space that it finds. With less than a third of the space having been filled, this means that the packer is still able to place relatively large spheres which are more uniformly distributed through the space than the smaller spheres which tend to cluster in the spaces between the larger ones. The other factor is the larger angular extent itself. As seen in Appendix A, when one increases the angular extent of a space with negative curvature, the volume of the space grows exponentially quickly. Consequently, each line of sight

contains a greater projected depth which serves to dilute the effect of a given void on the projected distribution, further contributing to the reduction in apparent structure.

The two packings shown in the bottom row of Figure 2.5 were both made in spaces that extended out to $\omega_{\max} = 0.49\pi$, but do so in different curvatures, with the one on the left filling 80% of the volume of a positively curved space, while that on the right fills this fraction of a flat space. Both display structure similar to that shown for the negative space, though it is not as well defined due to the smaller number of spheres in each case here.

An alternative way of visualizing the structure inside the packings is to view their projection onto a two-dimensional flat surface. Refer to Appendix C.1 for details regarding the conversion between angular and Cartesian coordinates. Note that in all cases shown here, the values have been computed with R set to 1. Figure 2.6 contains a couple of Cartesian projections for three different packings, each made in a space of different curvature, though all of which extended out to $\omega_{\max} = \pi/10$. The projections in the top row correspond to the packing for which $K = +1$ and involve the w axis, the fourth spatial dimension required for the embedding. Neither the $w - x$, nor the $w - z$ planes plotted (and the same holds true for the $w - y$ plane not displayed) show much in the way of structure, though this is not surprising considering the distortion implied by the axial scales. The only significant trend of note is that the points seem more densely clustered along the strip at the smallest values of w , which is in agreement with the earlier finding that there are more small spheres that are fit into the region near the boundary of the space. The second row contains projections of a packing in a space where $K = 0$. There is no artificial embedding dimension in this case so the projections are circular as expected. The lack of distortion makes the structure easier to see, with the voids caused by the larger spheres being quite apparent.

The packing in the bottom row was made in a negatively curved space. The projection on the left is the same plane as the upper left one involving positive curvature, though

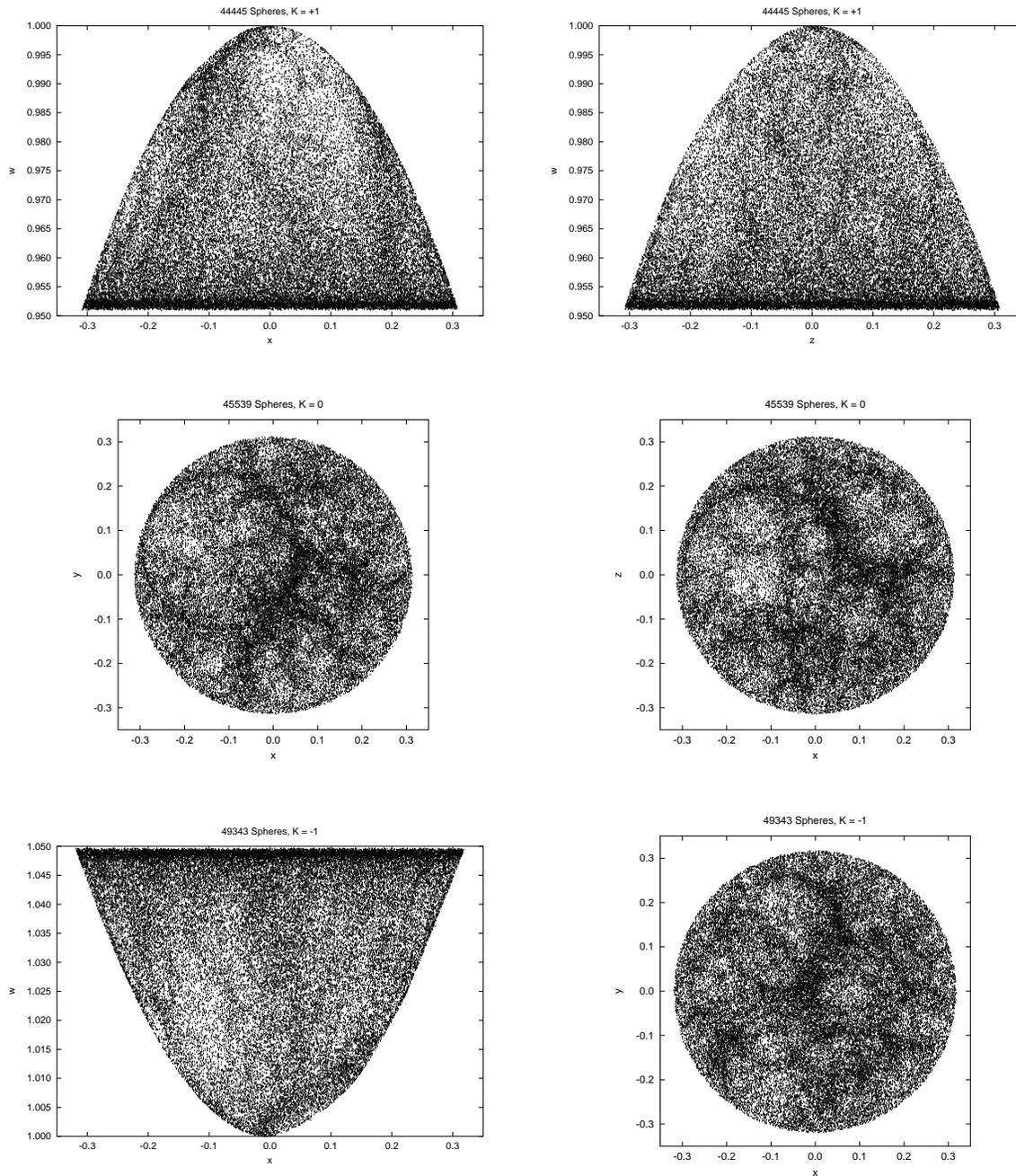


Figure 2.6: Various Cartesian projections of three packings, all of which filled 80% of the volume of a space with $\omega_{\max} = \pi/10$. The space in the top row had positive curvature, that in the middle was flat, while the packing in the bottom row was made in a negatively curved space.

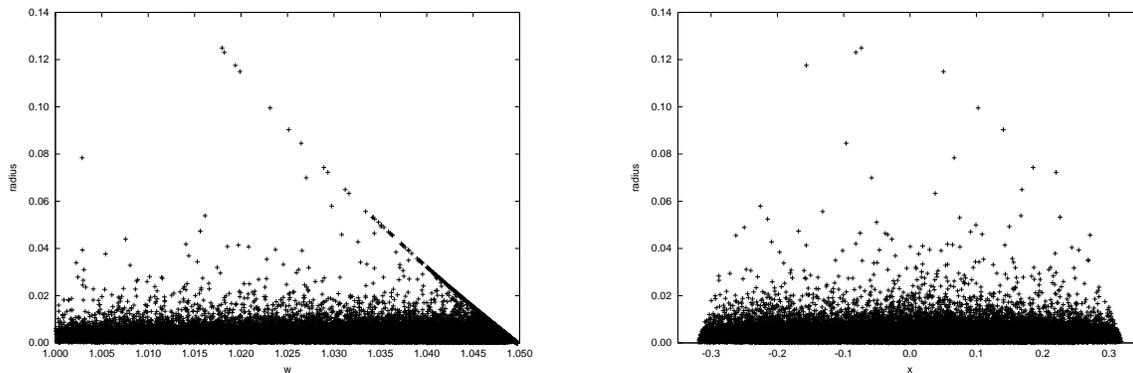


Figure 2.7: The distribution of sphere radii for the packing that was shown in the bottom row of Figure 2.6 (i.e. 49 343 spheres, $K = -1$, $\omega_{\max} = \pi/10$).

in this case the volume extends up from the centre at $w = 1$. Again, the points are distributed more densely along the boundary of the region. The projection onto the $x - y$ plane on the right is circular, as projections not involving w should be regardless of the curvature, and also has structure apparent. In fact, the structure in such projections for packings made with all three curvatures appears to be quite similar. Figure 2.7 shows the distribution of radii for this packing, with the plot on the left showing the variation along the w axis while the one on the right graphs the variation along the x axis. As is apparent in either of them, most of the spheres in the packing are small. While the majority of the spheres in the plot on the left seem to be fairly evenly distributed over the region, there is a set that form a definite boundary running diagonally upward from the bottom right side of the graph. These are the largest spheres that contact the boundary of the region for a given value of w . Those near the outer limit in w are necessarily small, but they get larger as one moves to smaller w and the centre of the space. The distribution along the x axis is more random, though the largest spheres can still be seen to fall near the centre of the range.

Instead of simply using a point to represent the location of the centre of each packed sphere, it is also possible to portray each of them as an opaque ball and plot what an observer looking at the packing from the outside would see. Figure 2.8 shows spheres

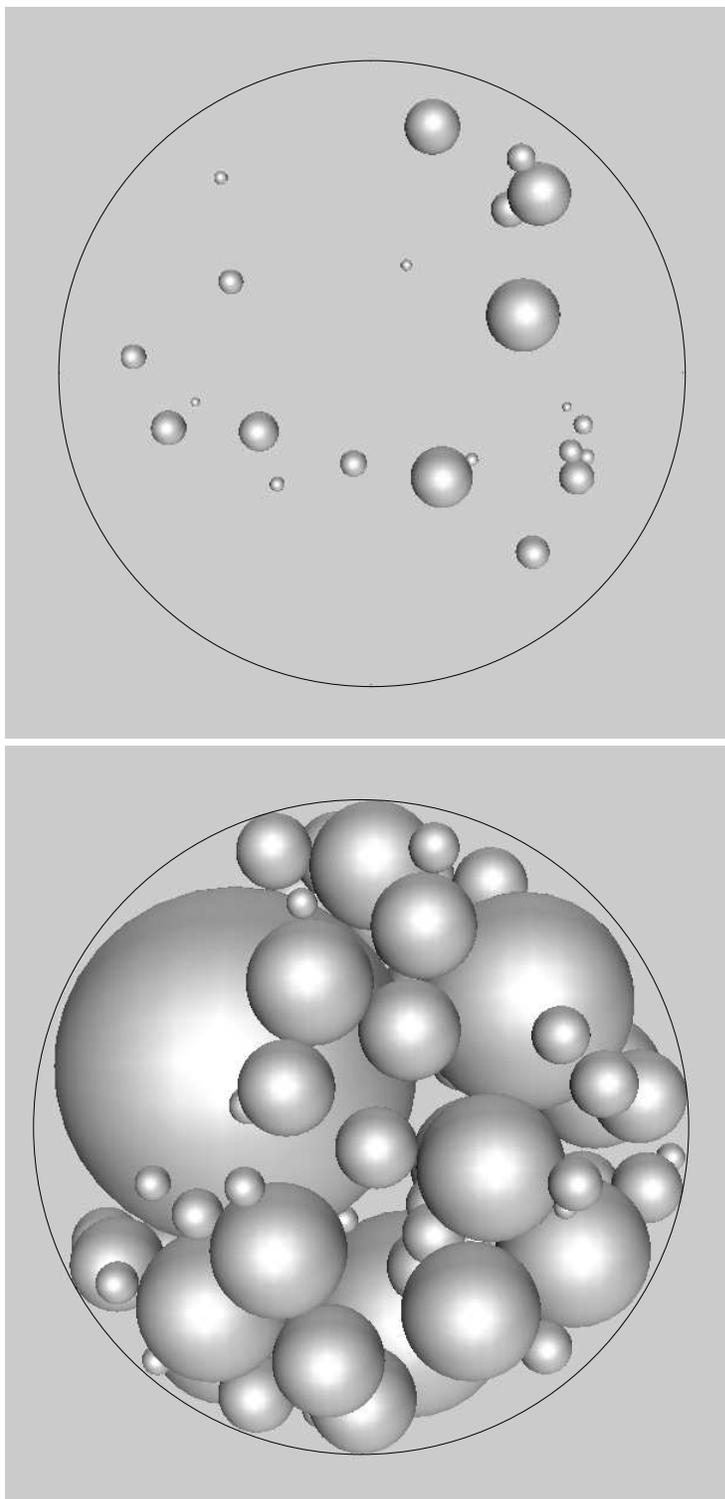


Figure 2.8: In this diagram, opaque balls have been used to represent the spheres of the packing shown in the middle row of Figure 2.6. The first 23 randomly placed spheres are shown in the top image, while the bottom one contains the first 100 that were packed. The view looking down onto the $x - y$ plane is depicted.

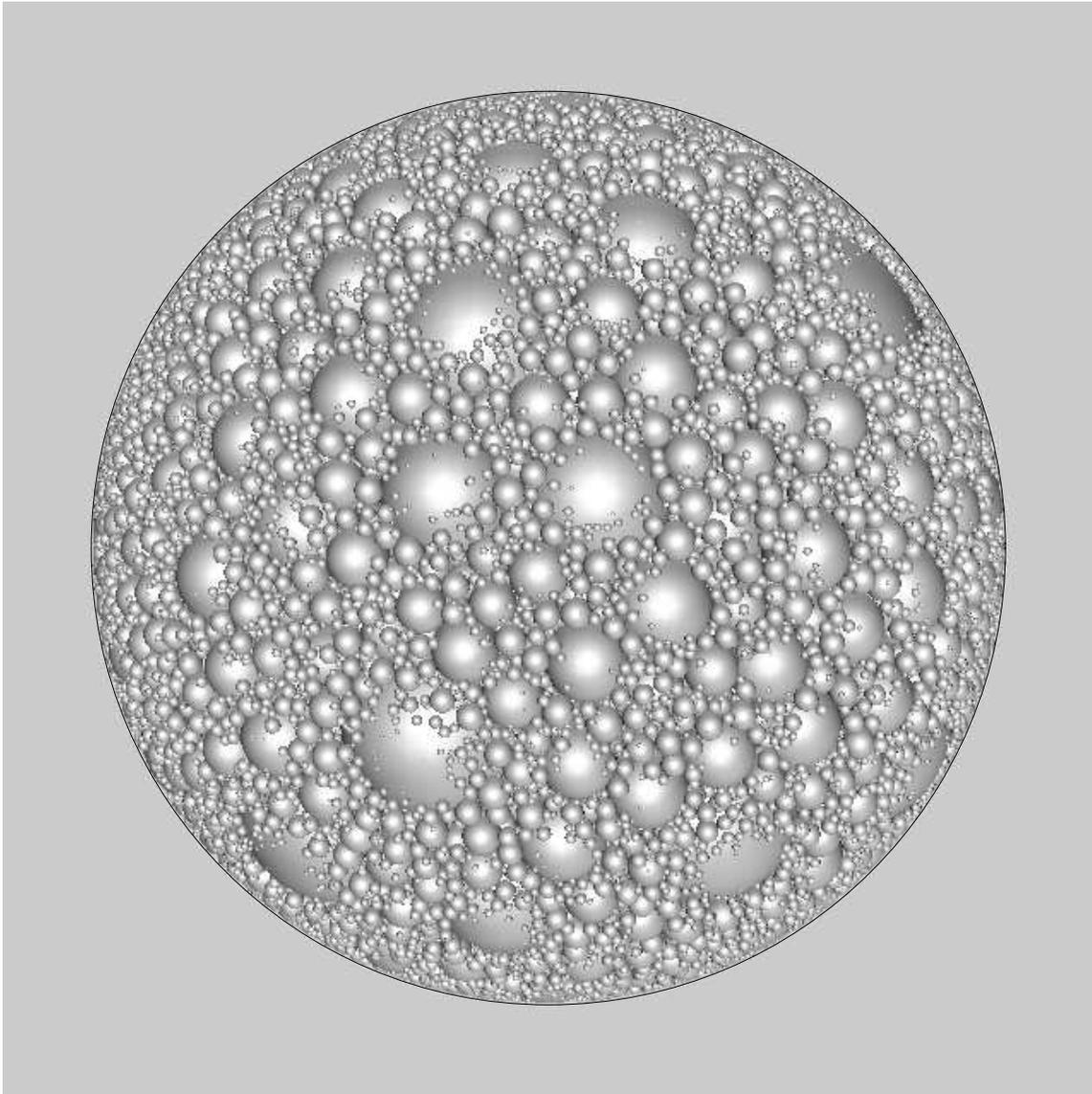


Figure 2.9: The $x - y$ plane projection of the full packing that was shown in Figure 2.8. The packing contains 45 539 spheres, though of course many are hidden in this projection.

from the packing depicted in the middle row of Figure 2.6, as they would be viewed by an observer looking onto the $x - y$ plane. The upper image shows the first 23 spheres that were placed randomly to “seed” the space for the packing algorithm, while the lower one also contains the first 76 packed to show the first 100 spheres placed in the space, the boundary of which is depicted by the thin black circle. Comparing this to the relevant $x - y$ plane plot in Figure 2.6, some of the same structure is apparent, particularly the location of the largest sphere. All 45 539 spheres in the packing were used to generate Figure 2.9, though most of them are hidden behind those at the surface closest to the observer. One can still make out some of the spheres that were shown in the image of the first 100 spheres, though the overall appearance of the packing is quite different with a much wider range of sphere sizes included.

The Two-Point Correlation Function

A more quantitative way of looking at the structure inside the packings involves computing the three-dimensional two-point spatial correlation function, ξ , for sphere centres. This function, parameterized using the form

$$\xi(r) = (r_0/r)^\gamma, \quad (2.1)$$

was computed for the spheres in a number of packings, using various estimators developed by Davis and Peebles (1983), Landy and Szalay (1993), and Hamilton (1993). These correlation functions were all found to have power law behaviour as is the case with observed matter distributions, though there were some indications that the functions might best be fit using more than one power law. The Landy and Szalay and Hamilton estimators produced very similar results, though the latter consistently gave larger values of r_0 and γ (i.e. less negative values of γ) than the Landy and Szalay estimator did. The $\frac{DD}{DR}$ estimator was very close to the Landy and Szalay estimator over shorter distances where the power was stronger, but diverged at larger separations. Refer to Figure 2.10 for a typical example. Though the correlation was found to be relatively weak for a

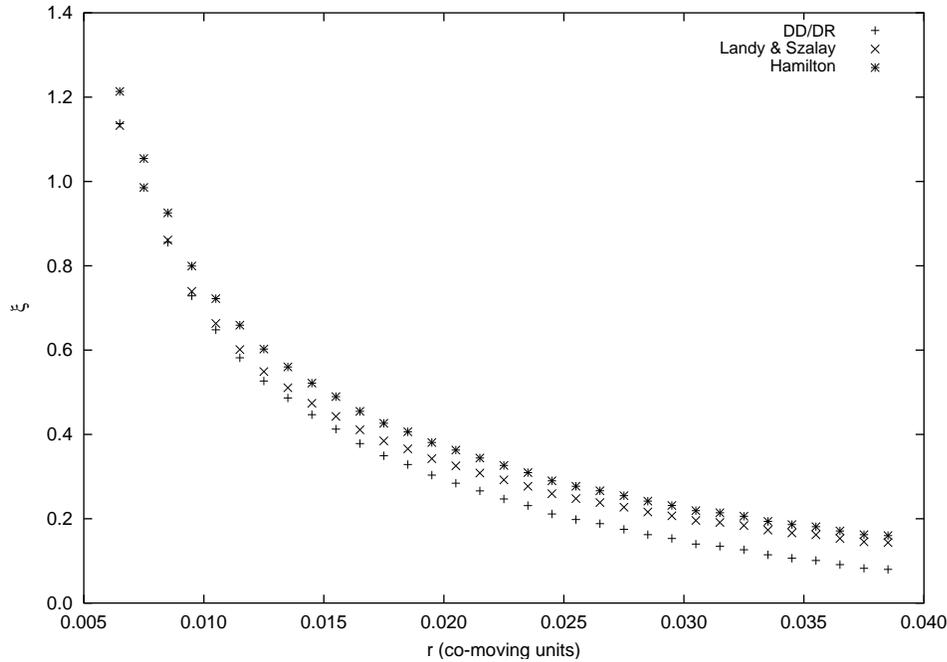


Figure 2.10: A comparison of a portion of the computed correlation function for a packing in positively curved space with 105 965 spheres.

single level packing, it should become stronger for the recursive model which will add a lot more power on smaller scales. Regardless of the estimator used, as the number of spheres in a packing increased, the slope of the correlation function became steeper and the correlation length decreased as expected.

Multifractal Structure

A related but more comprehensive means of quantifying the structure in a packing is via a multifractal analysis. A classical fractal (or monofractal) is an object with a self-similar or scale-invariant structure. In large part, the sphere packing problem boils down to trying to fit the largest spheres possible into spaces between other, still larger spheres. Consequently, it is fairly straightforward to see how the packing algorithm used can lead to self-similar structures appearing in the packings, with increasingly smaller spheres being used to surround the previous “generations” of spheres as the space being packed is filled. One would expect the fractal nature to be truncated – there is after

all a maximum sphere radius of ω_{\max} that can be used, and a lower limit to the size of the spheres needed to achieve a PVF that is less than 1. However, as was indicated by considering the correlation function, the packings are not simply truncated classical fractals, a point that can be understood intuitively since there are factors that cause the packing behaviour to vary as the space is filled. First of all, there is the random placement of spheres to “seed” the space, a procedure which does not involve space-filling optimization. Once the packing begins, the initial phase that involves fitting the larger spheres into the voids between the randomly placed ones operates in a space that looks quite different from that available when the smaller spheres are subsequently packed into the gaps left between the tightly packed larger ones. Further, those spheres packed next to the boundary of the space also experience conditions that differ from those further within. Consequently, rather than purely fractal behaviour, the packings exhibit structure that is *multifractal* in nature.

A monofractal has a single scaling dimension that is non-integer.³ A multifractal is a composite of monofractals, each of which has its own characteristic scaling dimension. There are a number of ways one can go about measuring the dimension of a fractal, with the most common being the *box-counting* method. In its simplest form, this involves placing a grid over the fractal, computing the fraction of cells occupied by the fractal, then refining the box size and repeating until a trend can be established. This method lends itself nicely to measuring the dimensionalities of multifractals.

Consider the factor $p_i(\epsilon)$, the relative proportion of a fractal inside box i of linear dimension ϵ ,

$$p_i(\epsilon) = \frac{n_i(\epsilon)}{n} ,$$

where $n_i(\epsilon)$ can, for instance, be the number of pixels from an image of the fractal that lie within box i , while n is the total number of pixels that define the fractal in the image.

³For instance, the area of a “flat” fractal does not increase in proportion to the square of its length, and the volume of one that extends into three dimensions does not increase as length cubed.

This can be used to define a partition function (Jensen et al., 1987),

$$\Gamma(q, \varrho) = \sum_i \frac{p_i^q}{\epsilon_i^q},$$

which, though strictly defined for balls of radius ϵ , is much easier in practice to compute using box counting as described previously. For a given value⁴ of q , consider the set of all partitions and find the supremum ($q > 0$) or infimum ($q < 0$) of Γ . In the limit $\epsilon_i \rightarrow 0$, there exists a value ϱ_q such that $\Gamma \rightarrow 0$ if $\varrho < \varrho_q$ and $\Gamma \rightarrow \infty$ if $\varrho > \varrho_q$, while converging to a constant at $\varrho = \varrho_q$. The *generalized dimension* of the multifractal, D_q , is then given as

$$D_q = \frac{\varrho_q}{q - 1} \tag{2.2}$$

(Jensen et al., 1987). Note that for integer quantities of $q \geq 0$, the dimension has physical meaning, with D_0 being the often used *fractal dimension*, D_1 being the information dimension, and D_2 the correlation dimension which is related to the exponent γ in the two-point correlation function. The set of all D_q characterizes the multifractal completely, with the measures all being equal for a monofractal, while

$$D_q > D_{q'} \quad \text{for} \quad q' > q$$

otherwise (Hentschel and Procaccia, 1983).

Mureika (2001) used the above method to estimate the dimensionality of the structure in a number of packings with $\omega_{\max} = \pi/10$ based on the locations of the packed sphere centres. He considered the structure present inside the largest box that could be completely contained within the spherical region of space that was packed, which makes the computation significantly easier, though eliminates much of the distinctive structure apparent at the boundary of the region. His results are summarized in Table 2.3. No significant differences were found between spaces of different curvature which is not too surprising considering that only packings with a small angular extent were considered.

⁴Not necessarily an integer value.

q	D_q Range
0	2.5 – 2.8
2	2.4 – 2.6
∞	2.1 – 2.3

Table 2.3: A summary of the results found by Mureika (2001) when considering the structure of a series of packings generated in a space defined by $\omega_{\max} = \pi/10$.

Note that D_∞ was found by computing D_{60} , a point where the dimension measure seemed to have become virtually constant. The limit $q \rightarrow \infty$ provides information about the scaling behaviour for the most dense regions of the multifractal. The range given in the table indicates that the densest structures in the packings considered tended towards two-dimensional formations, which is understandable given the tendency of the packing algorithm to form sheet-like collections of smaller spheres on the surface of larger ones. While these results for single packings aren't out of line with the wide ranges of values that have been found using N-body simulations, they do not agree with what has been found using Las Campanas Redshift Survey data where $D_\infty \approx 0.6 - 1.0$, though they can vary a fair bit if projection effects and a crude biasing factor based simply on sphere mass are introduced (Mureika, 2001, and references therein).

2.2 Generating an RSC Model

As indicated earlier, there are a variety of choices that can be made regarding the structure of the core at the centre of each SC hole. An obvious option is to choose the mass to consist of the same pressureless dust as in the FRW background but at a higher density, thus adding to the self-similar nature of the model. In fact, the self-similarity can be further extended by embedding spherical vacuoles containing their own cores into this higher density FRW phase, adding substructure to the lens in the process.

One way of accomplishing this would be to use the sphere packing code to fill each core

with spheres as it is encountered by some beam propagating through the model. However, as the sphere packing can take a significant amount of time to do, this would not be very practical. Another option, and the one actually implemented, makes recursive use of the library of SC packings available, taking advantage of the self-similar nature of the packings. The first time a core is hit by a beam, the code decides whether it is to be given substructure (refer to Section 3.3.3 for details). If so, then the chosen angular extent of the core, ω_c , is used to determine which of the available packings are suitable for use in describing this structure. In general, $\omega_c \neq \omega_{\max}$ for any of the packings in the library. Given that they are stored using only angular values, it may seem that scaling a packing up or down to the correct radius would be the obvious thing to do. However, for spaces with non-zero curvature, a packing cannot be scaled correctly by simply multiplying the radial ω coordinate and radius of each sphere by a constant factor as this does not preserve the distribution of tightly packed and non-overlapping spheres required.

Instead, the code uses a subsection of a randomly chosen packing that was made with the smallest value of ω_{\max} larger than ω_c . It picks a centre for the sub-packing by randomly selecting a sphere from within the packing, ensuring that it has a centre at least ω_c from the boundary. A check is also made to ensure that this sphere has a radius no larger than half ω_c in order to avoid selecting one that is uncharacteristically large for the region (i.e. that would not likely have been put in there by the packing code). If no such sphere can be found, the one furthest from the boundary is chosen, or the code can simply use the centre of the coordinate system that the packing is stored in. Next, all spheres that lie entirely within ω_c of the chosen sphere and which, when given the appropriate physical scale, will remain larger than the preset minimum physical radius to be considered, are selected. They are expressed in a coordinate system which has the chosen sphere at the centre (refer to Section 3.1.1), and the system is given a random orientation by spinning it by a random amount in a couple of planes. In the cases of non-zero curvature, these planes do not include the w axis so as to preserve the radial distance

of each sphere from the centre. The resultant sub-packing still has structure consistent with what would be generated by the packing code, and is thus used to describe any structure in the core.

2.2.1 Truncating the Model

Although in principle the recursive packing described above can continue ad infinitum while still satisfying the field equations, for any physical model there will be a lower limit below which it will no longer be reasonable to add structure. If the initial model density represents the mean density of some region of the universe, then clearly the packing cannot continue down to the point where the core of a sphere has a density equal to that of the Earth. The universe cannot be approximated by the RSC model at such a small scale so the recursion must be cut off at some appropriate level, such as at galaxy-mass objects. Furthermore, there is only a range of length scales that the model needs to account for. Suppose a typical galaxy has a radius of about 10 Kpc, about six orders of magnitude smaller than that of the observable universe. Since a single packing can have six orders of magnitude separating its smallest sphere from the size of its bounding surface, even at the top level of the hierarchy there are going to be spheres which do not need to be recursively filled. As one descends levels, the fraction of spheres that need to be re-packed will drop off very quickly since the majority of spheres have a radius that is less than a tenth the size of the sphere that they have been packed into. As a result, after about four levels, all branches of the hierarchical *tree* should have ended in *leaves*.

The RSC model is thus truncated based on mass and size limits. If during beam propagation, the mass of a potential lens is found to be below a predetermined mass limit, the sphere is not given further structure but treated as a leaf and ignored (i.e. the region occupied by the sphere is considered to be filled with dust at the same density as the background the sphere is embedded in). Similarly, as indicated earlier, spheres in a packing falling below a set length scale are not selected as candidates for substructure.

There are also two additional situations that lead to the model being truncated. If, when first hit by a beam, a sphere is found to have a radius that lies within the event horizon of the mass enclosed, then it is designated as a leaf. Further, if the code finds that the parameters describing the core need to be too finely tuned to be valid (refer to Section 3.3.3 for details), then the sphere is also designated as a leaf. It should be noted that the packing libraries typically have a small number of rather large spheres. If these spheres were to be denoted as leaves, then it is possible that a significant volume of the model universe will lack structure. If this proves to be a problem for some set of parameters, structure can always be added in a consistent manner by replacing these spheres with sub-packings as described above.

2.2.2 Storing the Model

Once an RSC model has been set up, light can be propagated through it from a variety of directions and observed at any vantage point that one may wish to place an observer. As such, it is important that the model remain self consistent. For instance, a beam encountering a particular lens from one direction at a particular time, should find the same lens (evolved by the appropriate amount) that a beam propagating through the region from another direction at an earlier time would have detected. Due to the recursive nature of the model, trying to initialize it by defining all possible lenses at the beginning can easily lead to a large data structure that is unwieldy to work with on a typical workstation. Furthermore, as most of the spheres are unlikely to be encountered by a beam, they would add to the computational effort required to propagate a beam without having any impact on the results produced. To address these issues, the RSC model is developed on the fly as it is “observed”. As a beam encounters embedded spheres, the lenses they contain are defined and enough information about them is stored so as to permit the evolution of the lens to be tracked properly should it be encountered by other beams. The result is an extensible model that only contains the data relevant for

the investigations it is being used for. Refer to the end of Section 3.3.4 for a list of the quantities actually stored to define the model.

Chapter 3

Propagating a Beam Through RSC

As indicated previously, a beam must follow a null geodesic of the spacetime through which it travels. The evolution of the 4-vector $x^a(\lambda)$ describing each point along this path is governed by the geodesic equations

$$\frac{d^2 x^a}{d\lambda^2} + \Gamma_{bc}^a \frac{dx^b}{d\lambda} \frac{dx^c}{d\lambda} = 0, \quad (3.1)$$

where Γ_{bc}^a is a Christoffel symbol given by

$$\Gamma_{bc}^a = \frac{1}{2} g^{am} (g_{bm,c} + g_{cm,b} - g_{bc,m}), \quad (3.2)$$

and λ is an affine parameter along the path. An equivalent form that is more practical to work with is that of the Euler-Lagrange equations¹

$$\frac{d}{d\lambda} \left(\frac{\partial L}{\partial \dot{x}^i} \right) - \frac{\partial L}{\partial x^i} = 0, \quad (3.3)$$

where the Lagrangian, L , is given by

$$L = \left(\frac{ds}{d\lambda} \right)^2, \quad (3.4)$$

and the derivative $\cdot \equiv \frac{d}{d\lambda}$.

¹Also known as the Lagrange equations of the second kind.

3.1 The FRW Regions

To propagate a beam through an FRW region of the model, it is useful to know whether it will encounter a packed sphere (lens), and if so, which one it will be. This will allow for the determination of a coordinate system in which the equations of motion are simplified, and will also avoid the code having to check for a sphere being hit after every step taken by the integration routine, a process that would be prohibitively time consuming considering the large number of spheres that can be present at a single level of the RSC model. The desired coordinate system has the following properties:

1. the origin of the coordinate system is placed at the centre of the lens under consideration
2. the beam starts at a location with $\theta = \frac{\pi}{2}$
3. the direction (tangent) vector for the beam lies in the $\theta = \frac{\pi}{2}$ plane – i.e. has a θ component of zero

A coordinate system oriented in this manner will be referred to as being in *propagation orientation*. Setting $\theta = \frac{\pi}{2}$ and $d\theta = 0$ in the RW line element (1.6),

$$ds^2 = dcT^2 - R^2(cT)[d\omega^2 + \mathcal{S}_K^2(\omega) d\phi^2].$$

Using this to define the Lagrangian via (3.4), the Euler-Lagrange equations (3.3) can be solved to find the form of the tangent vector k^a in this coordinate system:

$$\frac{dcT}{d\lambda} = \frac{p}{R} \tag{3.5a}$$

$$\frac{d\omega}{d\lambda} = \frac{\epsilon_\omega p}{R^2} \sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}} \tag{3.5b}$$

$$\frac{d\theta}{d\lambda} = 0 \tag{3.5c}$$

$$\frac{d\phi}{d\lambda} = \frac{pq}{R^2 \mathcal{S}_K^2(\omega)} \tag{3.5d}$$

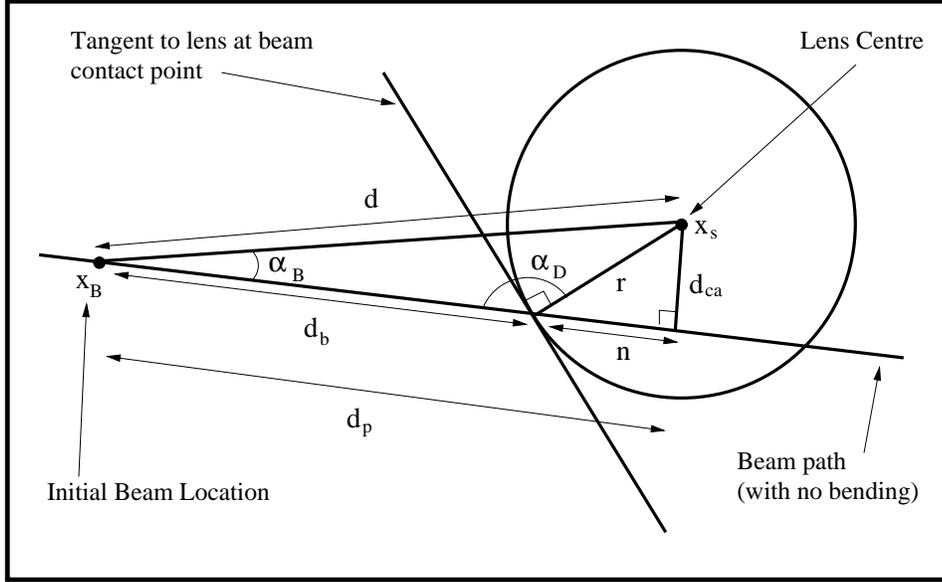


Figure 3.1: The various quantities used in the derivations to set up the propagation of a beam in FRW regions. Note that the beam path within the lens is shown as if there was no bending taking place.

where the nullity of the geodesic (i.e. $ds^2 = 0$) was used to find the ω component. The quantities p and q are constants of integration related to the energy and angular momentum of the beam respectively, and $\epsilon_\omega = \pm 1$ indicates whether the beam is headed towards or away from the lens. The constant p has units of length, while q is unit-less.

3.1.1 Orienting the System

As the conditions necessary to satisfy the propagation orientation described above only require modification of the spatial components of the beam's location and direction, one way to perform the necessary setup is via a series of $SO(2)$ rotations described in a Euclidian embedding space. The embedding of the potentially curved manifold into this space is described in Section 2.1.2 and Appendix C.1.1. When working in the embedding space, care must be taken that all transformations leave the points on the original manifold.

Consider, for instance, the case of positive curvature where the spatial sections take

the form of a 3-sphere. This can be chosen to be a unit 3-sphere for simplicity as the rotations to be performed will not involve the scale, so the embedding is described by equations (C.3) with $r = 1$. The centre of the lens under consideration is given in this system as

$$x_s = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix}, \quad (3.6)$$

which must be mapped to the origin of the angular system, $(\omega, \theta, \phi) = (0, 0, 0)$, or equivalently, to the pole at $(x, y, z, w) = (0, 0, 0, 1)$. To do this, start with a rotation in the xy plane to set the x component to zero,

$$m_1 \times x_s = \begin{pmatrix} \cos \alpha & \sin \alpha & 0 & 0 \\ -\sin \alpha & \cos \alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ y' \\ z \\ w \end{pmatrix}. \quad (3.7)$$

Now repeat the procedure, rotating in the yz and zw planes with matrices m_2 and m_3 to set the y and z components to zero respectively. Since x_s lies on the surface of the hypersphere, $w' = \pm 1$. As the angular coordinates only describe the positive w portion of the 4-space, if $w' = -1$, a further rotation in the zw plane of 180° must be performed to make w' positive. Calling this rotation (which is just the identity matrix if $w' = 1$ already) m_{flip} , the transformation thus far is given by

$$m_{\text{flip}} \times m_3 \times m_2 \times m_1 \times x_s = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}. \quad (3.8)$$

This takes care of the sphere centre. Next, rotations are needed to ensure that the beam at x_B has a θ value of $\frac{\pi}{2}$, which is equivalent to having a z coordinate of zero. Performing the transformation $m_{\text{flip}} \times m_3 \times m_2 \times m_1 \times x_B$ expresses x_B in the coordinate system centred on the lens, and gives it a w coordinate value equal to $\cos d$ (i.e. the angular separation between the sphere centre and x_B is d). A further rotation in the yz plane, m_5 , allows the z component to be set to zero as required without disturbing the w component, thus placing the beam at $(x_{\text{temp}}, y_{\text{temp}}, 0, \cos d)$, while leaving the sphere centre at the origin.

Finally, the coordinate system needs to be further manipulated so that the tangent vector describing the beam propagation direction from x_B lies in the $\theta = \frac{\pi}{2}$ plane, that is to say, it has a θ component of zero. To do this, start by applying a rotation in the xy plane, m_6 , which sets the y component of x_B to zero, and gives the beam a starting location of the form:

$$x_B = \begin{pmatrix} x_{Bf} \\ 0 \\ 0 \\ \cos d \end{pmatrix} = \begin{pmatrix} \sin d \\ 0 \\ 0 \\ \cos d \end{pmatrix}, \quad (3.9)$$

where the form of x_{Bf} is known since x_B must remain on the unit 3-sphere. As before, this rotation leaves the lens at the pole. Next, the components of the tangent vector need to be expressed in the Cartesian embedding coordinates. This transformation is given in equations (C.7), where r can again be set to 1 and $v^4 = 0$ since any vector under consideration will lie on the 3-sphere in question.

Formally, each rotation that has been performed on the system describes a linear coordinate transformation. To express a vector in the new coordinate system, the Jacobian of the transformation needs to be computed and applied to the vector appropriately. It is straightforward to show that for each of the rotations, the Jacobian is in fact the same as the rotation matrix, so in practice, the rotations can be applied to the components of a direction in the same way that they are used with those of a location. Applying the transformation used to re-orient the lens and x_B to the tangent vector of the beam,

$$m_6 \times m_5 \times m_{\text{flip}} \times m_3 \times m_2 \times m_1 \times \begin{pmatrix} \tilde{k}^1 \\ \tilde{k}^2 \\ \tilde{k}^3 \\ \tilde{k}^4 \end{pmatrix} = \begin{pmatrix} (\tilde{k}^1)' \\ (\tilde{k}^2)' \\ (\tilde{k}^3)' \\ (\tilde{k}^4)' \end{pmatrix}, \quad (3.10)$$

and the vector is expressed in the rotated system. The requirement that $d\theta = 0$ at the point with $\theta = \frac{\pi}{2}$ is equivalent to requiring that $\tilde{k}^3 = 0$. This can be accomplished by performing a final rotation in the yz plane, m_7 , so

$$m_7 \times \begin{pmatrix} (\tilde{k}^1)' \\ (\tilde{k}^2)' \\ (\tilde{k}^3)' \\ (\tilde{k}^4)' \end{pmatrix} = \begin{pmatrix} (\tilde{k}^1)' \\ (\tilde{k}^2)'' \\ 0 \\ (\tilde{k}^4)' \end{pmatrix}. \quad (3.11)$$

Once more, the rotation leaves the lens unmoved, and with $y = z = 0$, point x_B is also unaffected. The result is now a system where the lens is at the origin, the beam starts at a location that has $\theta = \frac{\pi}{2}$, and the tangent vector has $k^2 = 0$ as required. The values of the six rotation angles in m_1 through m_6 can be expressed completely in terms of the coordinates x_s and x_B , while the angle used in m_7 can be found since the values of the components of the tangent vector at x_B are known.

Using equations (C.3) to convert the position back to angular coordinates, x^a now has the form $(\omega_f, \theta_f, \phi_f) = (d, \frac{\pi}{2}, 0)$. Similarly, equations (C.8) can be used to transform the components of the tangent vector back to angular form, so

$$k^1 = -\frac{(\tilde{k}^4)'}{\sin d} \quad (3.12a)$$

$$k^2 = 0 \quad (3.12b)$$

$$k^3 = \frac{(\tilde{k}^2)''}{\sin d}. \quad (3.12c)$$

With the tangent vector now satisfying the conditions required to have the form given by (3.5), the value of p can be set from the known $\frac{dcT}{d\lambda}$, and q can then be computed by using the value of k^3 found above. These quantities must satisfy the expression for $\frac{d\omega}{d\lambda}$ since it was derived using the expressions for $\frac{dcT}{d\lambda}$, $\frac{d\phi}{d\lambda}$, and the nullity condition. In fact, note that from (3.5),

$$k_a = g_{ab}k^b = \left(\frac{p}{R}, -\epsilon_\omega p \sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}}, 0, -pq \right) \quad (3.13)$$

when $\theta = \frac{\pi}{2}$, and the inner product of k^a is then

$$k_a k^a = \frac{p^2}{R^2} - \frac{p^2}{R^2} \left(1 - \frac{q^2}{\mathcal{S}_K^2(\omega)} \right) - \frac{p^2 q^2}{R^2 \mathcal{S}_K^2(\omega)} = 0 \quad (3.14)$$

as expected for a null vector.

The procedure is analogous in the case of a negatively curved space, except that equations (C.4) are used for the embedding, while (C.9) and (C.10) are used to convert the direction vector. Further, while the rotations in the xy and yz planes in the embedding

space do not change, the rotation matrix for the zw plane does. This is because under the rotation, points must remain on the unit 3-hyperboloid, defined by $w^2 - x^2 - y^2 - z^2 = 1$, and not on the unit 3-sphere where $w^2 + x^2 + y^2 + z^2 = 1$. The new matrix describing a rotation through an angle α in the zw plane that satisfies this condition has the form

$$m_3 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cosh \alpha & \sinh \alpha \\ 0 & 0 & \sinh \alpha & \cosh \alpha \end{pmatrix}. \quad (3.15)$$

This matrix has a unit determinant and preserves inner products as required. Though a rotation in the xw or yw planes would also require a suitably re-arranged matrix of this form, these rotations aren't necessary for the re-orientation described. The form of m_{flip} found before still satisfies the rotation properties necessary and can thus be used in this case too, though it can only be derived from (3.15) using imaginary values of α .

For the case of a space with zero curvature, the calculations simplify considerably as there is no need to introduce a fourth embedding dimension. In fact, a set of simple translations are all that are required to have the lens at the origin. There is also no potential flip to worry about, and the subsequent rotations are performed using only 3×3 matrices. Again, refer to Appendix C.1.1 for conversion details.

3.1.2 Finding the Target

Find the First Lens Encountered

By assuming that no deviation from the FRW trajectory of the beam takes place (i.e. that none of the packed spheres have the mass they contain replaced by a vacuum region and dense core), the code can determine which spheres, if any, the beam will potentially encounter by simply comparing the radius (r) of each sphere to d_{ca} , the distance of closest approach to its centre that the beam would have. With the coordinate system re-oriented as described above, the unlensed point of closest approach of the beam to a

sphere placed at the origin occurs at $\dot{\omega} = 0$. Thus, from (3.5b),

$$\mathcal{S}_K(\omega) = \pm q$$

at closest approach. Using the definition of $\mathcal{S}_K(\omega)$ in (1.8) and solving for ω ,

$$d_{ca} = \omega = \begin{cases} \sinh^{-1}(\pm q) & : K = -1 \\ \pm q & : K = 0 \\ \sin^{-1}(\pm q) & : K = +1 \end{cases} . \quad (3.16)$$

In each case, the sign is chosen so that $d_{ca} \geq 0$.

While the geodesic distance between the beam and boundary of a sphere (d_b in Figure 3.1) can be computed to determine which potential lens encounter occurs first, whenever a sphere is found to lie in the path of the beam, it turns out to be quicker and easier to compute d_p , the distance to the hypothetical point of closest approach instead. One can quickly convince oneself that the sphere having the smallest such distance from the beam location x_B will also be the one that the beam path must intersect first.

To determine d_p , it is first necessary to compute the geodesic distance d from x_B to the sphere centre. Finding an expression for this distance in flat space is straightforward. Given two points in Cartesian coordinates, the distance between them is given by

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2} ,$$

or, converting to spherical coordinates via (C.2) and re-arranging the terms,

$$d = \sqrt{\omega_1^2 + \omega_2^2 - 2\omega_1\omega_2[\sin\theta_1 \sin\theta_2 \cos(\phi_1 - \phi_2) + \cos\theta_1 \cos\theta_2]} , \quad (3.17)$$

where in this case, ω is a linear radial coordinate. For the equivalent expression in positively curved space, describe two points P_1 and P_2 on the unit 3-sphere, using 4-vectors in the Euclidian embedding space, i.e. $P_1 = (x_1, y_1, z_1, w_1)$ and $P_2 = (x_2, y_2, z_2, w_2)$. In the plane defined by the origin of the embedding space and these two points, the geodesic connecting them along the 3-sphere will take the form of a circular segment (this must

be the case by definition of a 3-sphere). If this segment has length d and subtends an angle α at the origin, then

$$d = r\alpha = \alpha \quad (3.18)$$

since $r = 1$. By definition,

$$P_1 \cdot P_2 = |P_1||P_2| \cos \alpha . \quad (3.19)$$

Using $|P_1| = |P_2| = 1$, and (3.18) in (3.19),

$$\alpha = \cos^{-1}(P_1 \cdot P_2) ,$$

or

$$d = \cos^{-1}(x_1x_2 + y_1y_2 + z_1z_2 + w_1w_2) .$$

Transforming to spherical coordinates via (C.3),

$$d = \cos^{-1}(\cos \omega_1 \cos \omega_2 + \sin \omega_1 \sin \omega_2 [\sin \theta_1 \sin \theta_2 \cos(\phi_1 - \phi_2) + \cos \theta_1 \cos \theta_2]) , \quad (3.20)$$

where this time ω is an angular radial coordinate as always in a space of non-zero curvature. The procedure can be repeated for the case of negative curvature (or one can simply allow $\alpha \rightarrow -i\alpha$, and $\omega \rightarrow -i\omega$), to find

$$d = \cosh^{-1}(\cosh \omega_1 \cosh \omega_2 - \sinh \omega_1 \sinh \omega_2 [\sin \theta_1 \sin \theta_2 \cos(\phi_1 - \phi_2) + \cos \theta_1 \cos \theta_2]) . \quad (3.21)$$

Note that the geodesic distances (3.17), (3.20), and (3.21) are all unit-less in this form and are only given a physical scale when multiplied by the scale factor R .

As is apparent from Figure 3.1, knowing d allows the distance d_p to be computed using trigonometry. Consider the triangle with sides d , d_{ca} , and d_p ($\triangle d_{ca}d_p d$). In the case of positive curvature, the Law of Cosines from spherical trigonometry relates these values via:

$$\cos(d) = \cos(d_{ca}) \cos(d_p) + \sin(\omega_m) \sin(d_p) \cos\left(\frac{\pi}{2}\right) ,$$

so

$$d_p = \cos^{-1}\left(\frac{\cos d}{\cos d_{ca}}\right) . \quad (3.22)$$

In the flat case, it is easy to see from the right angled triangle that

$$d_p = \sqrt{d^2 - d_{ca}^2}, \quad (3.23)$$

while in the case of negative curvature, the hyperbolic Law of Cosines shows

$$\cosh(d) = \cosh(d_{ca}) \cosh(d_p) - \sinh(d_{ca}) \sinh(d_p) \cos\left(\frac{\pi}{2}\right),$$

so

$$d_p = \cosh^{-1}\left(\frac{\cosh d}{\cosh d_{ca}}\right) \quad (3.24)$$

Note that each of these expressions for d_p is monotonic, so to save computing time, when trying to find the smallest value of d_p in the curved spaces, only the fractions need to be compared, while in the flat case, the comparison can be made using values of d_p^2 directly.

Computing the Distance to the Lens and the Angle of Intersection

For the first sphere that is intersected, it is possible to compute the distance d_b from x_B to the point of intersection, and the impact angle α_D . Again, from Figure 3.1, it is clear that trigonometry is needed.

Consider the triangle $\triangle d_{ca}d_p d$. In the case of positive curvature, spherical trigonometry gives:

$$\cos(d_{ca}) = \cos(d_p) \cos(d) + \sin(d_p) \sin(d) \cos(\alpha_B),$$

which can be rearranged to yield

$$\alpha_B = \cos^{-1}\left(\frac{\cos d_{ca} - \cos d_p \cos d}{\sin d_p \sin d}\right). \quad (3.25)$$

Considering now $\triangle r d_b d$, the spherical Law of Sines provides the relation:

$$\frac{\sin \alpha_D}{\sin d} = \frac{\sin \alpha_B}{\sin r},$$

which, solving for α_D gives:

$$\alpha_D = \sin^{-1}\left(\frac{\sin(d) \sin(\alpha_B)}{\sin(r)}\right). \quad (3.26)$$

This is the angle between the beam and the radial vector of the sphere. To find the distance to the point of intersection, the Law of Cosines can be used on $\triangle d_{ca}rn$ to give

$$\cos(r) = \cos(n) \cos(d_{ca}) + \sin(n) \sin(d_{ca}) \cos\left(\frac{\pi}{2}\right) ,$$

which can be rearranged to give an expression for n ,

$$n = \cos^{-1}\left(\frac{\cos r}{\cos d_{ca}}\right) .$$

Since $d_b = d_p - n$, the distance from x_B to the point of intersection is thus

$$d_b = d_p - \cos^{-1}\left(\frac{\cos r}{\cos d_{ca}}\right) . \quad (3.27)$$

The standard Euclidian space trigonometric relations can be used in the flat case which results in significant simplification. Considering $\triangle rd_{ca}n$,

$$\sin(\pi - \alpha_D) = \frac{d_{ca}}{r} ,$$

so the angle α_D is simply

$$\alpha_D = \sin^{-1}\left(\frac{d_{ca}}{r}\right) . \quad (3.28)$$

To find the distance to the point of intersection, note that n is

$$n = r \cos(\pi - \alpha_D) ,$$

which results in the expression for d_b ,

$$d_b = d_p - r \cos \alpha_D . \quad (3.29)$$

The case of negative curvature is very similar to the case of positive curvature, though again, hyperbolic trigonometric relations must be used in place of spherical trigonometric ones. Applying the hyperbolic Law of Cosines to $\triangle d_{ca}d_p d$,

$$\cosh(d_{ca}) = \cosh(d_p) \cosh(d) - \sinh(d_p) \sinh(d) \cos(\alpha_B) ,$$

so the angle α_B is given by

$$\alpha_B = \cos^{-1} \left(\frac{\cosh(d_p) \cosh(d) - \cosh(d_{ca})}{\sinh(d_p) \sinh(d)} \right). \quad (3.30)$$

The hyperbolic Law of Sines can then be applied to $\Delta r d_b d$ to find

$$\frac{\sin \alpha_D}{\sinh d} = \frac{\sin \alpha_B}{\sinh r},$$

so α_D , the angle of interest, is given by

$$\alpha_D = \sin^{-1} \left(\frac{\sinh(d) \sin(\alpha_B)}{\sinh(r)} \right). \quad (3.31)$$

Considering now $\Delta d_{ca} r n$,

$$\cosh(r) = \cosh(n) \cosh(d_{ca}) - \sinh(n) \sinh(d_{ca}) \cos\left(\frac{\pi}{2}\right),$$

so

$$n = \cosh^{-1} \left(\frac{\cosh r}{\cosh d_{ca}} \right)$$

and thus d_b , the distance from x_B to the point of intersection with the sphere, is given by

$$d_b = d_p - \cosh^{-1} \left(\frac{\cosh r}{\cosh d_{ca}} \right). \quad (3.32)$$

Note that in each of the above cases, we require $\alpha_D \geq \frac{\pi}{2}$ so that $\alpha_D - \frac{\pi}{2}$ gives the incident angle with which the beam hits the sphere (i.e. the acute angle between the beam and the tangent to the sphere at the point of contact).

Finding the Coordinates of Intersection

Using the form of k^a given in (3.5), it is possible to analytically determine the location that the beam ends up having on the surface of the embedded sphere that it encounters. In this coordinate system, the value of ω_f at the end of propagation is known and the θ coordinate remains $\frac{\pi}{2}$, leaving only ϕ_f to be computed. Using (3.5b) to replace $d\lambda$ in (3.5d),

$$\int_{\phi_i}^{\phi_f} d\phi = \epsilon_\omega q \int_{\omega_i}^{\omega_f} \frac{d\omega}{\mathcal{S}_K(\omega) \sqrt{\mathcal{S}_K^2(\omega) - q^2}},$$

which can be solved for each value of the curvature indicator to get

$$\phi_f = \begin{cases} \phi_i - \epsilon_\omega \left[\tan^{-1} \left(\frac{q \cosh(\omega)}{\sqrt{\sinh^2(\omega) - q^2}} \right) \right]_{\omega_i}^{\omega_f} & : K = -1 \\ \phi_i + \epsilon_\omega \left[\cos^{-1} \left(\frac{q}{\omega} \right) \right]_{\omega_i}^{\omega_f} & : K = 0 \\ \phi_i - \epsilon_\omega \left[\tan^{-1} \left(\frac{q \cos(\omega)}{\sqrt{\sin^2(\omega) - q^2}} \right) \right]_{\omega_i}^{\omega_f} & : K = +1 \end{cases} . \quad (3.33)$$

Aside from the location of the beam, the code must also track the evolution in the scale factor R during the propagation. Since the total amount of matter in the model is constant, $\rho R^3 = \rho_0 R_0^3$, where the subscripted values are evaluated at some instant in time such as at the beginning of the beam propagation. Then from the Friedmann Equation (1.12),

$$\frac{dR}{d(cT)} = \epsilon_R \sqrt{\frac{8\pi G \rho_0 R_0^3}{3c^2 R} + \frac{\Lambda R^2}{3} - K} , \quad (3.34)$$

where $\epsilon_R = \pm 1$. Combining this with (3.5a),

$$\frac{dR}{d\lambda} = \epsilon_R p \sqrt{\frac{8\pi G \rho_0 R_0^3}{3c^2 R^3} + \frac{\Lambda}{3} - \frac{K}{R^2}} , \quad (3.35)$$

which can be numerically integrated simultaneously with (3.5a) for the time cT , and (3.5b) to determine the value of λ at the point of contact with the lens. At that instant, the beam will thus have coordinate values $(cT_f, \omega_f, \frac{\pi}{2}, \phi_f)$, and the FRW region a scale factor of R_f , all of which can be determined.

The Case if no Sphere is Hit

If the beam is found not to encounter any spheres in the region it is to propagate through, then the centre of the region is designated as the target “lens” and the orientation of the coordinate system as described in Section 3.1 can still be carried out. If, once oriented, the beam is found to have $\dot{\omega} > 0$, it must be headed towards the boundary of the FRW region so the target value of ω_f is simply the comoving radius of the boundary and propagation can continue as usual. If, on the other hand, the beam is found to be headed

towards the centre of the region, the propagation is broken up into two steps to properly deal with the closest approach. Refer to Section 5.3.2 for details.

3.2 The Schwarzschild Regions

For any single encounter with a spherical object, the beam is confined to motion in a plane that contains the centre of this object. Hence, as was the case in the background before, when the beam travels through a Schwarzschild vacuum in the RSC model, it is possible to choose the coordinates to be oriented in such a way that they simplify the equations of motion. In fact, the same propagation orientation defined for the FRW regions in Section 3.1, with the lens at the centre of the coordinate system and the beam propagating in the plane with $\theta = \pi/2$, can also be used here.

With this orientation, the equations of motion of the beam can be determined in the same way that they were for the FRW regions. The Schwarzschild line element becomes

$$ds^2 = f(r) d(ct)^2 - \frac{dr^2}{f(r)} - r^2 d\phi^2 ,$$

which leads to a Lagrangian that can be used in the Euler-Lagrange equations (3.3). Not surprisingly, $\frac{\partial L}{\partial t} = 0$ and $\frac{\partial L}{\partial \phi} = 0$, which is reasonable since the beam is under the influence of a central force so energy and angular momentum should be conserved. Using r_γ to represent the radial coordinate of the beam in the vacuum, k^a , the null tangent vector to the geodesic, has components:

$$\frac{dct}{d\lambda} = \frac{l_\gamma}{f(r_\gamma)} \tag{3.36a}$$

$$\frac{dr_\gamma}{d\lambda} = \epsilon_\gamma \sqrt{l_\gamma^2 - \frac{h_\gamma^2}{r_\gamma^2} f(r_\gamma)} \tag{3.36b}$$

$$\frac{d\theta}{d\lambda} = 0 \tag{3.36c}$$

$$\frac{d\phi}{d\lambda} = \frac{h_\gamma}{r_\gamma^2} \tag{3.36d}$$

The quantities l_γ and h_γ are constants of integration and $\epsilon_\gamma = \pm 1$ indicates whether the beam is headed towards or away from the core at the centre of the vacuum. l_γ is related to the energy of the beam and is unit-less, while h_γ , with units of length, describes an impact parameter and is thus related to the angular momentum of the beam.

Besides the motion of the beam in the vacuum, it will also be useful to determine the 4-velocity, u^a , of a timelike observer in this region. In particular, we will be interested in the motion of observers moving with the inner and outer boundaries of the vacuum. Since these move in a radial fashion in the Schwarzschild coordinates, the observers will have

$$u^2 = u^3 = 0. \quad (3.37)$$

Using the Euler-Lagrange equations again, but with an affine parameter τ , a timelike observer moving with the outer boundary (the *boundary observer*) has

$$u^0 = \frac{dct}{d\tau} = \frac{l_b}{f(r_b)}, \quad (3.38)$$

analogous to (3.36a), but with r_b describing the radial coordinate of this observer and l_b being a constant describing the energy of their motion. Using (3.37) and (3.38) in the Schwarzschild metric (1.2),

$$u^1 = \frac{dr_b}{d\tau} = \epsilon_b \sqrt{l_b^2 - f(r_b)}, \quad (3.39)$$

where $\epsilon_b = \pm 1$ depending on whether the observer is moving away from the centre or towards it, and this time, the relation $\frac{ds^2}{d\tau^2} = 1$ for a timelike observer was used, instead of $\frac{ds^2}{d\tau^2} = 0$ as was the case for the beam. The same will hold true for a timelike observer at the inner border of the vacuum on the edge of the core mass (the *core observer* who will have quantities denoted by a subscript “c” rather than a “b”).

In an FRW region of the RSC model, the location and size of each embedded sphere are fixed using comoving coordinates. These fixed targets allow the ultimate destination of a beam being propagated thorough the region to be determined a priori. The situation

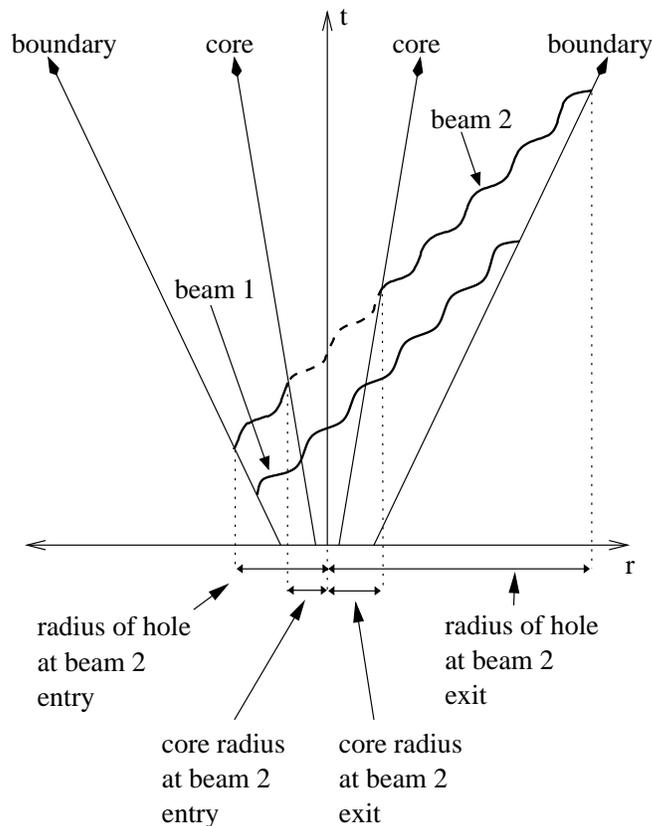


Figure 3.2: A spacetime diagram using vacuum (not comoving) coordinates to describe the propagation of a couple of beams through an expanding hole with an expanding core. Beam 1 enters the hole first but does not encounter the core, while beam 2 does travel through it (shown as a dashed wavy line). Note that the wavy lines are used simply as representations of the beams of light and do not imply varying propagation speeds. Furthermore, for simplicity, the beams are shown travelling in straight lines (i.e. with no lensing depicted).

is more complicated once a beam enters a vacuum region because the potential targets at the inner and outer boundary are both moving. The radius of the core and of the background region in which it is embedded are evolving at different rates, so in general, it is not immediately apparent whether a beam with a given direction vector will hit the core or simply propagate past it before returning to the outer boundary again.² Refer to Figure 3.2 for a spacetime diagram showing the two scenarios, with Figures 3.3 and 3.4 providing alternate graphical depictions.

²There is also the possibility that beams do not make it all the way through a hole, but this will be dealt with in Section 3.4.

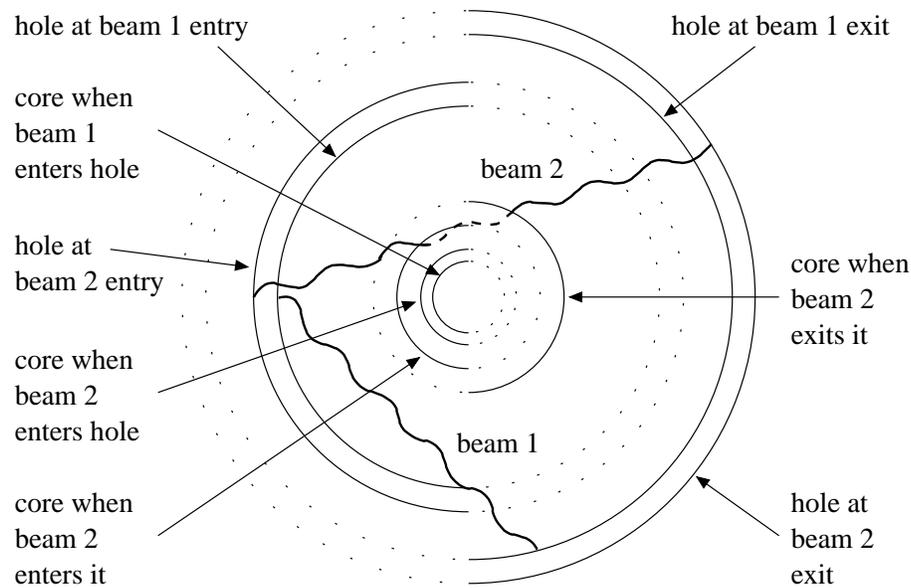


Figure 3.3: A top view of Figure 3.2. The beams enter from the left side and propagate to the right. Solid lines are used to depict the boundaries of the hole and core as relevant to the propagation of the beam on the given half of the diagram. As before, the beams are shown as heavy solid lines except where beam 2 propagates inside the core. There it is displayed using a heavy dashed line. Once more, the wavy lines do not imply variations in the path of a beam – they are simply used as a means of distinguishing the beams from the boundaries. There is no lensing depicted in the diagram.

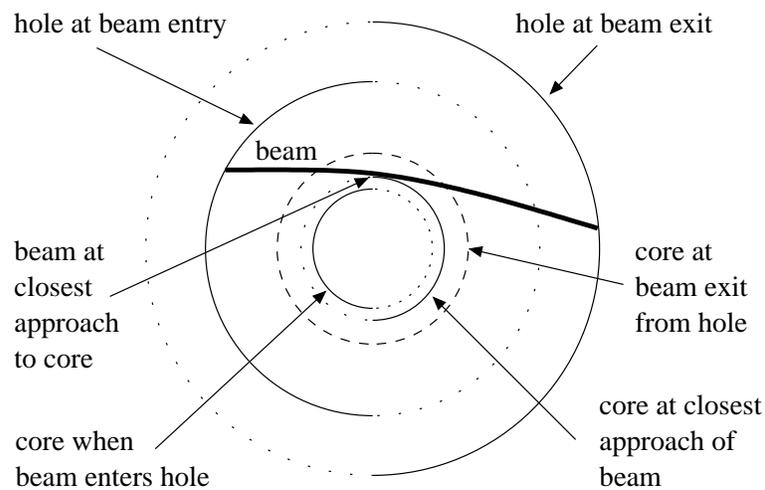


Figure 3.4: This figure is similar to Figure 3.3, but it depicts the critical case where a beam just grazes the core tangentially. Here, the beam is depicted as a heavy solid line, and an attempt is made to depict the bent path that the beam would take as it travels through the dynamic hole. The beam enters on the left and as it follows a curved path through the vacuum, the core expands. At its point of closest approach to the core, the beam just avoids contact and continues on to exit on the right side, travelling further through the vacuum as the hole has expanded.

3.2.1 Determining the Potential Closest Approach

The radial equation of motion of the beam in the vacuum (3.36b) has an overall sign, ϵ_γ , that needs to be changed if the beam reaches a minimum radius without encountering the core. As such, it is necessary for the code to determine this potential point of closest approach when the beam enters a vacuum. The point is reached when $\frac{dr_\gamma}{d\lambda} = 0$, which implies

$$r_\gamma^3 - \frac{r_\gamma}{\Gamma} + \frac{2m}{\Gamma} = 0 ,$$

where $\Gamma \equiv \frac{l_\gamma^2}{h_\gamma^2} + \frac{\Lambda}{3}$. A cubic of the form $x^3 + a_1x^2 + a_2x + a_3 = 0$ has roots

$$\begin{aligned} x_1 &= S + T - \frac{1}{3}a_1 \\ x_2 &= -\frac{1}{2}(S + T) - \frac{1}{3}a_1 + \frac{i\sqrt{3}}{2}(S - T) \\ x_3 &= -\frac{1}{2}(S + T) - \frac{1}{3}a_1 - \frac{i\sqrt{3}}{2}(S - T) \end{aligned}$$

where

$$\begin{aligned} S &= \sqrt[3]{V + \sqrt{U^3 + V^2}} & T &= \sqrt[3]{V - \sqrt{U^3 + V^2}} \\ U &= \frac{3a_2 - a_1^2}{9} & V &= \frac{9a_1a_2 - 27a_3 - 2a_1^3}{54} \end{aligned}$$

With $\Lambda > 0$

$$a_1 = 0 \quad a_2 = -\frac{1}{\Gamma} < 0 \quad a_3 = \frac{2m}{\Gamma} > 0$$

so $U = -\frac{1}{3\Gamma}$, $V = -\frac{m}{\Gamma}$, and the discriminant

$$D = U^3 + V^2 = \frac{27\Gamma m^2 - 1}{27\Gamma^3} .$$

If $D > 0$, there is one real root,

$$r = S + T = \sqrt[3]{V + \sqrt{D}} + \sqrt[3]{V - \sqrt{D}} . \quad (3.40)$$

If $D = 0$, all roots are real and the last 2 are equal, with the positive one being given by

$$r = \begin{cases} -\frac{1}{2}(S + T) = -\sqrt[3]{V} = \sqrt[3]{\frac{m}{\Gamma}} & : V < 0 \\ S + T = 2\sqrt[3]{V} = -2\sqrt[3]{\frac{m}{\Gamma}} & : V > 0 \end{cases} . \quad (3.41)$$

If $D < 0$, all roots are real and unequal, and the solutions can be written as:

$$x_1 = 2\sqrt{-U} \cos\left(\frac{\theta}{3}\right) - \frac{a_1}{3} \quad (3.42a)$$

$$x_2 = 2\sqrt{-U} \cos\left(\frac{\theta}{3} + \frac{2\pi}{3}\right) - \frac{a_1}{3} \quad (3.42b)$$

$$x_3 = 2\sqrt{-U} \cos\left(\frac{\theta}{3} + \frac{4\pi}{3}\right) - \frac{a_1}{3} \quad (3.42c)$$

where

$$\cos(\theta) \equiv \frac{V}{\sqrt{-U^3}} = -3m\sqrt{3\Gamma}$$

in this case. This means $\frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}$, or $\frac{\pi}{6} \leq \frac{\theta}{3} \leq \frac{3\pi}{6}$, so $\cos(\frac{\theta}{3})$, and hence x_1 , is positive. Similarly, $\frac{\theta}{3} + \frac{2\pi}{3}$ lies between $\frac{5\pi}{6}$ and $\frac{7\pi}{6}$, so x_2 is negative. Finally, $\frac{\theta}{3} + \frac{4\pi}{3}$ lies between $\frac{9\pi}{6}$ and $\frac{11\pi}{6}$, so x_3 is positive. This means that the correct root is one of x_1 or x_3 :

$$r_1 = \frac{2}{\sqrt{3\Gamma}} \cos\left(\frac{\theta}{3}\right) \quad (3.43)$$

$$r_3 = \frac{2}{\sqrt{3\Gamma}} \cos\left(\frac{\theta + 4\pi}{3}\right) \quad (3.44)$$

The valid one for the purpose at hand is the largest one less than the value of r_γ upon beam entry to the vacuum.

3.3 Connecting the Regions

3.3.1 FRW Background to Schwarzschild Vacuum

Matching the Location and Direction of the Beam

Consider the situation of a beam moving from an FRW region to a Schwarzschild vacuum in the model. As indicated in Appendix C.1.2, the θ and ϕ coordinates used in each region are identical, so the beam can remain at

$$\theta_{\text{Sch}} = \frac{\pi}{2} \quad \text{and} \quad \phi_{\text{Sch}} = \phi_f \quad (3.45)$$

in the Schwarzschild coordinate system, where ϕ_f is the final ϕ value computed in the FRW region via (3.33). Similarly, (C.11a) can be used to write

$$r_b = R(cT_f)\mathcal{S}_K(\omega_f) , \quad (3.46)$$

while the absolute value of the time coordinate in this system is not important so ct can be set to zero initially. Hence, at the boundary to the vacuum, a beam that has travelled through an FRW region in the propagation orientation (refer to Section 3.1) would have Schwarzschild coordinates of $(0, r_b, \frac{\pi}{2}, \phi_f)$. In fact, as shall be seen below, trying to follow the Schwarzschild coordinate time can lead to computational difficulties so as it is not actually needed, it does not need to be followed.

Aside from the location of the beam, its direction of propagation described by the tangent vector k^a , also needs to be expressed using Schwarzschild coordinates. As the FRW and Schwarzschild regions are matched at the boundaries between them, the null geodesic trajectory must be smoothly continuous across these boundaries. In the propagation orientation, the components of k^a take the form given in equations (3.36). Considering the case at the outer boundary where $r_\gamma = r_b$, all that remains is to determine the values of h_γ and l_γ . Matching the components of k^a at this boundary,

$$\begin{aligned} k_{\text{FRW}}^3 &= k_{\text{Sch}}^3 \\ \frac{pq}{\mathcal{S}^2(\omega_f)R^2(cT_f)} &= \frac{h_\gamma}{r_b^2} \\ \therefore h_\gamma &= pq . \end{aligned} \quad (3.47)$$

To find the value of l_γ , we consider the frequency scalar $u^a k_a$, where u^a is the 4-velocity of a timelike observer. In the FRW region, the comoving boundary observer has a 4-velocity of $u^a = (1, 0, 0, 0)$. From (3.13), $k_0 = \frac{p}{R}$, so at the boundary,

$$(u^a k_a)_{\text{FRW}} = \frac{p}{R(cT_f)} . \quad (3.48)$$

Using the expression for k^a in the Schwarzschild region given in (3.36),

$$k_a = g_{ab}k^b = \left(l_\gamma, -\frac{\dot{r}_\gamma}{f(r_\gamma)}, 0, -h_\gamma \right) , \quad (3.49)$$

which can be combined with the vacuum expression for the 4-velocity of the boundary observer given in (3.37) to (3.39) to find the frequency scalar has the form

$$(u^a k_a)_{Sch} = \frac{l_b}{f(r_b)} l_\gamma - \epsilon_b \sqrt{l_b^2 - f(r_b)} \frac{\dot{r}_\gamma}{f(r_\gamma)} \quad (3.50)$$

there. Equating the scalars from the two coordinate systems at the boundary between them,

$$\frac{pf(r_b)}{R} - l_b l_\gamma = -\epsilon_b \epsilon_\gamma \sqrt{l_b^2 - f(r_b)} \sqrt{l_\gamma^2 - \frac{h_\gamma^2}{r_b^2} f(r_b)}, \quad (3.51)$$

where $R \equiv R(ct_f)$ and use was made of the fact that at the boundary, $r_\gamma = r_b$. Squaring and re-arranging further yields a quadratic in l_γ that can be solved to give

$$l_\gamma = \frac{p r_b l_b \pm \sqrt{(p^2 r_b^2 - h_\gamma^2 R^2) (l_b^2 - f(r_b))}}{r_b R}. \quad (3.52)$$

The choice of sign can be made by comparing this to similar expressions in [Dyer \(1973\)](#). When a beam is being followed backwards from observer to source through an expanding universe, $\frac{dR}{dcT} < 0$ and the positive sign is the appropriate one to use when it enters a hole. On the other hand, if the beam is leaving a core whose radius is evolving in the same direction as that of the hole boundary, the negative sign should be chosen. In practice then, this sign is given by the product $\epsilon_\gamma \epsilon_R$, where the two factors are from (3.36b) and (3.34) respectively.

Determining the Geometrized Mass

In order to evaluate $f(r)$, the geometrized mass m needs to be computed via (1.5) using the gravitating mass M . This quantity is determined by the matching condition at the boundary of the FRW and Schwarzschild regions:

$$M = \frac{4\pi}{3} r_b^3 \rho_b, \quad (3.53)$$

where the Schwarzschild radius can be found from (3.46), and the background FRW mass density can be evaluated using the relation $R^3 \rho_b = R_0^3 \rho_0$ ([Dyer, 1973](#)). In the case of zero

curvature, this mass is the same quantity one would get by taking the volume integral over the hole were it filled with dust at the FRW background density,

$$M_{vol} = \rho_b R^3 \int_0^{\omega_b} dV ,$$

since $r_b = R\omega_b$. However, this is not the case in general. When the curvature is positive, (3.53) can be written as

$$M = \frac{4\pi}{3} R^3 \sin^3(\omega_b) \rho_b ,$$

while the volume integral (refer to Appendix A) gives

$$M_{vol} = 4\pi R^3 \left(\frac{\omega_b}{2} - \frac{\sin(2\omega_b)}{4} \right) \rho_b .$$

Here, $M < M_{vol}$ as the binding energy present lowers the gravitational source term. Correspondingly, when the curvature is negative, (3.53) becomes

$$M = \frac{4\pi}{3} R^3 \sinh^3(\omega_b) \rho_b ,$$

which is larger than the volume integrated mass

$$M_{vol} = 4\pi R^3 \left(\frac{\sinh(2\omega_b)}{4} - \frac{\omega_b}{2} \right) \rho_b$$

as a consequence of the kinetic energy present increasing the gravitational source.

Determining l_b

The only other unknown in (3.52) is l_b . Using (C.11a), along with the definition of the geometrized mass, the Friedmann equation (1.12) can be written as

$$\frac{dr_b}{dcT_b} = \epsilon_b \sqrt{\frac{2m}{r_b} + \frac{\Lambda r_b^2}{3} - K_b \mathcal{S}_K^2(\omega_b)} . \quad (3.54)$$

Since the motion of the boundary observer can be described using their proper time, T_b , (3.54) can be compared to u^1 in (3.39), and the two are seen to agree provided that

$$l_b = \pm \sqrt{1 - K_b \mathcal{S}_K^2(\omega_b)} . \quad (3.55)$$

From (3.38), it is clear that the positive sign should be chosen for a beam outside an event horizon to be travelling forward in coordinate time. Equation (3.55) is in agreement with

$$\frac{dct}{dcT_b} = \pm \frac{\sqrt{1 - K_b \mathcal{S}_K^2(\omega_b)}}{f(r_b)},$$

the connection condition at the boundary (Dyer, 1973).

3.3.2 Schwarzschild Vacuum to FRW Background

Once a beam has propagated through a hole and reached the boundary of the vacuum again, its location and direction must be converted back into the RW coordinates describing the space outside. With the θ and ϕ coordinates being common to the two systems, and the comoving angular size of the hole, ω_b , remaining constant, the location of the beam on the boundary is known. The scale parameter can be computed via (C.11a), while the new value of the RW coordinate time can be set to zero since its absolute value does not matter. The procedure to convert the beam direction between the two coordinate systems as it leaves the hole is exactly analogous to that used when the beam entered. In this case though, it is the constants p and q that need to be found, rather than h_γ and l_γ . Since the frequency scalar $u^a k_a$ can be computed, (3.51) can be used to find p , which in turn can be used in (3.47) to find q . With the location and direction vector of the beam known, it can be rotated out of the propagation orientation it was put into to travel to the lens. This is accomplished by inverting the rotations described in Section 3.1.1, putting the beam back into unoriented background coordinates so the search can begin for the next sphere encounter.

3.3.3 Core Dynamics

Slow Versus Lagging Cores

Though the RSC model itself does not impose any restriction on the dynamics of the cores it contains, in this work, the choice was made to describe these cores as FRW

regions. As such, their evolution is governed by the Friedmann equation (1.12). While one should not attach too much significance to the early evolutionary behaviour of the model, since shortly after the FRW “Big Bang”, the pressure is not negligible as has been assumed, it is nonetheless useful to develop a consistent picture of how the cores might have evolved.

Perhaps the simplest scenario envisions the cores as being spherical regions of matter which started expanding in the Bang with the rest of the universe, but that had an initial velocity that was slightly below that of the background, due, for instance, to a slight density perturbation. These *slow* cores would continue to expand at a slower rate, allowing vacuum regions to grow between their outermost dust shell and the innermost shell of the background dust. While aesthetically pleasing due to its simplicity, this picture ends up being rather restrictive in practice. As the model evolves, there is a narrowing of the range of core properties that allows them to grow to a reasonable fraction of the boundary size without recollapsing or remaining very close to the boundary for most of their evolutionary history. This results in a model with few non-collapsed regions of significantly higher density. An alternative picture which solves this problem is to lift the restriction that the cores and background begin expanding simultaneously. In this situation, the vacuum region around a singular core grows as the background universe does, until the delayed or *lagging* core experiences its own bang and subsequent expansion. Care must be taken in this case to ensure that the matter in the core does not end up expanding into the matter in the background, but allowing the beginning of the core expansion to be a free parameter results in an RSC model that displays interesting structure over a wider timescale.

Following the Motion of the Hole Boundary and the Core

As was seen in Section 3.2, when expressed using Schwarzschild coordinates, the 4-velocity of a timelike observer comoving with the FRW dust at the edge of a vacuum region takes

the form:

$$u^a = \left(\frac{l}{f(r)}, \epsilon \sqrt{l^2 - f(r)}, 0, 0 \right). \quad (3.56)$$

This applies to both the boundary and core observers, though each of these measure their 4-velocity with respect to their own proper time τ . In order to follow the motion of these observers while at the same time tracking a beam whose trajectory is governed by (3.36), a consistent way of synchronizing the different systems is required. The code needs a common basis to work with in order to be able to determine by how much the boundary and core have evolved each time the beam is stepped by some amount $\Delta\lambda$.

The most obvious candidate for this basis may seem to be the Schwarzschild coordinate time ct , which is what a Killing observer (an observer remaining at a constant radius in the vacuum) would measure as their proper time. However, there are problems with this measure. Considering the case where $\Lambda = 0$ for simplicity, there is an event horizon at $r = 2m$, and as this is approached, the value of ct becomes infinite as is clear from the Kruskal diagram in Figure 3.5, or by considering u^0 in the limit that $f(r) \rightarrow 0$. As can be seen from the metric (1.2), when $r < 2m$, t and r are no longer properly defined as timelike and spacelike quantities respectively. Even if the code need not be concerned with beams that fall inside an event horizon, it will need to follow the motion of the boundary and cores from $r = 0$, and though it is possible to integrate over the problematic point, this system does not easily lend itself to numerical integration. As such, it is useful to establish a different coordinate system in which the quantities of interest behave more reasonably and a relativistically valid synchronization over the vacuum can be implemented.

The Infalling Coordinate System

To define this new system, consider a Killing observer located at some large value of $r = r_K$ outside a core with geometrized mass m . This observer will actually lie inside the background FRW region in the model, but this is not a problem as neither the beam,

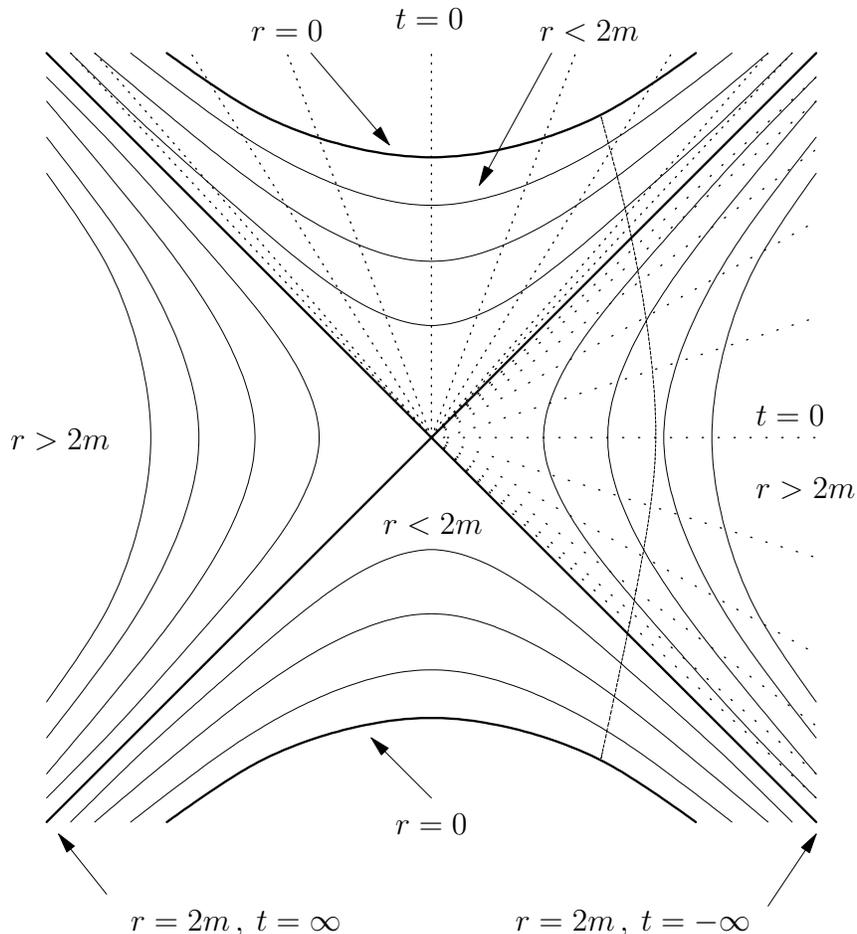


Figure 3.5: A Kruskal diagram for $\Lambda = 0$. The solid lines represent curves of constant r , with the heavy solid lines denoting $r = 0, 2m$. The dotted lines denote curves of constant t , while the dashed line shows the motion of a timelike particle leaving the “white hole” at the bottom of the diagram and eventually ending up in the “black hole” at the top.

nor the boundary, nor the core will need to travel through the vacuum to r_K . At regular intervals of their proper time, this observer releases an infalling radial “signal photon” and a “pebble” (which is simply a test mass representing a freely-falling, timelike observer). The photon and pebble fall at different rates, the former following a null geodesic while the latter follows a timelike path, and events occurring in the vacuum radially inward from $r = r_K$ can be described in terms of what photon and what pebble are at the particular location of the event at the instant it happens.

This description can be generalized to have a group of Killing observers, one at each

point of the spherical surface at $r = r_K$, all behaving in the same way and releasing photons and pebbles at regular intervals as measured by their clocks. Hence, instead of having photons and pebbles falling along a single axis through the space, there exist ingoing wavefronts and spherical shells of test masses. In this way, any event occurring within the volume enclosed by the group of Killing observers can be described in terms of the wavefront and shell that it occurs at. Since physically, the photons and pebbles must pass through the coordinate singularity at $f(r) = 0$ without any difficulty, they can be used to make measurements in any region of the Schwarzschild vacuum.

To formally establish this coordinate system, consider the motion of a signal photon. Choosing it to be travelling in the plane with $\theta = \pi/2$, from (3.36), the radial motion means that $h_\gamma = 0$, so the trajectory is described by:

$$\frac{dct}{d\lambda} = \frac{l_{\gamma r}}{f(r)} \quad (3.57a)$$

$$\frac{dr}{d\lambda} = \epsilon_{\gamma r} l_{\gamma r} \quad (3.57b)$$

where the γr subscripts on the constants are used to differentiate the signal photon from a general one. Eliminating the affine parameter λ ,

$$dct = \frac{\epsilon_{\gamma r}}{f(r)} dr, \quad (3.58)$$

and the same can be done using (3.56), the 4-velocity for a timelike radial observer, to get

$$dct = \frac{\epsilon_p l_p}{f(r) \sqrt{l_p^2 - f(r)}} dr, \quad (3.59)$$

which applies to the pebble.

Taking $\Lambda = 0$ again for simplicity, $f(r) = 1 - \frac{2m}{r}$. The signal photons have $\epsilon_{\gamma r} = -1$, so from (3.58),

$$\int dct = - \int \frac{dr}{1 - \frac{2m}{r}}$$

and hence,

$$ct + r + 2m \ln |r - 2m| = \nu, \quad (3.60)$$

where ν is an integration constant. Using (3.59) to do the same for the pebbles,

$$\int dct = - \int \frac{l_p dr}{\left(1 - \frac{2m}{r}\right) \sqrt{l_p^2 - 1 + \frac{2m}{r}}}. \quad (3.61)$$

If the Killing observer drops the pebbles with an initial velocity such that $l_p = 1$, then

$$\int dct = \frac{1}{\sqrt{2m}} \int \frac{r\sqrt{r}}{2m - r} dr$$

and as a result,

$$ct + 2\sqrt{2mr} + \frac{2r^{\frac{3}{2}}}{3\sqrt{2m}} - 2m \ln \left| \frac{\sqrt{r} + \sqrt{2m}}{\sqrt{r} - \sqrt{2m}} \right| = \mu \quad (3.62)$$

for the integration constant μ . Note that when considering the motion of the pebbles, the requirement that $l_p = 1$ is equivalent to saying that they are released by the Killing observer with the same velocity that they would have if they fell in from an initial zero velocity at $r = \infty$. In fact, any value could have been chosen for l_p and the expressions would still be integrable analytically. The case chosen just results in equations that are easier to work with.

Examining equations (3.60) and (3.62), it is clear that the Killing observers can use the values of ν and μ to respectively label the photon and pebble that are released at a given value of their proper time t . In other words, at time t_0 , each observer releases photon ν_0 and pebble μ_0 , with photon ν_n and pebble μ_n being released at time t_n . These equations allow us to take an event occurring at a given Schwarzschild (ct, r, θ, ϕ) coordinate, and convert it to an “infalling” (ν, μ, θ, ϕ) coordinate. Though the system can just as easily be constructed with outgoing beams and pebbles, taking care to manage the signs properly will allow almost all situations to be covered using this infalling system. Refer to Figure 3.6 to see this system depicted in a Kruskal diagram. Note that the above procedure can still be followed for non-zero values of Λ . The only difference is that (3.60) and (3.62) will involve elliptic integrals that need to be performed numerically in order to get ν and μ from values of t and r .

As the angular θ and ϕ coordinates are common to the Schwarzschild and infalling coordinate systems, the conversion of some tensorial quantity v^a between the two only

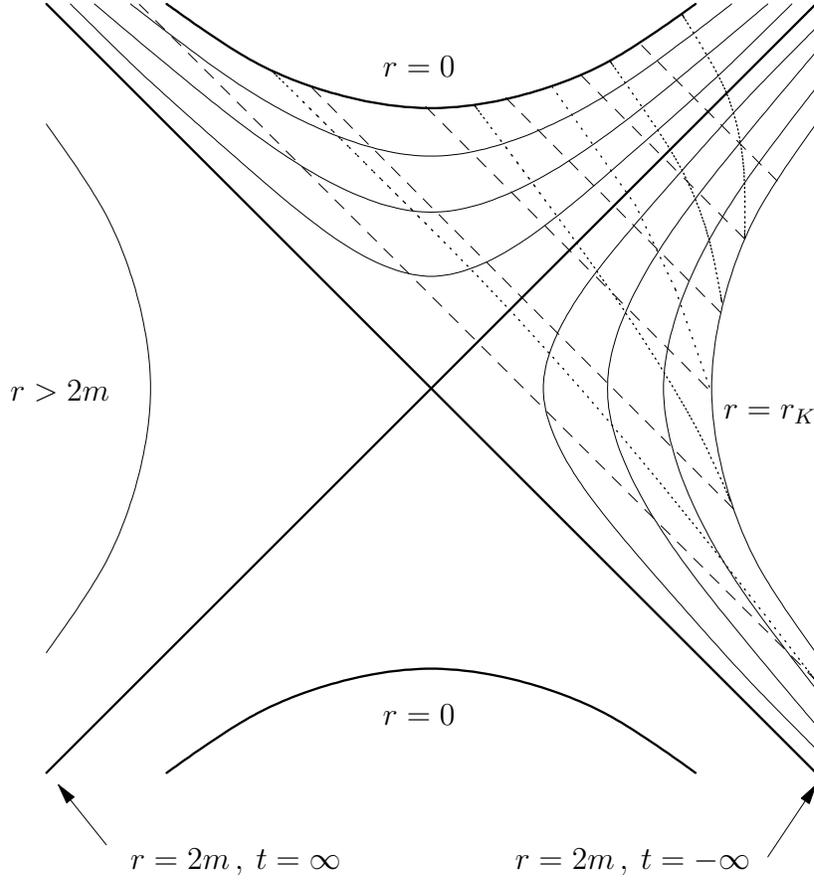


Figure 3.6: The infalling coordinates depicted in a Kruskal diagram. The dashed lines represent constant values of ν for the ingoing beams, while the dotted lines denote constant values of μ for the infalling pebbles. They all originate from a Killing observer at $r = r_K \gg 2m$.

involves the transformation between (ct, r) and (ν, μ) . Using (3.60) and (3.62) in (C.1a),

$$\begin{aligned}\tilde{v}^0 &= v^0 + \frac{r}{r-2m}v^1 \\ \tilde{v}^1 &= v^0 + \frac{r\sqrt{r}}{\sqrt{2m}(r-2m)}v^1,\end{aligned}$$

where \tilde{v}^a is expressed in the infalling system. Applying this to the 4-velocity (3.56),

$$\begin{aligned}\frac{d\nu_b}{d\tau_b} &= \frac{r}{r-2m} \left(l_b + \epsilon_b \sqrt{l_b^2 - f(r_b)} \right) \\ \frac{d\mu_b}{d\tau_b} &= \frac{r}{r-2m} \left(l_b + \epsilon_b \sqrt{\frac{(l_b^2 - 1)r}{2m} + 1} \right),\end{aligned}$$

giving

$$\frac{d\mu_b}{d\nu_b} = \sqrt{\frac{r_b}{2m}} \left(\frac{\sqrt{2m} l_b + \epsilon_b \sqrt{(l_b^2 - 1) r_b + 2m}}{\sqrt{r_b} l_b + \epsilon_b \sqrt{(l_b^2 - 1) r_b + 2m}} \right) \quad (3.63)$$

for the boundary observer. The same expression holds true for the core observer (simply replace each “ b ” subscript with a “ c ”). Note that in this expression, the quantity $(l_b^2 - 1)r_b + 2m$ cannot be negative as otherwise the u^1 component of the 4-velocity will be complex, which cannot happen for physical observers. Hence, all quantities in this equation remain real.

Repeating the process for k^a , the tangent vector to the null geodesic followed by the beam given in (3.36),

$$\frac{d\nu_\gamma}{d\lambda} = \frac{l_\gamma r_\gamma \sqrt{r_\gamma} + \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)}}{\sqrt{r_\gamma} (r_\gamma - 2m)} \quad (3.64a)$$

$$\frac{d\mu_\gamma}{d\lambda} = \frac{l_\gamma r_\gamma \sqrt{2m} + \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)}}{\sqrt{2m} (r_\gamma - 2m)} \quad (3.64b)$$

$$\frac{d\theta_\gamma}{d\lambda} = 0 \quad (3.64c)$$

$$\frac{d\phi_\gamma}{d\lambda} = \frac{h_\gamma}{r_\gamma^2} \quad (3.64d)$$

and eliminating the affine parameter λ ,

$$\frac{d\mu_\gamma}{d\nu_\gamma} = \sqrt{\frac{r_\gamma}{2m}} \left(\frac{l_\gamma r_\gamma \sqrt{2m} + \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)}}{l_\gamma r_\gamma \sqrt{r_\gamma} + \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)}} \right) \quad (3.65a)$$

$$\frac{d\theta_\gamma}{d\nu_\gamma} = 0 \quad (3.65b)$$

$$\frac{d\phi_\gamma}{d\nu_\gamma} = \frac{h_\gamma (r_\gamma - 2m)}{r_\gamma \sqrt{r_\gamma} \left(l_\gamma r_\gamma \sqrt{r_\gamma} + \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)} \right)} \quad (3.65c)$$

Again, the quantity $r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)$ must be greater than or equal to zero in order for the k^1 component of the tangent vector to remain real. Hence, the quantities in these expressions are all real. Since a given set of signal photons released by the Killing observers is used to synchronize all events, $\nu_b = \nu_c = \nu_\gamma = \nu$, so the differential equations (3.65),

along with (3.63) and a similar expression for the core, can be integrated simultaneously to track the locations of the boundary observer, core observer, and beam, using ν as an independent variable.

Note that for the boundary (or core) observer, if $\epsilon_b = -1$ and $l_b = 1$ as is the case for the pebbles, $\frac{d\mu}{d\nu} = 0$ as expected since each test mass has a constant value of μ . Furthermore, for the beam, if $h_\gamma = 0$ and $\epsilon_\gamma = -1$ as is true for the signal photons, then $\frac{d\mu_\gamma}{d\nu} = \infty$ and L'Hopital's rule shows that the same is true for $\frac{d\phi_\gamma}{d\nu}$ as is expected for these beams since they are defined by a constant value of ν . This does imply that a purely radial ingoing beam cannot be treated properly in this coordinate system, but in practice, this should not be a common occurrence.

In order to use the propagation equations above, the relevant values of r are still needed for the computations. Combining (3.60) and (3.62) to eliminate ct , yields

$$\nu - \mu - r - 4m \ln \left(\sqrt{r} + \sqrt{2m} \right) + 2\sqrt{2mr} + \frac{2r^{\frac{3}{2}}}{3\sqrt{2m}} = 0. \quad (3.66)$$

Though this cannot be solved directly for r , given values of ν and μ , it can easily be solved numerically as it is well defined for all positive r and is a monotonically decreasing function.

While the infalling coordinate system (ν, μ, θ, ϕ) can be used to follow the evolution of the system throughout the Schwarzschild vacuum, if our interest lies only outside the event horizon at $2m$, then things can be simplified considerably. In particular, there is no longer any need for the μ coordinate, and the various values of r can be integrated directly instead of being computed using (3.66) at every step.³ Converting (3.56) to the (ν, r, θ, ϕ) system and eliminating the affine parameter, the equation of interest becomes

$$\frac{dr_{bc}}{d\nu} = -\frac{\sqrt{r_{bc}l_{bc}^2 - (r_{bc} - 2m)}}{r_{bc}} \left(\sqrt{r_{bc}l_{bc}^2 - (r_{bc} - 2m)} - \epsilon_{bc}l_{bc}\sqrt{r_{bc}} \right), \quad (3.67)$$

where the subscripts indicate that this can be used for both the boundary and core

³The system effectively becomes one described by Eddington-Finkelstein coordinates – see Appendix C.2.

observers. This expression has been written in a form that is numerically stable, even as an observer approaches $2m$. Equation (3.64a) can be re-written as

$$\frac{d\lambda}{d\nu} = \frac{\sqrt{r_\gamma}}{h_\gamma^2} \left(l_\gamma r_\gamma \sqrt{r_\gamma} - \epsilon_\gamma \sqrt{r_\gamma^3 l_\gamma^2 - h_\gamma^2 (r_\gamma - 2m)} \right), \quad (3.68)$$

which also has no problem near $2m$. It can be used to convert the differential equations describing the trajectory of the beam and the optical scalars from using λ as an independent parameter to using ν .

There is an interesting observation that should be made regarding the motion of the core and boundary observers. If they are comoving with a dynamically closed FRW region and they begin their trajectory moving outward from an initial singularity at $r = 0$, then there is a minimum radius out to which they must travel before they can fall inward again. Consider the radial component of their 4-velocity in the simple case where $\Lambda = 0$:

$$\frac{dr}{d\tau} = \epsilon \sqrt{l^2 - \left(1 - \frac{2m}{r}\right)}.$$

Initially, $\epsilon = +1$ and the observer moves outward until $\frac{dr}{d\tau} = 0$, at which point ϵ changes sign and the observer starts to fall inward. The maximum value of r attained is clearly dependent on l , the constant of motion related to energy. For instance,

$$\left. \frac{dr}{d\tau} \right|_{r=4m} = 0 \quad \implies \quad l = \sqrt{0.5},$$

so a closed trajectory that turns around at $4m$ requires $l = \sqrt{0.5}$. As the turnaround point is moved further and further out, l must increase, until in the limit,

$$\left. \frac{dr}{d\tau} \right|_{r=\infty} = 0 \quad \implies \quad l = 1,$$

which is the marginally bound case. As l is increased beyond 1, the derivative is never zero so there is no turn-around and the observer simply has an increasing velocity at $r = \infty$ corresponding to an increasingly unbound system as can be expected.

However, as l is reduced, the system becomes additionally bound until a turn-around point of $r = 2m$ is reached, in which case $l = 0$. Hence, any outgoing physical observer

must reach a radial value of at least $2m$ before turning around. This is of course easy to see in a Kruskal diagram (refer to Figure 3.5), where a timelike particle travelling outward in the “white hole” region at the bottom of the diagram cannot return to $r = 0$ without at least reaching $2m$ and passing through to the “black hole” region at the top of the diagram. Since the motion of these observers can also be described using the RW metric, the turnaround point must also correspond to the extremum of the Friedmann equation when $\frac{dR}{dcT} = 0$, giving

$$R_{\max} = \frac{8\pi G\rho_0 R_0^3}{3c^2}$$

for $\Lambda = 0$ and $K = 1$. On the surface, this may appear to exclude the possibility of a uniform FRW region expanding then collapsing within a vacuum. However, if considered from within a Schwarzschild coordinate system, the smaller the radial distance of an FRW dust particle from the centre, the smaller the value of $2m$ it experiences, and the smaller the minimum turnaround radius is, so there is actually nothing inconsistent with the description given. As an interesting aside, the minimum time to live for an expanding region is $2m\pi$, which, for a region with a galactic-type mass of $6 \times 10^{11} M_\odot$, corresponds to about 20 days.

3.3.4 Defining the Core

With the RSC model being constructed as it is observed, the first time that a sphere is encountered by a beam, the code must choose the properties of the lens it contains inside. Since the value of M , the mass inside the hole, is already determined, r_c and l_c , the radius of the core and a constant related to its energy, are all that remain to be selected. Care must be taken in choosing these values, to ensure that they do not allow for the core to overlap with the boundary at any point in their evolution. At first pass, we assume for simplicity that the cores, like the background universe, do not reach a turn-around point during a beam propagation interval. Though this condition extends the self-similarity of the model, it is not essential and dropping it will in fact add a bit

more flexibility to the model.

Choosing r_c

When the beam being propagated reaches the boundary observer, the radius of the core will lie between 0 and r_b , the radius of the boundary. To choose a value, a positive scale factor $s < 1$ is picked at random. Currently, this factor is chosen from among a small number of values, each of which is an inverse power of two. However, this procedure can easily be made more sophisticated to produce results based on some predetermined distribution. While the most straightforward choice would be to use $r_c = sr_b$, as the beam moves further back in redshift, it is to be expected that on average, the boundaries and cores should be closer together as there has been less time elapsed for them to evolve apart. To roughly account for this, r_c is chosen via

$$r_c = \left(s + (1 - s) \frac{z}{z_{\max}} \right) r_b .$$

In this way, the full scale factor is applied when $z = 0$, while in the limit of the maximum redshift that the beam will be propagated to, $r_c = r_b$ and the hole has no core (it becomes a leaf as described in Section 2.2.1). It should be emphasized that the value of r_c chosen is not the result of an instantaneous measurement at a distance. Rather, it is the radius that the boundary observer would measure the core to have by communicating with the core observer using radial light rays. In other words, it is not the size of the core as measured on an acausal spacelike “sheet” across the vacuum region, as it takes into account signal propagation time in order to ensure that the system is properly synchronized at the location of the propagated beam.

Choosing l_c

With r_c chosen, a suitable value of l_c needs to be selected. The larger the value of this constant, the more energetic the core is and the quicker a given value of r_c can be reached.

From the conditions imposed above, the maximum value that could be used is l_b . Any larger, and over the course of its evolution, the core will expand into the boundary. In general, the minimum value of l_c for a dynamic core is anything greater than zero. In the case where $K = 1$ though, the core will at some point reach a maximum extent and hence, must be given enough energy to attain the designated value of r_c before turning around and collapsing. While a trial value of l_c can be chosen at random to lie between these limits, this scheme is not sufficient. Though the core may be given enough energy to eventually reach the chosen radius, it may still expand too slowly, reaching the required size only if it is allowed to start its expansion earlier than the boundary of the hole does, an unreasonable proposition. As a result, before settling on the value of l_c , the code must ensure that evolution can occur in a sensible manner.

In the infalling coordinate system adopted, elapsed time can be measured in a consistent way by considering intervals in ν . Applying (3.67) to the boundary observer and rearranging,

$$\Delta\nu = \int_{r_{b_i}}^{r_{b_f}} \frac{r_b}{r_b - 2m} \left(\frac{\epsilon_b l_b \sqrt{r_b}}{\sqrt{(l_b^2 - 1)r_b + 2m}} + 1 \right) dr_b. \quad (3.69)$$

As the elapsed interval in ν can always be computed by considering the boundary to be infalling and setting the limits accordingly, we can set $\epsilon_b = -1$. In the case where $l_b = 1$, that is, where the background FRW region is flat, the integral simplifies and

$$\Delta\nu \Big|_{l_b=1} = \left[r_b + 4m \ln \left(\sqrt{r_b} + \sqrt{2m} \right) - \frac{\sqrt{2mr_b}}{3m} (6m + r_b) \right]_{r_{b_i}}^{r_{b_f}}, \quad (3.70)$$

which is well behaved for all r_b . For other values of l_b , the analytic integral contains complex values, but since these are constants that cancel out when considering differences in ν , the result is

$$\Delta\nu \Big|_{l_b \neq 1} = \left[r_b - 2m \ln \left(\frac{2m + r_b(2l_b^2 - 1) - 2l_b \sqrt{(2m + (l_b^2 - 1)r_b)r_b}}{(r - 2m)^2} \right) + F(l_b, m, r_b) - \frac{l_b}{l_b^2 - 1} \sqrt{r_b(2m + (l_b^2 - 1)r_b)} \right]_{r_{b_i}}^{r_{b_f}}, \quad (3.71)$$

where

$$F(l_b, m, r_b) = \begin{cases} \frac{2ml_b(2l_b^2-3)}{\sqrt{(1-l_b^2)^3}} \operatorname{atan} \left(\sqrt{\frac{(1-l_b^2)r_b}{2m+(l_b^2-1)r_b}} \right) & : l_b < 1 \\ -\frac{ml_b(2l_b^2-3)}{\sqrt{(l_b^2-1)^3}} \ln \left(\frac{m+r_b(l_b^2-1)+\sqrt{r_b(l_b^2-1)(2m+(l_b^2-1)r_b)}}{m} \right) & : l_b > 1 \end{cases},$$

all of which is real. While the logarithm in (3.71) appears to be undefined when $r = 2m$, a couple of applications of L'Hopital's rule shows that it is indeed finite, with

$$\left[\frac{2m + r_b(2l_b^2 - 1) - 2l_b\sqrt{(2m + (l_b^2 - 1)r_b)r_b}}{(r - 2m)^2} \right]_{r_b=2m} = \frac{1}{8ml_b^2}, \quad (3.72)$$

so the expression for $\Delta\nu$ is valid for all values of r_b as expected from Figure 3.6. Whenever the function needs to be evaluated at a value of r_b close to $2m$, the potentially problematic argument can be expanded as a third order Taylor series about this point,

$$\begin{aligned} \frac{2m + r_b(2l_b^2 - 1) - 2l_b\sqrt{(2m + (l_b^2 - 1)r_b)r_b}}{(r - 2m)^2} &\approx \\ \frac{1}{8ml_b^2} \left(1 + \frac{(1 - 2l_b^2)(r_b - 2m)}{4ml_b^2} + \frac{(16l_b^2(l_b^2 - 1) + 5)(r_b - 2m)^2}{64m^2l_b^4} \right. \\ &\quad \left. + \frac{(-32l_b^6 + 48l_b^4 - 30l_b^2 + 7)(r_b - 2m)^3}{256m^3l_b^6} \right), \end{aligned} \quad (3.73)$$

and used inside the logarithm instead. In practice, using the Taylor series when $1.98m \leq r_b \leq 2.02m$ seems to avoid any numerical problems that can otherwise arise trying to evaluate the analytic solution directly.

The above expressions allow $\Delta\nu_b$ and $\Delta\nu_c$ to be computed. If the values of r_c and l_c chosen result in $\Delta\nu_c > \Delta\nu_b$, then l_c is increased via $l_c = 0.5(l_b + l_c)$ until $\Delta\nu_c$ drops below $\Delta\nu_b$, though if the increment in l_c reaches floating point precision, then the core is simply designated as a leaf since it would end up very near the boundary anyhow. Otherwise, the combination of r_c and l_c obtained is valid, and the difference between the elapsed intervals in ν is just the amount by which the core is delayed. M and r_c can be used to set ρ_{c0} , the fiducial density of the core, while l_c allows the curvature indicator of the core

to be set via an expression analogous to that for the boundary in (3.55), so

$$K_c = \begin{cases} -1 & : l_c > 1 \\ 0 & : l_c = 1 \\ +1 & : l_c < 1 \end{cases} .$$

Knowing K_c , the expression can then be solved for ω_c to find the angular extent of the core.

Subsequent Lens Encounters

If a beam hits a hole containing a lens that has already been defined by a previous beam encounter, then a similar approach can be followed to find the appropriate value of r_c (l_c , a constant of the motion of the core, would have already been fixed). It must lie between zero and r_b , or the radius of maximum extent if $K_c = 1$. This turn-around radius can be found by following the same procedure described in Section 3.2.1, but using the radial component of the core observer's 4-velocity from (3.56) instead of the expression for $dr_\gamma/d\lambda$ used in that section. With the limits on r_c known, the value that yields a ν_c satisfying the expression

$$\Delta\nu_b - (\nu_c|_{(r_c=0)} - \nu_c) - \text{core_delay} = 0$$

can be computed using a root-finding scheme based on the `zbrent` routine in Numerical Recipes (Press et al., 1992). This is a bracketing method that uses quadratic interpolation to make it quicker than a conventional Regula-Falsi method without requiring the evaluation of derivatives.

There is a potential problem that can arise with the chosen method of defining lenses if the first beam to hit a large lens only grazes it. Since this beam is used to set the core size, the core delay that is subsequently implied may end up being too short to allow the relation

$$\Delta\nu_b = \Delta\nu_c + \text{core_delay} \tag{3.74}$$

to hold for a later beam that hits the hole in a much more direct fashion. This more direct beam, travelling from the final observer back to its source, will reach the hole when it is much larger than it was when encountered by the grazing beam. Consequently, the interval $\Delta\nu_b$ that needs to elapse in order for the boundary to evolve to the larger size may be longer than the combination of the core delay established earlier and the interval needed by the core to reach its maximum extent. That is to say, the core should have turned around by the time the direct beam reaches the boundary of the hole. If this proves to be a significant problem in practice, the simplifying restriction that cores not be allowed to turn around can be lifted, thereby allowing the model to continue following the core as it re-contracts, either to remain a final collapsed object such as a black hole, or even re-expanding again in a cyclic fashion.

Storing the Definition

As indicated earlier, the observed RSC model can be stored in memory as a hierarchical tree structure, with each node in the structure describing a portion of the model at a particular level. The uppermost node describes the background universe that the model is embedded in, while the lowest levels represent holes that either contain a spherical lens with no further substructure, or that are uniformly filled with matter at the same density as the background in which they are inscribed (i.e. they are not holes at all, but the leaves described in Section 2.2.1). Each node then, can contain the following information:

R_0 : The scale factor of the core determined when it is first defined.

ρ_0 : The density of the core determined when it is first defined.

ω_c : The angular extent of the core.

l_c : The constant of motion for the core.

K_c : The curvature indicator of the core.

delay : The interval in ν that elapses between the beginning of the boundary expansion and the beginning of the core expansion.

id : The sphere representing the boundary of the hole in the packing used to describe the FRW region in which it is embedded.

pack : An identifier for the packing used to describe any substructure within the core.

num : The number of spheres loaded from **pack** (used as a check when a packing is re-loaded).

centre : The sphere in **pack** chosen to be used as the centre of coordinates.

orientation : A set of 4 numbers describing the orientation chosen for **pack**, consisting of the 2 planes and corresponding angles used for the random spins (refer to Section 2.2 for more).

links : Two pointers that are used to connect the node to the rest of the model. One pointer is used to connect to the next discovered hole at the same level as the current node, while the other is used to connect to the nodes describing substructure within the core of the current node.

Together, these values allow everything that needs to be known about the RSC model to be computed as needed.

3.3.5 Moving Between the Vacuum and Core

Before a beam is propagated through a core that it has hit, its location and direction must be converted to the RW coordinate system describing the region. As was the case for the outside boundary of the vacuum, the boundary of the core lies at a constant comoving radius when measured in these coordinates. In fact, with the appropriate

choice of signs, the same procedure used to follow the beam across the outer boundary can also be employed at this inner boundary.

As indicated above, when a beam passes from an FRW background region into a Schwarzschild vacuum, the system is left in propagation orientation (i.e. with the lens at the centre of coordinates and the beam having $\theta = \pi/2$ and $d\theta = 0$ as explained in Section 3.1). However, once the core has been hit, this orientation cannot be maintained as otherwise the structure inside would not be consistent for any subsequent beams that enter with different orientations. As such, it is necessary to define a core coordinate system that can be kept consistent for all beams. This is actually quite straightforward. The transformation required to place a lens at the centre of coordinates (involving m_1 to m_3 and m_{flip}) is only dependent on the coordinates of the centre of the sphere containing the lens, and thus will be identical for any beam. It is only the latter rotations (m_5 to m_7), required to place the beam at $\theta = \pi/2$ with $d\theta = 0$, that vary. Hence, once a beam hits the core, inverting these rotations will express its location and direction vector in the desired coordinate system. The beam can then be propagated through the FRW region as before, potentially recursing deeper into the model if substructure is encountered inside the lens, repeating the transformations as required.

Once the beam has propagated through the core and is set to move back into the vacuum, it will be propagated to the boundary with an orientation set by using the centre of the lens (as opposed to the centre of some substructure within the lens) as the centre of coordinates. This core propagation orientation can also be used to follow the beam back through the vacuum as it satisfies the conditions needed for that purpose. When the beam reaches the outer boundary of the vacuum, some care is needed to ensure the correct re-orientation is performed there. First, the system needs to be taken out of the core propagation orientation. Since the centre of the lens used to establish this orientation was already at the centre of the coordinate system, this only involves the inversion of rotations m_5 to m_7 . Then, the system needs to be returned to the unoriented

coordinate system of the background. This is done by inverting the transformations m_1 to m_3 and m_{flip} that were originally used to place the lens at the centre of coordinates before the beam was propagated to it. The other rotations required to establish the background propagation orientation last used were inverted when the beam hit the core as was described above.

If the core has a different curvature than the background in which it is embedded, it may seem that an issue arises as to which value of K to use to perform the transformations required. Since the rotations m_5 to m_7 do not involve a change in the ω coordinate, they are independent of K , and it is only the embedding necessary to perform them that is affected by it. For this reason, when rotating coordinate systems, the code always uses the curvature appropriate to the location of the beam (i.e. K_b at the boundary and K_c at the core). In this way, while it is possible for the rotations m_5 to m_7 to be performed using one curvature in the background and inverted using another at the core, the rotations that do involve the curvature explicitly and not just in the embedding, are always applied and reversed using the same value of K .

To find r_b when the beam emerges from a core, virtually the same procedure that was described in Section 3.3.4, to re-synchronize a core for the propagation of subsequent beams hitting the hole boundary, can be followed. First, $\Delta\nu_c$ is computed to determine the interval that elapses for the core to go from the known radius r_c to zero (remember that the computation is always done assuming infall). This is then added to the stored core delay, and the `zbrent` routine is used to find the value of r_b that leads to a $\Delta\nu_b$ satisfying (3.74). In this case, the lower limit for r_b is r_c , while the upper limit is the value of r_b that was measured when the beam first entered the hole.

3.3.6 Tracking the Redshift

One of the parameters the code requires at run time is the specification of a maximum redshift value that the beams are to be propagated back to – i.e. the redshift of their

source. In order to compute the redshift at various points along the beam path, one can compare the frequency of the beam measured by an observer at its source to that measured by an observer at its destination. As each of these observers makes the measurement in their own frame of reference, the frequency can be determined by taking the inner product of their 4-velocity with the null tangent vector to the beam at their location. Hence, using (3.48), the redshift factor between two observers comoving with an FRW region can be written as

$$1 + z = \frac{(u^a k_a)_{\text{emit}}}{(u^a k_a)_{\text{obs}}} = \frac{p_e R_o}{p_o R_e}.$$

For a beam propagating through a single FRW region in the RSC model, such as from the boundary of one hole to that of another, the propagation constant p does not change so the redshift factor is simply

$$1 + z = \frac{R_o}{R_e}, \quad (3.75)$$

which is the standard form for an FRW universe with no holes. Since the beam is being propagated backwards from the observer to the source in an expanding universe, the scale factor at the end of the propagation (when the beam was emitted) is smaller than at the beginning, giving $z \geq 0$ as required. As the scale factor is already followed via (3.35) in order to keep track of the null tangent vector k^a , the redshift across such regions can be computed at any point desired. Notice that equation (3.75) is only dependent on the scale factor, emphasizing that the cosmological redshift between two comoving observers in an FRW region is due simply to photons travelling through the expanding space between them, and is not the result of a Doppler shift. In fact, it is not possible to give an unambiguous meaning to the relative velocities of two objects separated by a large distance in a non-Euclidian space (see Narlikar, 1993, for instance).

While it is easy to define a comoving observer at any point of an FRW region, the situation is not as clear-cut in a Schwarzschild vacuum. The natural observers to use there would probably be Killing observers that remain at constant r , though this is only

valid outside $2m$ since there are no stationary observers inside. With this choice, the beam would experience a potentially significant jump in redshift each time it passes from an FRW region to a vacuum due to the relative motion between the comoving observers at the FRW boundary and the Killing observers in the vacuum. This jump might then be largely cancelled for the same reason once the beam makes the transition from the vacuum to the core.

An alternate and more reasonable approach, given the lack of physical observers in the vacuum, is to compute the change in redshift as it would be determined between two comoving FRW observers at the edge of each vacuum that the beam is to pass through. These would be the boundary observer and core observer if the beam propagates from the boundary through the vacuum to the core, and two boundary observers if the beam leaves the hole without hitting the core. The 4-velocities of these observers have the same form, so the frequency they would measure is given by (3.50), with the appropriate replacement of subscripts being made when dealing with the core observer. Taking the ratio of these scalars as was done above, we find the redshift factor for the beam when propagating between two comoving FRW observers separated by a Schwarzschild vacuum is

$$1 + z = \frac{r_o f(r_o) \left(r_e l_e l_\gamma - \epsilon_e \epsilon_{\gamma e} \sqrt{(l_e^2 - f(r_e)) (r_e^2 l_{\gamma e}^2 - h^2 f(r_e))} \right)}{r_e f(r_e) \left(r_o l_o l_\gamma - \epsilon_o \epsilon_{\gamma o} \sqrt{(l_o^2 - f(r_o)) (r_o^2 l_{\gamma o}^2 - h^2 f(r_o))} \right)}, \quad (3.76)$$

where the “ γe ” and “ γo ” subscripts are used to describe quantities related to the beam at the emitter and observer respectively, and use was made of the relation $r_{bc} = r_\gamma$ at the measurement location. Note that there can be a blueshift across the vacuum experienced by a beam that hits a core that is expanding fast enough, but there will still be a net redshift across the entire hole. Equation (3.76) is valid everywhere except possibly where $f(r_e) = 0$, but since the code does not try to follow beams that are emitted at an event horizon (see Section 3.4 below), this is not a problem in practice.

Hence, when propagating through the model, the code simply computes the redshift

factor experienced by the beam as it passes through each region and determines the net redshift via

$$(1 + z)_{\text{net}} = (1 + z)_1 \times (1 + z)_2 \times \cdots \times (1 + z)_n ,$$

where n is the total number of FRW and vacuum regions traversed.

3.4 Ending the Propagation

In most circumstances, propagation of the beam is expected to stop when it reaches the specified source redshift z_{max} as described in the previous section. However, there are three situations that can arise which will result in the propagation being terminated sooner, with the beam being designated as invalid and its information discarded.

The first was mentioned earlier and involves those beams which are found to have originated at an event horizon. Such beams must have come from within the core and as such, do not contain information about background objects. They can only provide information about possible substructure of the core, of which there is none since the evolution of cores that lie within their own event horizon is not followed. In practice, the propagation of a beam is ended if the beam is found to have approached an event horizon to within some tolerance⁴ in order to avoid forcing the integrator to take many small steps trying to reach the horizon exactly.

Finding that a beam hits an event horizon is the only reason that propagation is stopped inside a vacuum region. Since redshifts are only computed between comoving FRW observers, a beam will never be found to have reached z_{max} while in a vacuum region. In this way, ingoing beams that would have been found to exceed the the redshift limit as measured by a Killing observer, yet would fall back within the limit once measured by the core observer (meaning they could physically have originated from within the core) are not discarded prematurely. However, for the same reason given above, beam

⁴This is set to be the distance the code considers small enough to step linearly – refer to Section 5.3.1.

propagation is also terminated if the beam is found to be trying to emerge from a core that has fallen within its own event horizon.

A beam found to hit the boundary of the background FRW region in which the RSC model is embedded is also rendered to be invalid and discarded. As the “real” observer can be placed anywhere in the background of the model and be made to look in any direction, it is possible to have them placed sufficiently close to the edge of the region considered that the beams are unable to reach z_{\max} before running out of a defined universe to propagate through. It would be straightforward to rectify this by just having the code assume that the smooth background extends as far as it needs to for the beam propagation to finish, but this would make the region containing the embedded structure a special part of the universe which goes against the cosmological principle.

Note that in practice, even valid beams will end up being propagated past the redshift limit during the final step taken by the integrator. To correct for this, a linear step is taken from the beginning or end of the interval, depending on which state is closer to the redshift limit, to place the beam at the appropriate location so that all relevant final quantities may be calculated there. Since the steps involved are small, this does not have a significant impact on the results as can be seen in Section [6.2.1](#).

Chapter 4

The Optical Scalars

4.1 The Driving Terms

In the FRW sections of the model, the Weyl tensor $C^{ab}{}_{cd}$ is identically zero, but the presence of matter within beams passing through means that Ricci focusing still takes place. As indicated earlier, for the comoving RW coordinates used, the only non-zero component of the energy momentum tensor is $T_{00} = \rho c^2$. Hence, from (1.22) and (1.23), the driving terms in the optical scalar equations take the form:

$$\mathcal{R} = -\frac{4\pi G \rho p^2}{c^2 R^2} \quad (4.1)$$

$$\mathcal{F} = 0, \quad (4.2)$$

where ρ is the mass density and p is the constant from (3.5a), not a pressure term. Since $\mathcal{F} = 0$, (1.28b) and (1.28c) imply that the quantities ξ and φ are constant here. If a beam has been shearing, it will continue to do so, but its rate of shear will not change.

There is no Ricci focusing in the vacuum regions of the model as the energy momentum tensor is equal to zero inside them. However, the dense spherical FRW cores at the centre of these regions do produce tidal forces that cause beams in the vacuum to shear. Computing the Riemann curvature tensor in this case and using it, along with the tangent vector k^a given in (3.36), the expressions for the OS driving terms (1.18) and (1.19) can

be written as:

$$\mathcal{R} = 0 \tag{4.3}$$

$$\mathcal{F} = \frac{3mh_\gamma^2}{r_\gamma^5}, \tag{4.4}$$

where h_γ is the constant defined in (3.36d), and m is the geometrized mass of the central object (Dyer, 1973).

4.2 Analyzing the Beam Cross-Section

We have already found a set of equations that, when properly integrated and connected, will allow the trajectory of a beam and the value of the optical scalars to be computed everywhere in the RSC model. At any point though, we may wish to know the cumulative effect that the expansion and shearing have had on the beam. The effect of the expansion is determined by simply tracking the area of the beam's cross-section, while the shear requires information about the shape of this cross-section. As indicated in Section 1.3.2, the shear distorts an initially circular cross-section to an elliptical one. It will be compressed along an axis in the direction of the first lens it encounters and stretched at an angle perpendicular to it. A subsequent lens with a different orientation from the original one would continue to distort the beam, resulting in a rotation of the cross-sectional ellipse as found by Dyer (1973) and Pineault (1975). Thus, the cumulative effect of shear can be determined by computing the distortion of this ellipse, along with the orientation it has. If the lengths of the semi-axes of the ellipse, a_e and b_e , are known, then the distortion, \mathcal{D} , is given by

$$\mathcal{D} = \begin{cases} 1 - \frac{a_e}{b_e} & : a_e < b_e \\ 0 & : a_e = b_e \\ 1 - \frac{b_e}{a_e} & : a_e > b_e \end{cases} . \tag{4.5}$$

Hence, when the beam is unsheared, $a_e = b_e$ and $\mathcal{D} = 0$, while at a line caustic, either $a_e = 0$ or $b_e = 0$ so $\mathcal{D} = 1$, indicating maximum distortion.

We now consider how these quantities can be followed using each of the OSE formulations described in Section 1.3.2.

4.2.1 The (A, ξ, φ) Form

In this formulation, the area of the beam is tracked directly via equation (1.28a). While it appears that numerically integrating this expression will be problematic when the beam passes through a caustic, this is not the case. As demonstrated by Harper (1991), A can be expanded in a Taylor Series about the caustic at λ_c :

$$A = \sum_{i=0}^{\infty} a_i (\lambda - \lambda_c)^i . \quad (4.6)$$

Since $A = 0$ when $\lambda = \lambda_c$, $a_0 = 0$, so to second order:

$$A = a_1 (\lambda - \lambda_c) + a_2 (\lambda - \lambda_c)^2 \quad (4.7)$$

$$\dot{A} = a_1 + 2a_2 (\lambda - \lambda_c) \quad (4.8)$$

$$\ddot{A} = 2a_2 . \quad (4.9)$$

Multiplying (1.28a) by A and using these series expansions,

$$2a_2 [a_1 (\lambda - \lambda_c) + a_2 (\lambda - \lambda_c)^2] = \frac{[a_1 + 2a_2 (\lambda - \lambda_c)]^2}{2} - 2\xi^2 - 2\mathcal{R} [a_1 (\lambda - \lambda_c) + a_2 (\lambda - \lambda_c)^2]^2 .$$

At the caustic, $\lambda \rightarrow \lambda_c$ so

$$\frac{a_1^2}{2} - 2\xi^2 = 0$$

and hence, $a_1 = \pm 2\xi$. This means that at this point,

$$A = 0 \quad (4.10)$$

$$\dot{A} = \pm 2\xi \quad (4.11)$$

$$\ddot{A} = 2a_2 , \quad (4.12)$$

where a_2 is undetermined since the OSE is satisfied for all values. Thus, A , \dot{A} , and \ddot{A} are well defined even at caustic points.

Considering now the quantities describing the shear, taking the quotient of (1.28b) and (1.28c),

$$\frac{\dot{\xi}}{\xi\dot{\varphi}} = \cot(\beta - \varphi) ,$$

which upon integration yields:

$$\xi \sin(\beta - \varphi) = \mathcal{C} , \quad (4.13)$$

where the constant $\mathcal{C} \equiv \xi_0 \sin(\beta - \varphi_0)$ is set by the values of ξ , φ , and β when a beam first enters a hole. Before encountering the first lens, a beam will be unsheared and so $\beta = \varphi_0$ (refer to Figure 1.4). After the initially circular beam starts shearing, it will in general have an elliptical shape and Dyer (1973) has shown that the minor axis of this ellipse has an orientation angle of $\varphi/2$.

While (1.28b) and (1.28c) can be used to compute ξ and φ respectively, it is possible to reduce these two equations to a single one. Equations (1.26) and (1.27) can be used to rewrite (1.17b) which, after some re-arranging, gives

$$[\xi e^{i(\varphi-\beta)}]^\cdot = A\mathcal{F} .$$

Considering the real part of this (as the imaginary part just gives back the definition of \mathcal{C}),

$$[\xi \cos(\beta - \varphi)]^\cdot = A\mathcal{F}$$

or, using (4.13) to replace the cosine term,

$$[\epsilon_\Delta \sqrt{\xi^2 - \mathcal{C}^2}]^\cdot = A\mathcal{F} ,$$

where $\epsilon_\Delta = \pm 1$ is determined by the sign of $\cos(\beta - \varphi)$. Defining a new variable

$$Y \equiv \epsilon_\Delta \sqrt{\xi^2 - \mathcal{C}^2} , \quad (4.14)$$

then

$$\frac{dY}{d\lambda} = \frac{3mh^2 A}{r^5} , \quad (4.15)$$

which can be integrated and used to find ξ and φ as necessary.

To compute the distortion of the beam, the axial ratio of its elliptical cross-section is required. In a Cartesian coordinate system (x, y) that is parallel propagated along with the shadow plane, the beam ellipse must satisfy the equation

$$A_e x^2 + 2H_e xy + B_e y^2 + C_e = 0, \quad (4.16)$$

where the coefficient of the rotation term is related to the others via

$$H_e^2 = A_e B_e + C_e.$$

The lengths of the semi-axes of the ellipse are given in terms of these coefficients by

$$a_e = \sqrt{\frac{1 + j^2}{B_e j^2 + 2H_e j + A_e}} \quad (4.17a)$$

$$b_e = \sqrt{\frac{1 + j^2}{A_e j^2 - 2H_e j + B_e}}, \quad (4.17b)$$

where

$$j = \frac{(B_e - A_e) + \sqrt{(B_e - A_e)^2 + 4H_e^2}}{2H_e}, \quad (4.18)$$

and the orientation of the ellipse with respect to the x -axis is simply

$$\alpha_e = \tan^{-1}(j) \quad (4.19)$$

(Harper, 1991). Note that $j = 0$ when $H_e = 0$. It should also be pointed out that α_e is in general different from the orientation of the shear rate, $\varphi/2$. The two are only the same in the case of a single lens encounter for an initially unsheared beam.

As a beam is propagated through the model, the change in the ellipse coefficients can be tracked via

$$\frac{d\bar{A}_e}{d\lambda} = -2\xi \left(\frac{\bar{A}_e}{A} \cos \varphi + \frac{\bar{H}_e}{A} \sin \varphi \right) + 2 \frac{\dot{A}\bar{A}_e}{A} \quad (4.20a)$$

$$\frac{d\bar{B}_e}{d\lambda} = 2\xi \left(\frac{\bar{B}_e}{A} \cos \varphi - \frac{\bar{H}_e}{A} \sin \varphi \right) + 2 \frac{\dot{A}\bar{B}_e}{A} \quad (4.20b)$$

$$\frac{d\bar{H}_e}{d\lambda} = -\xi \sin \varphi \frac{(\bar{A}_e + \bar{B}_e)}{A} + 2 \frac{\dot{A}\bar{H}_e}{A}, \quad (4.20c)$$

where the barred variables are equal to the equivalent ellipse coefficient multiplied by the square of the beam area (i.e. $\bar{A}_e \equiv A_e A^2$) in order to ensure the expressions are formally well behaved at caustics.¹ As $A \rightarrow 0$, H_e approaches a constant so $\frac{\bar{H}_e}{A} \rightarrow 0$. Similarly, the quantity $\frac{\bar{A}_e}{A}$ tends towards zero at the caustic if the ellipse is growing along the x -axis, and towards a constant if it is shrinking in that direction. The same is true for $\frac{\bar{B}_e}{A}$, but depending on the growth or contraction along the y -axis. Unfortunately, though the limits for the ratios in (4.20) are well defined at caustics in theory, attempting to numerically integrate them in practice can lead to results that are overly sensitive to the step size chosen by the integrator. In some situations, this can lead to invalid results when calculating the axial ratio of the ellipse after a caustic has been passed through. Attempts to track the semi-axes of the ellipse directly instead of through the ellipse coefficients also lead to expressions involving ratios with A in the denominator, and hence, which could potentially experience the same sensitivity.

4.2.2 The (C_{\pm}, α_{\pm}) Form

In the most simple case described in Section 1.3.2, where $\dot{\alpha}_+ = \dot{\alpha}_- = 0$, the principal curvatures C_+ and C_- are proportional to the radial and tangential axes of the beam respectively (Dyer, 1977). Hence, to within a constant, the area of the beam is just $A = C_+ C_-$, and the axial ratio can be found by taking the ratio of these curvatures. Furthermore, the differential equations (1.32) are clearly well-behaved for all real curvature values. However, this simple setup relies on the great deal of symmetry present in the encounter with a single, spherically-symmetric lens, and cannot be used if multiple lenses are encountered.

Moving to a more general case where α is allowed to vary, it is clear that while the expressions for \ddot{C}_{\pm} can easily be rewritten to be well-behaved everywhere, equation

¹As shown by Harper (1991) using a power series expansion of these quantities about the caustic. This is required because at least one of the coefficients of the ellipse will not be well defined when the beam passes through a caustic.

(1.31b) for $\ddot{\alpha}$ does have potential problems at a caustic where one of the curvatures goes to zero. Further, even if the expression can be made suitable for numerical integration, it is not clear how α can be related to the orientation of the beam ellipse, which is ultimately one of the quantities of interest.

4.2.3 The (ψ, η) Form

With this formulation, the semi-major and semi-minor axes of the elliptical cross-section of a distorted beam are given by

$$a_e = |\psi| + |\eta| \quad (4.21a)$$

$$b_e = |\psi| - |\eta|, \quad (4.21b)$$

while the orientation of this distortion is given by

$$\alpha_e = \frac{1}{2} \arg(\psi\eta) = \frac{1}{2} \tan^{-1} \left(\frac{xw + yz}{xz - yw} \right), \quad (4.22)$$

as shown by Pineault (1975). In this case though, the sign conventions used result in α_e being offset by $\frac{\pi}{2}$ from the value determined by Pineault, implying that it describes the orientation of the semi-minor axis instead of the semi-major one. With x , y , z , and w known by integrating equations (1.38), ψ and η can be computed everywhere. Thus, a_e and b_e can be determined, allowing the area and distortion of the beam to be computed as needed. Note that as was mentioned when discussing the (A, ξ, φ) formalism, it is also possible in this case for the cross-sectional area to be negative when $|\eta| > |\psi|$, with the beam turning “inside-out” in a sense. As was mentioned before, this simply means that an absolute value needs to be taken whenever a physical area is desired. More importantly, unlike the differential equations derived from the (A, ξ, φ) or (C_{\pm}, α) formalisms, all the relevant equations in this case are well behaved everywhere, including at caustics where the beam cross-section has zero area. Consequently, it is this form that is ultimately used by the lensing code to track the beam cross-section.

In order to initialize the system at the focal point of the observer, we use the fact that the beam has zero area there and choose it to start unsheared. Hence, $a_e = b_e = 0$, implying $\psi = \eta = 0$, and thus,

$$x = y = z = w = 0$$

to begin with. Considering the area of the beam and its derivatives at initialization,

$$A = \pi a_e b_e = 0 \tag{4.23}$$

$$\dot{A} = \pi (\dot{a}_e b_e + a_e \dot{b}_e) = 0 \tag{4.24}$$

$$\ddot{A} = 2\pi \dot{a}_e \dot{b}_e = 2\pi \dot{a}_e^2, \tag{4.25}$$

where the fact that $\dot{a}_e = \dot{b}_e$ for a non-shearing beam starting in an FRW region was used in (4.25). These expressions are in agreement with equations (4.10) to (4.12) using the (A, ξ, φ) formalism, which found a single free parameter related to the initial second derivative of the beam area. This formalism makes it clear that this free parameter is the initial rate of change of the radius of the circular beam cross-section. In order to translate this into the remaining initial conditions to impose on the variables being integrated, consider that $|\eta|$ must remain zero while there is no shear being experienced. Hence, we must set $\dot{z} = \dot{w} = 0$ initially, as along with equations (1.38c) and (1.38d), this ensures that z and w will remain zero in the first FRW region. The initial values of \dot{x} and \dot{y} can be set arbitrarily as they are related to the free parameter discussed above. However, since the beam must begin evolving purely radially in the shadow plane, the real and imaginary components of ψ must grow equally, so while they are arbitrary, the initial values of \dot{x} and \dot{y} must be equal.

4.3 Spinor Properties

Adopting a spinor formalism when dealing with null vector fields often makes it more convenient to work with them. A brief review of spinors and some of their properties is

given here as these will be needed in order to determine how to follow the orientation of a distorted beam through the model.

First-rank or 1-spinors are elements of a two-dimensional complex vector space, such as that required to describe the null 4-vectors associated with the beam (refer to Figure 4.1). A capital letter will be used for spinor indices, which can have the value of ① or ②, where the numbers are circled as is customary in order to distinguish them from values of tensor indices. The complex conjugate of a spinor will be noted by a bar over it and a dot over its index, so the complex conjugate of \mathcal{K}^A appears as $\bar{\mathcal{K}}^{\dot{A}}$. Like tensors, spinors can be written in covariant or contravariant notation and the indices can be raised or lowered via the metric spinor ϵ_{AB} . This is an antisymmetric object so

$$\epsilon_{AB} = -\epsilon_{BA} , \quad (4.26)$$

and the non-zero components are

$$\epsilon_{\textcircled{1}\textcircled{2}} = 1 \quad \text{and} \quad \epsilon_{\textcircled{2}\textcircled{1}} = -1 . \quad (4.27)$$

When contracting indices, the order matters so

$$\mathcal{K}^A \epsilon_{AB} = \mathcal{K}_B \quad \text{while} \quad \epsilon_{BA} \mathcal{K}^A = -\mathcal{K}_B . \quad (4.28)$$

In other words, when reading the expression from left to right, if the indices are being contracted upward, a negative sign is introduced, while contraction downward behaves as contraction of tensor indices does.

Taken with its inverse, the metric spinor can be used to form the two-dimensional Kronecker delta symbol,

$$\epsilon^{AC} \epsilon_{BC} = \delta_B^A . \quad (4.29)$$

Note that contraction of a spinor with the Kronecker delta does not introduce a negative

sign regardless of the direction of the contraction:

$$\begin{aligned}
\mathcal{K}_A \delta_B^A &= \mathcal{K}_A \epsilon^{AC} \epsilon_{BC} \\
&= -\mathcal{K}^C \epsilon_{BC} \\
&= \mathcal{K}^C \epsilon_{CB} \\
\therefore \mathcal{K}_A \delta_B^A &= \mathcal{K}_B
\end{aligned}$$

Another object that will prove to be useful is the Spinor Christoffel Symbol,

$$\Gamma_a{}^B{}_C = \frac{1}{2} \sigma_e{}^{B\dot{X}} \left[\sigma^b{}_{C\dot{X}} \Gamma_{ab}^e + \partial_a (\sigma^e{}_{C\dot{X}}) \right]$$

(Dyer, 1973), which can be re-arranged to a more usable form,

$$\Gamma_a{}^B{}_C = -\frac{1}{2} \sigma_e{}^{B\dot{X}} \left[g^{bf} \epsilon_{CP} \sigma_f{}^{P\dot{Q}} \epsilon_{\dot{Q}\dot{X}} \Gamma_{ab}^e + \partial_a \left(g^{eg} \epsilon_{CP} \sigma_g{}^{P\dot{Q}} \epsilon_{\dot{Q}\dot{X}} \right) \right]. \quad (4.30)$$

The $\sigma_a{}^{B\dot{C}}$ are *spin connections*, which are generalizations of the Pauli spin matrices. These can be obtained via a coordinate transformation appropriate to the coordinate system under consideration (Pirani, 1964).

4.4 A Parallel Propagated Shadow Plane Basis

4.4.1 The FRW Regions

As the quantities used to describe the distortion of a beam are defined in the shadow plane, a set of basis vectors for this plane that are parallel propagated along with the beam are needed in order to have a consistent means of following the quantities over multiple lens encounters. If one extends the definition of k^a , taking it from a null tangent vector to single a beam, to a real null vector field describing the tangent vectors to a congruence of beams, a 1-spinor, \mathcal{K}^A , can be defined to within an arbitrary phase via

$$\mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}} = \pm \sigma_a{}^{B\dot{X}} k^a,$$

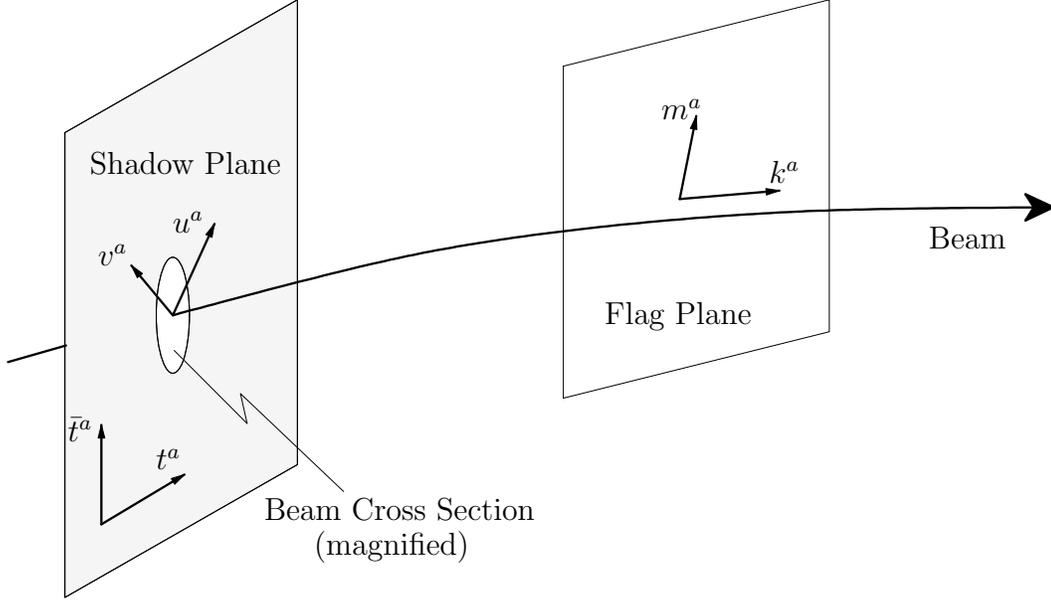


Figure 4.1: A depiction of the flag and shadow planes associated with a beam, along with the 4-vectors that span each of them.

or equivalently,

$$k^a = \pm \sigma^a_{B\dot{X}} \mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}} \quad (4.31)$$

(Pirani, 1964). The sign in these equations is determined by the time direction in which the null tangent vectors being considered are pointing. As future-pointing null tangent vectors are the ones of interest, the positive sign is appropriate so

$$\mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}} = \sigma_a^{B\dot{X}} k^a \quad (4.32)$$

will be used.

The spin connections in an FRW region can be written in matrix form as:

$$\begin{aligned} \sigma_0^{A\dot{B}} &= \sqrt{\frac{1}{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \sigma_1^{A\dot{B}} &= \frac{\epsilon_\omega R}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_2^{A\dot{B}} &= \frac{iRS}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \sigma_3^{A\dot{B}} &= \frac{RS \sin \theta}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \quad (4.33)$$

Here, the first spinor index specifies the row of the matrix and the second the column. As the coordinate system will always be oriented in such a way that the propagation of

a beam takes place in the $\theta = \pi/2$ plane, $k^2 = 0$. Since $\sigma_2^{A\dot{B}}$ is the only spin connection with an imaginary component, this means that $\mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}}$ must be real. Writing \mathcal{K}^B with explicit magnitude and phase terms,

$$\mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}} = M^B N^{\dot{X}} e^{i(\alpha_B - \alpha_{\dot{X}})},$$

so for $\mathcal{K}^B \bar{\mathcal{K}}^{\dot{X}}$ to be real, $\alpha_B - \alpha_{\dot{X}} = 0$, which means that $\mathcal{K}^{\textcircled{1}}$ and $\mathcal{K}^{\textcircled{2}}$ must have the same phase. This phase is arbitrary, so for convenience we can choose it in such a way that \mathcal{K}^A is both real and positive.

Introducing a second 1-spinor, \mathcal{U}^A , a basis for spinors can be established if we require

$$\mathcal{K}_A \mathcal{U}^A = 1 \tag{4.34}$$

(Dyer, 1973). Since \mathcal{K}^A is real, this implies that \mathcal{U}^A must also be real. This new spinor can be related to a new null 4-vector, m^a , through an analogous relation to (4.31),

$$m^a = \sigma^a_{B\dot{X}} \mathcal{U}^B \bar{\mathcal{U}}^{\dot{X}}. \tag{4.35}$$

The 4-vectors k^a and m^a form a basis for the *flag plane*, which is a plane that contains the beam as it propagates. The shadow plane is normal to the flag plane and has a basis described by a complex null 4-vector, t^a , which is defined via an expression that is similar to (4.31) and (4.35) but using both of the 1-spinors defined above:

$$t^a = \sigma^a_{B\dot{X}} \mathcal{K}^B \bar{\mathcal{U}}^{\dot{X}} \tag{4.36}$$

(Dyer, 1973). Refer to Figure 4.1 for a depiction of these quantities that make up a *null tetrad*.

It is useful to consider a few scalar products between these basis vectors as they can serve as a check for the components computed later on. Lowering the index of k^a in

(4.31) gives an expression for k_a using spinor quantities:

$$\begin{aligned}
g_{ab}k^b &= g_{ab}\sigma^b_{B\dot{X}}\mathcal{K}^B\bar{\mathcal{K}}^{\dot{X}} \\
k_a &= \sigma_{aB\dot{X}}\mathcal{K}^B\bar{\mathcal{K}}^{\dot{X}} \\
&= -\epsilon_{BA}\sigma_a^{A\dot{Y}}\epsilon_{\dot{Y}\dot{X}}\mathcal{K}^B\bar{\mathcal{K}}^{\dot{X}} \\
&= \epsilon_{AB}\mathcal{K}^B\sigma_a^{A\dot{Y}}\epsilon_{\dot{Y}\dot{X}}\bar{\mathcal{K}}^{\dot{X}} \\
\therefore k_a &= \mathcal{K}_A\sigma_a^{A\dot{Y}}\bar{\mathcal{K}}_{\dot{Y}}. \tag{4.37}
\end{aligned}$$

Using this with (4.35),

$$\begin{aligned}
k_a m^a &= \mathcal{K}_A\sigma_a^{A\dot{Y}}\bar{\mathcal{K}}_{\dot{Y}}\sigma^a_{B\dot{X}}\mathcal{U}^B\bar{\mathcal{U}}^{\dot{X}} \\
&= \mathcal{K}_A\delta_B^A\delta_{\dot{X}}^{\dot{Y}}\bar{\mathcal{K}}_{\dot{Y}}\mathcal{U}^B\bar{\mathcal{U}}^{\dot{X}} \\
&= \mathcal{K}_B\bar{\mathcal{K}}_{\dot{X}}\mathcal{U}^B\bar{\mathcal{U}}^{\dot{X}} \\
\therefore k_a m^a &= 1, \tag{4.38}
\end{aligned}$$

where the relation $\sigma_a^{A\dot{Y}}\sigma^a_{B\dot{X}} = \delta_B^A\delta_{\dot{X}}^{\dot{Y}}$ (Pirani, 1964) was used. Similarly, it can be determined that

$$k_a t^a = 0, \tag{4.39}$$

while

$$t_a t^a = 0 \quad \text{and} \quad \bar{t}_a t^a = -1. \tag{4.40}$$

In order to be able to follow the orientation of a beam numerically, the components of the 4-vectors listed above are needed in the coordinate system that is being used for the propagation. While the values of k^a in this region are given in (3.5), the form of \mathcal{U}^A is required to determine m^a and t^a . To find this spinor, it is instructive to first consider the form of \mathcal{K}^A . Using equation (4.32) and setting $B=\dot{X}=\textcircled{1}$,

$$\mathcal{K}^{\textcircled{1}}\bar{\mathcal{K}}^{\textcircled{1}} = \sigma_a^{\textcircled{1}\textcircled{1}}k^a.$$

However, $\mathcal{K}^A = \bar{\mathcal{K}}^{\dot{A}}$ and only $\sigma_0^{A\dot{B}}$ and $\sigma_3^{A\dot{B}}$ are symmetric in the spinor indices, so using

(3.5) and (4.33),

$$\begin{aligned}
(\mathcal{K}^{\textcircled{1}})^2 &= \sigma_0^{\textcircled{1}\dot{\textcircled{1}}} k^0 + \sigma_3^{\textcircled{1}\dot{\textcircled{1}}} k^3 \\
\mathcal{K}^{\textcircled{1}} &= \pm \sqrt{\sqrt{\frac{1}{2}} \left(\frac{p}{R}\right) + \left(\frac{RS}{\sqrt{2}}\right) \left(\frac{pq}{R^2 S^2}\right)} \\
\therefore \mathcal{K}^{\textcircled{1}} &= 2^{-\frac{1}{4}} \sqrt{\frac{p}{R}} \sqrt{1 + \frac{q}{S}}.
\end{aligned} \tag{4.41}$$

The positive sign was used for (4.41) due to the choice of phase for \mathcal{K}^A made previously.

In a similar way, setting $B=\dot{X}=\textcircled{2}$ in (4.32) provides the relation

$$\mathcal{K}^{\textcircled{2}} = 2^{-\frac{1}{4}} \sqrt{\frac{p}{R}} \sqrt{1 - \frac{q}{S}}. \tag{4.42}$$

For convenience, we can define two symbols, \boxplus and \boxminus , so that expressions (4.41) and (4.42) can be written as:

$$\mathcal{K}^{\textcircled{1}} = 2^{-\frac{1}{4}} \boxplus \quad \text{and} \quad \mathcal{K}^{\textcircled{2}} = 2^{-\frac{1}{4}} \boxminus, \tag{4.43}$$

which is more compact.

In order to drag the shadow plane along with the beam, we require that t^a undergo parallel propagation, so using (4.36),

$$\frac{D t^a}{D\lambda} = \sigma^a_{B\dot{X}} \left(\frac{D\mathcal{K}^B}{D\lambda} \bar{u}^{\dot{X}} + \mathcal{K}^B \frac{D\bar{u}^{\dot{X}}}{D\lambda} \right) = 0, \tag{4.44}$$

where $\frac{D}{D\lambda}$ is the covariant derivative along the affine parameter λ . Through (4.31), the parallel propagation of k^a implies the parallel propagation of \mathcal{K}^A , so for t^a to be parallel propagated, (4.44) implies that \mathcal{U}^A must also undergo parallel propagation. This requires

$$k^a \nabla_a \mathcal{U}^A = k^a (\partial_a \mathcal{U}^A + \Gamma_a^A_B \mathcal{U}^B) = 0 \tag{4.45}$$

(Dyer, 1973).

In an FRW region, the non-vanishing Christoffel Symbols are²:

$$\begin{aligned}
\Gamma_{11}^0 &= \dot{R}R & \Gamma_{22}^0 &= \dot{R}R\mathcal{S}^2 & \Gamma_{33}^0 &= \dot{R}R\mathcal{S}^2 \sin^2 \theta \\
\Gamma_{01}^1 &= \Gamma_{10}^1 = \frac{\dot{R}}{R} & \Gamma_{22}^1 &= -\mathcal{S}'\mathcal{S} & \Gamma_{33}^1 &= -\mathcal{S}'\mathcal{S} \sin^2 \theta \\
\Gamma_{02}^2 &= \Gamma_{20}^2 = \frac{\dot{R}}{R} & \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{\mathcal{S}'}{\mathcal{S}} & \Gamma_{33}^2 &= -\sin \theta \cos \theta \\
\Gamma_{03}^3 &= \Gamma_{30}^3 = \frac{\dot{R}}{R} & \Gamma_{13}^3 &= \Gamma_{31}^3 = \frac{\mathcal{S}'}{\mathcal{S}} & \Gamma_{23}^3 &= \Gamma_{32}^3 = \frac{\cos \theta}{\sin \theta}
\end{aligned} \tag{4.46}$$

where for this section, the dot and prime represent derivatives with respect to cT and ω respectively, so $\dot{R} \equiv \frac{dR}{dcT}$ and $\mathcal{S}' \equiv \frac{d\mathcal{S}}{d\omega}$. Using these values, along with the RW metric (1.7) and the spin connections (4.33) in (4.30), the spinor Christoffel Symbols in the FRW region can be computed to be:

$$\begin{aligned}
\Gamma_0^B{}_C &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Gamma_1^B{}_C &= \frac{\epsilon_\omega \dot{R}}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\Gamma_2^B{}_C &= i \begin{pmatrix} -\epsilon_\omega \mathcal{S}' & -\frac{1}{4} \dot{R}\mathcal{S} \\ \frac{1}{4} \dot{R}\mathcal{S} & \epsilon_\omega \mathcal{S}' \end{pmatrix} \\
\Gamma_3^B{}_C &= \frac{\sin \theta + 1}{4} \begin{pmatrix} \dot{R}\mathcal{S} \sin \theta & \epsilon_\omega \mathcal{S}' \sin \theta - i \cos \theta \\ -\epsilon_\omega \mathcal{S}' \sin \theta - i \cos \theta & -\dot{R}\mathcal{S} \sin \theta \end{pmatrix} \\
&= \frac{1}{2} \begin{pmatrix} \dot{R}\mathcal{S} & \epsilon_\omega \mathcal{S}' \\ -\epsilon_\omega \mathcal{S}' & -\dot{R}\mathcal{S} \end{pmatrix} \quad \text{for} \quad \theta = \frac{\pi}{2}
\end{aligned} \tag{4.47}$$

Note that again, the first spinor index of $\Gamma_a^B{}_C$ refers to the row of the matrix while the second refers to the column. Since \mathcal{K}^A is a function of $R(cT)$ and $\mathcal{S}_K(\omega)$, (4.34) implies that \mathcal{U}^A must also be a function of only cT and ω . Hence, these are the only two non-zero derivatives in (4.45), which simplifies to

$$k^0 \dot{\mathcal{U}}^A + k^1 \mathcal{U}^{A'} + k^a \Gamma_a^A{}_B \mathcal{U}^B = 0.$$

From (4.47), when $A=\textcircled{1}$, the only non-zero components of this expression are

$$k^0 \dot{\mathcal{U}}^{\textcircled{1}} + k^1 \mathcal{U}^{\textcircled{1}'} + k^3 \Gamma_3^{\textcircled{1}}{}_{\textcircled{1}} \mathcal{U}^{\textcircled{1}} + \left(k^1 \Gamma_1^{\textcircled{1}}{}_{\textcircled{2}} + k^3 \Gamma_3^{\textcircled{1}}{}_{\textcircled{2}} \right) \mathcal{U}^{\textcircled{2}} = 0, \tag{4.48}$$

²Found using the REDTEN package of the REDUCE computer algebra programming system.

and when $A=\textcircled{2}$,

$$k^0\dot{\mathcal{U}}^{\textcircled{2}} + k^1\mathcal{U}^{\textcircled{2}'} + k^3\Gamma_3^{\textcircled{2}}{}_{\textcircled{2}}\mathcal{U}^{\textcircled{2}} + \left(k^1\Gamma_1^{\textcircled{2}}{}_{\textcircled{1}} + k^3\Gamma_3^{\textcircled{2}}{}_{\textcircled{1}}\right)\mathcal{U}^{\textcircled{1}} = 0. \quad (4.49)$$

Substituting in the values for k^a and the components of the spinor Christoffel Symbols, (4.48) and (4.49) become

$$\frac{p}{R}\dot{\mathcal{U}}^{\textcircled{1}} + \frac{\epsilon_\omega p}{R^2}\sqrt{1 - \frac{q^2}{S^2}}\mathcal{U}^{\textcircled{1}'} + \frac{pq\dot{R}}{2R^2\mathcal{S}}\mathcal{U}^{\textcircled{1}} + \left(\frac{p\dot{R}}{2R^2}\sqrt{1 - \frac{q^2}{S^2}} + \frac{\epsilon_\omega pq\mathcal{S}'}{2R^2\mathcal{S}^2}\right)\mathcal{U}^{\textcircled{2}} = 0 \quad (4.50a)$$

and

$$\frac{p}{R}\dot{\mathcal{U}}^{\textcircled{2}} + \frac{\epsilon_\omega p}{R^2}\sqrt{1 - \frac{q^2}{S^2}}\mathcal{U}^{\textcircled{2}'} - \frac{pq\dot{R}}{2R^2\mathcal{S}}\mathcal{U}^{\textcircled{2}} + \left(\frac{p\dot{R}}{2R^2}\sqrt{1 - \frac{q^2}{S^2}} - \frac{\epsilon_\omega pq\mathcal{S}'}{2R^2\mathcal{S}^2}\right)\mathcal{U}^{\textcircled{1}} = 0 \quad (4.50b)$$

respectively. Noticing that $k^1 = \frac{\epsilon_\omega}{R}\boxplus\boxminus$, and using the relations

$$\begin{aligned} \boxplus^2 &= \frac{p}{R}\left(1 + \frac{q}{S}\right) & \boxminus^2 &= \frac{p}{R}\left(1 - \frac{q}{S}\right) \\ (\boxplus^2) &= -\frac{p\dot{R}}{R^2}\left(1 + \frac{q}{S}\right) & (\boxminus^2) &= -\frac{p\dot{R}}{R^2}\left(1 - \frac{q}{S}\right) \\ (\boxplus^2)' &= -\frac{pq\mathcal{S}'}{S^2} & (\boxminus^2)' &= \frac{pq\mathcal{S}'}{S^2} \end{aligned} \quad (4.51)$$

we can re-write (4.50a) and (4.50b) as

$$2p\dot{\mathcal{U}}^{\textcircled{1}} + 2\epsilon_\omega\boxplus\boxminus\mathcal{U}^{\textcircled{1}'} + \left[R(\boxplus^2) + \frac{p\dot{R}}{R}\right]\mathcal{U}^{\textcircled{1}} + \left[\dot{R}\boxplus\boxminus + \epsilon_\omega(\boxplus^2)'\right]\mathcal{U}^{\textcircled{2}} = 0 \quad (4.52a)$$

and

$$2p\dot{\mathcal{U}}^{\textcircled{2}} + 2\epsilon_\omega\boxplus\boxminus\mathcal{U}^{\textcircled{2}'} + \left[R(\boxminus^2) + \frac{p\dot{R}}{R}\right]\mathcal{U}^{\textcircled{2}} + \left[\dot{R}\boxplus\boxminus + \epsilon_\omega(\boxminus^2)'\right]\mathcal{U}^{\textcircled{1}} = 0. \quad (4.52b)$$

The symmetry of these equations indicates that the definitions of $\mathcal{U}^{\textcircled{1}}$ and $\mathcal{U}^{\textcircled{2}}$ have very similar forms, with one simply using \boxplus where the other uses \boxminus . This suggests solutions of the form

$$\mathcal{U}^{\textcircled{1}} = \mathcal{A}\boxplus + \mathcal{B}\boxminus \quad \text{and} \quad \mathcal{U}^{\textcircled{2}} = \mathcal{A}\boxminus + \mathcal{B}\boxplus \quad (4.53)$$

be tried, where \mathcal{A} and \mathcal{B} are potentially functions of $R(cT)$ and $\mathcal{S}(\omega)$. When used with the expressions for \mathcal{K}^A inside the spinor inner product (4.34), these relations yield an

expression for \mathcal{B} :

$$\begin{aligned}
\mathcal{K}^B \epsilon_{BA} \mathcal{U}^A &= 1 \\
\mathcal{K}^{\textcircled{1}} \mathcal{U}^{\textcircled{2}} - \mathcal{K}^{\textcircled{2}} \mathcal{U}^{\textcircled{1}} &= 1 \\
2^{-\frac{1}{4}} \boxplus (\mathcal{A} \boxplus + \mathcal{B} \boxplus) - 2^{-\frac{1}{4}} \boxminus (\mathcal{A} \boxplus + \mathcal{B} \boxplus) &= 1 \\
\therefore \mathcal{B} &= \frac{R\mathcal{S}}{2^{\frac{3}{4}}pq}. \tag{4.54}
\end{aligned}$$

Using (4.53) in (4.52a) and grouping the terms involving a common derivative,

$$\begin{aligned}
&\left[2p \left(\dot{\mathcal{A}} \boxplus + \mathcal{A} \dot{\boxplus} + \dot{\mathcal{B}} \boxplus + \mathcal{B} \dot{\boxplus} \right) + 2R\mathcal{A} \boxplus \boxminus \dot{\boxplus} + 2R\mathcal{B} \boxminus \boxplus \dot{\boxplus}^2 \right. \\
&\quad \left. + \frac{p\dot{R}\mathcal{A}}{R} \boxplus + \frac{p\dot{R}\mathcal{B}}{R} \boxplus + \dot{R}\mathcal{A} \boxplus \boxplus^2 + \dot{R}\mathcal{B} \boxplus \boxplus^2 \right] \\
&\quad + \left[2\epsilon_\omega \boxplus \boxplus \left(\mathcal{A}' \boxplus + \mathcal{A} \boxplus' + \mathcal{B}' \boxplus + \mathcal{B} \boxplus' \right) \right. \\
&\quad \left. + 2\epsilon_\omega \boxplus \boxplus' \left(\mathcal{A} \boxplus + \mathcal{B} \boxplus \right) \right] = 0. \tag{4.55}
\end{aligned}$$

Since it is unreasonable to expect terms containing an \dot{R} to cancel with terms containing an \mathcal{S}' , we can assume the dotted and primed terms in (4.55) go to zero separately. Symmetry arguments can again be invoked to suggest that the form of \mathcal{A} is similar to that found for \mathcal{B} . Setting the dotted part of (4.55) equal to zero, and using

$$\mathcal{A} = \mathcal{C} \frac{R\mathcal{S}}{pq} \tag{4.56}$$

along with (4.54) inside, produces the constraint

$$\frac{\dot{R}\mathcal{S}}{qR} \sqrt{\frac{p}{R}} \sqrt{1 + \frac{q}{\mathcal{S}}} \left(\mathcal{C} + 2^{-\frac{3}{4}} \sqrt{1 - \frac{q^2}{\mathcal{S}^2}} \right) = 0, \tag{4.57}$$

where \mathcal{C} is potentially a function of both R and \mathcal{S} . For this to be true in general,

$$\mathcal{C} = -2^{-\frac{3}{4}} \sqrt{1 - \frac{q^2}{\mathcal{S}^2}}, \tag{4.58}$$

so after some factoring,

$$\mathcal{A} = -\frac{R\sqrt{\mathcal{S}^2 - q^2}}{2^{\frac{3}{4}}pq}. \tag{4.59}$$

Direct substitution of this into the primed portion of (4.55) shows that it also becomes zero as needed. Similarly, substitution of (4.59) and (4.54) for \mathcal{A} and \mathcal{B} respectively into (4.52b) shows that it is also satisfied, which means that these values can be used to determine the components of a parallel propagated \mathcal{U}^A as required. Substituting them into (4.53) yields the expressions:

$$\mathcal{U}^{\textcircled{1}} = -2^{-\frac{3}{4}}\boxminus = -2^{-\frac{3}{4}}\sqrt{\frac{R}{p}}\sqrt{1 - \frac{q}{\mathcal{S}}} \quad (4.60a)$$

$$\mathcal{U}^{\textcircled{2}} = 2^{-\frac{3}{4}}\boxplus = 2^{-\frac{3}{4}}\sqrt{\frac{R}{p}}\sqrt{1 + \frac{q}{\mathcal{S}}}, \quad (4.60b)$$

which only differ from $\mathcal{K}^{\textcircled{2}}$ and $\mathcal{K}^{\textcircled{1}}$ by constant factors.

Knowing \mathcal{U}^A , the 4-vector m^a can now be computed. Expanding (4.35) to a more convenient form,

$$m^a = -g^{ab}\epsilon_{BA}\sigma_b^{A\dot{Y}}\epsilon_{\dot{Y}\dot{X}}\mathcal{U}^B\bar{\mathcal{U}}^{\dot{X}}, \quad (4.61)$$

and hence,

$$m^a = \left(\frac{R}{2p}, -\frac{\epsilon_\omega}{2p}\sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}}, 0, -\frac{q}{2p\mathcal{S}_K^2(\omega)} \right). \quad (4.62)$$

Using (3.13) to take the inner product of this with k^a ,

$$k_a m^a = \frac{1}{2} + \frac{1}{2} \left(1 - \frac{q^2}{\mathcal{S}^2} \right) + \frac{q^2}{2\mathcal{S}^2} = 1, \quad (4.63)$$

in agreement with (4.38). Similarly, writing (4.36) as

$$t^a = -g^{ab}\epsilon_{BA}\sigma_b^{A\dot{Y}}\epsilon_{\dot{Y}\dot{X}}\mathcal{K}^B\bar{\mathcal{U}}^{\dot{X}}, \quad (4.64)$$

the components of t^a are

$$t^a = \left(0, \frac{\epsilon_\omega q}{\sqrt{2}R\mathcal{S}_K(\omega)}, \frac{i}{\sqrt{2}R\mathcal{S}_K(\omega)}, -\frac{1}{\sqrt{2}R\mathcal{S}_K(\omega)}\sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}} \right). \quad (4.65)$$

As before, using (3.13) to take the inner product of this with k^a ,

$$k_a t^a = \frac{\epsilon_\omega q}{\sqrt{2}RS} \left(-\epsilon_\omega p \sqrt{1 - \frac{q^2}{\mathcal{S}^2}} \right) - \frac{1}{\sqrt{2}RS} \sqrt{1 - \frac{q^2}{\mathcal{S}^2}} (-pq) = 0, \quad (4.66)$$

in agreement with (4.39), while taking the inner product of t^a with itself and its complex conjugate can easily be seen to satisfy relations (4.40) as required.

Though t^a and \bar{t}^a span the shadow plane, the fact that they are null makes them inconvenient to use directly as a basis. However, t^a can be written in terms of a real and imaginary component, u^a and v^a respectively:

$$t^a = \frac{1}{\sqrt{2}} (u^a + iv^a) , \quad (4.67)$$

where the factor of $\frac{1}{\sqrt{2}}$ is for normalization. Taking the complex conjugate, we see that

$$u^a = \frac{1}{\sqrt{2}} (t^a + \bar{t}^a) \quad \text{and} \quad v^a = \frac{i}{\sqrt{2}} (\bar{t}^a - t^a) \quad (4.68)$$

or, expanding them with (4.65),

$$u^a = \left(0 , \frac{\epsilon_\omega q}{R\mathcal{S}_K(\omega)} , 0 , -\frac{1}{R\mathcal{S}_K(\omega)} \sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}} \right) \quad (4.69)$$

$$v^a = \left(0 , 0 , \frac{1}{R\mathcal{S}_K(\omega)} , 0 \right) . \quad (4.70)$$

It is easy to confirm that u^a and v^a are spacelike and normal to one another, with

$$u^a u_a = v^a v_a = -1 \quad \text{and} \quad u^a v_a = 0 , \quad (4.71)$$

so these two vectors, which are related to the orientation of the beam and also lie in the shadow plane, can be used to establish a more practical basis there.

With the orientation of the coordinate system chosen, at any point in the propagation of the beam through an FRW region, the normal to the lens of interest is given by

$$n^a = (0, \epsilon_n, 0, 0) ,$$

which can be either inward or outward pointing. In order to establish the angle β (refer to Figure 1.4), this vector needs to be projected into the shadow plane, an operation that can be accomplished through the use of the projection tensor h_{ab} ,

$$h_{ab} = g_{ab} - k_a m_b - k_b m_a . \quad (4.72)$$

Using the nullity of k^a and m^a , along with relations (4.38) to (4.40), it is easy to see that h_{ab} does what is expected:

$$\begin{aligned} h_{ab}k^a &= 0 & h_{ab}t^a &= t_b \\ h_{ab}m^a &= 0 & h_{ab}\bar{t}^a &= \bar{t}_b \end{aligned}$$

acting like a metric tensor in the shadow plane. Using it to find the projected normal vector,

$$\tilde{n}_a = h_{ab}n^b = \epsilon_n g_{a1} + \epsilon_n \epsilon_\omega \sqrt{1 - \frac{q^2}{\mathcal{S}^2}} \left(pm_a - \frac{R^2}{2p} k_a \right),$$

so

$$\tilde{n}_a = \left(0, -\frac{\epsilon_n R^2 q^2}{\mathcal{S}_K^2(\omega)}, 0, \epsilon_n \epsilon_\omega q R^2 \sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}} \right). \quad (4.73)$$

Writing the contravariant components of this,

$$\tilde{n}^a = \left(0, \frac{\epsilon_n q^2}{\mathcal{S}_K^2(\omega)}, 0, -\frac{\epsilon_n \epsilon_\omega q}{\mathcal{S}_K^2(\omega)} \sqrt{1 - \frac{q^2}{\mathcal{S}_K^2(\omega)}} \right), \quad (4.74)$$

and comparing it to the expression for u^a in (4.69), we see that

$$u^a = \frac{\epsilon_\omega \mathcal{S}}{\epsilon_n q R} \tilde{n}^a.$$

Hence, in the FRW regime, the vector u^a always lies in the same direction as the projection of the normal to the lens onto the shadow plane. In fact, u^a is the normalized value of \tilde{n}^a . It is straightforward to check that $\tilde{n}^a k_a = \tilde{n}^a m_a = 0$ as is required.

4.4.2 The Schwarzschild Regions

Here we follow exactly the same procedure as was done for the FRW regime, but derive the quantities needed to follow the orientation of a beam in the vacuum regions of the model. The spin connections in the Schwarzschild region outside an event horizon can be written as:

$$\begin{aligned} \sigma_0^{A\dot{B}} &= \sqrt{\frac{f}{2}} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \sigma_1^{A\dot{B}} &= \frac{\epsilon_\gamma}{\sqrt{2f}} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ \sigma_2^{A\dot{B}} &= \frac{ir}{\sqrt{2}} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \sigma_3^{A\dot{B}} &= \frac{r \sin \theta}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \quad (4.75)$$

where f is given in (1.3). As the coordinate system will again be oriented in such a way that beam propagation takes place in the $\theta = \pi/2$ plane and $\sigma_2^{A\dot{B}}$ is the only spin connection with an imaginary component, we can once more use k^a to define a spinor \mathcal{K}^A that is both real and positive. Setting $B=\dot{X}=\dot{\ominus}$ in (4.32), and using the spin connections above with the components of k^a given in (3.36),

$$\begin{aligned}\mathcal{K}^{\dot{\ominus}}\bar{\mathcal{K}}^{\dot{\ominus}} &= \sigma_a^{\dot{\ominus}\dot{\ominus}}k^a \\ (\mathcal{K}^{\dot{\ominus}})^2 &= \sigma_0^{\dot{\ominus}\dot{\ominus}}k^0 + \sigma_3^{\dot{\ominus}\dot{\ominus}}k^3 \\ \mathcal{K}^{\dot{\ominus}} &= \pm\sqrt{\frac{l_\gamma}{f}\sqrt{\frac{f}{2}} + \frac{h_\gamma}{r^2}}\left(\frac{r}{\sqrt{2}}\right) \\ \mathcal{K}^{\dot{\ominus}} &= 2^{-\frac{1}{4}}\sqrt{\frac{l_\gamma}{\sqrt{f}} + \frac{h_\gamma}{r}},\end{aligned}\tag{4.76}$$

where the positive sign was used as before. Doing the same for $B=\dot{X}=\dot{\otimes}$ yields the other component,

$$\mathcal{K}^{\dot{\otimes}} = 2^{-\frac{1}{4}}\sqrt{\frac{l_\gamma}{\sqrt{f}} - \frac{h_\gamma}{r}},\tag{4.77}$$

and introducing the symbols \oplus and \ominus , these can be simplified to a more compact form:

$$\mathcal{K}^{\dot{\oplus}} = 2^{-\frac{1}{4}}\oplus \quad \text{and} \quad \mathcal{K}^{\dot{\ominus}} = 2^{-\frac{1}{4}}\ominus.\tag{4.78}$$

To find the components of the second 1-spinor \mathcal{U}^A , we again need to consider what is required for it to undergo parallel propagation along a beam trajectory.

The non-vanishing Christoffel Symbols in the external Schwarzschild solution are:³

$$\begin{aligned}\Gamma_{01}^0 &= \Gamma_{10}^0 = \frac{f'}{2f} \\ \Gamma_{00}^1 &= \frac{f'f}{2} & \Gamma_{11}^1 &= -\frac{f'}{2f} & \Gamma_{22}^1 &= -fr & \Gamma_{33}^1 &= -fr\sin^2\theta \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{1}{r} & \Gamma_{33}^2 &= -\sin\theta\cos\theta \\ \Gamma_{13}^3 &= \Gamma_{31}^3 = \frac{1}{r} & \Gamma_{23}^3 &= \Gamma_{32}^3 = \frac{\cos\theta}{\sin\theta}\end{aligned}\tag{4.79}$$

³Again found using REDUCE's REDTEN package.

where the prime represents a derivative with respect to the r coordinate in this section. Using these values and the metric (1.2), along with the spin connections (4.75), the spinor Christoffel Symbols (4.30) can be computed to be:

$$\begin{aligned}
\Gamma_0^B{}_C &= \frac{\epsilon_\gamma f'}{4} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
\Gamma_1^B{}_C &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \\
\Gamma_2^B{}_C &= \frac{i\epsilon_\gamma \sqrt{f}}{2} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \\
\Gamma_3^B{}_C &= \frac{1}{2} \begin{pmatrix} 0 & \epsilon_\gamma \sqrt{f} \sin \theta - i \cos \theta \\ -\epsilon_\gamma \sqrt{f} \sin \theta - i \cos \theta & 0 \end{pmatrix} \\
&= \frac{\epsilon_\gamma \sqrt{f}}{2} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{for} \quad \theta = \frac{\pi}{2}
\end{aligned} \tag{4.80}$$

in the Schwarzschild vacuum regions. Unlike the FRW regime, \mathcal{K}^A is only a function of $f(r)$ here, which means that the same is true for \mathcal{U}^A , and so (4.45) can be written as

$$k^1 \mathcal{U}^{A'} + k^a \Gamma_a^A{}_B \mathcal{U}^B = 0. \tag{4.81}$$

Setting $A=\oplus$ and expanding this using known quantities,

$$k^1 \mathcal{U}^{\oplus'} + \frac{\epsilon_\gamma \sqrt{f}}{2} \left(\frac{l_\gamma f'}{2f\sqrt{f}} + \frac{h_\gamma}{r^2} \right) \mathcal{U}^\ominus = 0. \tag{4.82}$$

Using the relations

$$\begin{aligned}
\oplus^2 &= \frac{l_\gamma}{\sqrt{f}} + \frac{h_\gamma}{r} & \ominus^2 &= \frac{l_\gamma}{\sqrt{f}} - \frac{h_\gamma}{r} \\
(\oplus^2)' &= - \left(\frac{l_\gamma f'}{2f\sqrt{f}} + \frac{h_\gamma}{r^2} \right) & (\ominus^2)' &= - \left(\frac{l_\gamma f'}{2f\sqrt{f}} - \frac{h_\gamma}{r^2} \right)
\end{aligned} \tag{4.83}$$

along with the expression $k^1 = \epsilon_\gamma \sqrt{f} \oplus \ominus$, (4.82) can be simplified to

$$\ominus \mathcal{U}^{\oplus'} - \oplus' \mathcal{U}^\ominus = 0. \tag{4.84a}$$

Similarly, setting $A=\ominus$ in (4.81) and following through gives

$$\oplus \mathcal{U}^{\ominus'} - \ominus' \mathcal{U}^\oplus = 0. \tag{4.84b}$$

As was the case in the FRW regions, the symmetry of (4.84a) and (4.84b) suggests that $\mathcal{U}^{\textcircled{2}}$ can be obtained by switching \oplus and \ominus in the definition of $\mathcal{U}^{\textcircled{1}}$, so solutions of the form

$$\mathcal{U}^{\textcircled{1}} = \mathcal{A} \oplus + \mathcal{B} \ominus \quad (4.85a)$$

$$\mathcal{U}^{\textcircled{2}} = \mathcal{A} \ominus + \mathcal{B} \oplus \quad (4.85b)$$

can be tried, where this time \mathcal{A} and \mathcal{B} are potentially functions of $f(r)$. Using these inside the inner product (4.34) again yields an expression for \mathcal{B} :

$$2^{-\frac{1}{4}} \oplus (\mathcal{A} \ominus + \mathcal{B} \oplus) - 2^{-\frac{1}{4}} \ominus (\mathcal{A} \oplus + \mathcal{B} \ominus) = 1$$

$$\therefore \mathcal{B} = \frac{r}{2^{\frac{3}{4}} h_\gamma} . \quad (4.86)$$

Substituting (4.85) into (4.84),

$$\mathcal{A}' \oplus \ominus + \mathcal{B}' \ominus^2 + \mathcal{B} (\ominus \ominus' - \oplus \oplus') = 0 \quad (4.87a)$$

and

$$\mathcal{A}' \oplus \ominus + \mathcal{B}' \oplus^2 + \mathcal{B} (\oplus \oplus' - \ominus \ominus') = 0 . \quad (4.87b)$$

Solving for \mathcal{A}' in (4.87a), and using the relations

$$\oplus' = -\frac{1}{2\oplus} \left(\frac{l_\gamma f'}{2f\sqrt{f}} + \frac{h_\gamma}{r^2} \right) \quad \text{and} \quad \ominus' = -\frac{1}{2\ominus} \left(\frac{l_\gamma f'}{2f\sqrt{f}} - \frac{h_\gamma}{r^2} \right) , \quad (4.88)$$

results in the expression

$$\mathcal{A}' = -\frac{l_\gamma r}{2^{\frac{3}{4}} h_\gamma \sqrt{l_\gamma^2 r^2 - h_\gamma^2 f}} . \quad (4.89)$$

Direct substitution shows that this also satisfies (4.87b). Due to the way in which f depends on r , this is in general an elliptic integral which does not have a simple analytic solution. However, (3.36b) can be used to convert this to a derivative with respect to λ instead of r_γ ,

$$\frac{d\mathcal{A}}{d\lambda} = \frac{d\mathcal{A}}{dr_\gamma} \frac{dr_\gamma}{d\lambda} = -\frac{\epsilon_\gamma l_\gamma}{2^{\frac{3}{4}} h_\gamma} , \quad (4.90)$$

which can easily be solved to give

$$\mathcal{A}_{\text{Final}} = \mathcal{A}_{\text{Initial}} + \frac{\epsilon_\gamma l_\gamma}{2^{\frac{3}{4}} h_\gamma} (\lambda_{\text{Initial}} - \lambda_{\text{Final}}) . \quad (4.91)$$

Consequently, as in the FRW case, it will be possible to compute the values of \mathcal{U}^A at any point during the propagation of a beam.

Knowing the spinors \mathcal{K}^A and \mathcal{U}^A , (4.61) and (4.64) can be evaluated to give m^a and t^a respectively:

$$\begin{aligned} m^0 &= \frac{1}{\sqrt{2f}} [(\mathcal{A}^2 + \mathcal{B}^2) (\oplus^2 + \ominus^2) + 4\mathcal{A}\mathcal{B} \oplus \ominus] \\ m^1 &= \epsilon_\gamma \sqrt{2f} [(\mathcal{A}^2 + \mathcal{B}^2) \oplus \ominus + \mathcal{A}\mathcal{B} (\oplus^2 + \ominus^2)] \\ m^2 &= 0 \\ m^3 &= \frac{1}{r\sqrt{2}} (\mathcal{A}^2 - \mathcal{B}^2) (\oplus^2 - \ominus^2) \end{aligned} \quad (4.92)$$

and

$$\begin{aligned} t^0 &= \frac{1}{2^{\frac{3}{4}} \sqrt{f}} [\mathcal{A} (\oplus^2 + \ominus^2) + 2\mathcal{B} \oplus \ominus] & t^2 &= \frac{i}{\sqrt{2} r} \\ t^1 &= \frac{\epsilon_\gamma \sqrt{f}}{2^{\frac{3}{4}}} [2\mathcal{A} \oplus \ominus + \mathcal{B} (\oplus^2 + \ominus^2)] & t^3 &= \frac{2^{\frac{1}{4}} h_\gamma \mathcal{A}}{r^2} \end{aligned} \quad (4.93)$$

in the Schwarzschild exterior. It is straightforward to show that these expressions satisfy the relations (4.38) to (4.40) as required. Knowing t^a , (4.68) can be used to find the components of the real spacelike vectors

$$\begin{aligned} u^0 &= \frac{1}{\sqrt{f}} \left[\frac{r}{h_\gamma} \oplus \ominus + \frac{\mathcal{A}}{2^{\frac{1}{4}}} (\oplus^2 + \ominus^2) \right] & u^2 &= 0 \\ u^1 &= \epsilon_\gamma \sqrt{f} \left[\frac{r}{2h_\gamma} (\oplus^2 + \ominus^2) + 2^{\frac{3}{4}} \mathcal{A} \oplus \ominus \right] & u^3 &= \frac{2^{\frac{3}{4}} h_\gamma \mathcal{A}}{r^2} \end{aligned} \quad (4.94)$$

and

$$v^a = \left(0, 0, \frac{1}{r}, 0 \right) \quad (4.95)$$

which span the shadow plane in the vacuum regions of the model.

The normal to the lens in this regime is also given by $n^a = (0, \epsilon_n, 0, 0)$, so constructing the projection tensor h_{ab} as in (4.72), the orientation of the lens in the shadow plane can

be written in contravariant form:

$$\begin{aligned}
\tilde{n}^0 &= \frac{\epsilon_\gamma \epsilon_n 2\sqrt{2}}{f} \left[\frac{l_\gamma}{\sqrt{f}} (\mathcal{A}^2 + \mathcal{B}^2) \oplus \ominus + \mathcal{A} \mathcal{B} \left(\frac{2l_\gamma^2}{f} - \frac{h^2}{r^2} \right) \right] \\
\tilde{n}^1 &= \epsilon_n \left[2\sqrt{2} (\mathcal{A}^2 + \mathcal{B}^2) \oplus^2 \ominus^2 + \frac{4\sqrt{2} l_\gamma}{\sqrt{f}} \mathcal{A} \mathcal{B} \oplus \ominus + 1 \right] \\
\tilde{n}^2 &= 0 \\
\tilde{n}^3 &= \frac{\epsilon_\gamma \epsilon_n 2\sqrt{2} h_\gamma \mathcal{A}}{r^2} \left(\frac{\mathcal{A}}{\sqrt{f}} \oplus \ominus + \frac{l_\gamma \mathcal{B}}{f} \right).
\end{aligned} \tag{4.96}$$

Comparing this to the components of u^a in (4.94), it is possible to relate the two via

$$u^a = \frac{\epsilon_\gamma \epsilon_n f h_\gamma l_\gamma r}{2^{-\frac{1}{4}} \sqrt{l_\gamma^2 r^2 - f h_\gamma^2} + f h_\gamma^2} \tilde{n}^a$$

provided that

$$\mathcal{A} = \frac{1 - 2^{\frac{1}{4}} \sqrt{l_\gamma^2 r^2 - f h_\gamma^2}}{2 h_\gamma l_\gamma}.$$

However, this does not satisfy (4.89) which is required for the parallel propagation of \mathcal{U}^A . Hence, unlike the relationship found in the FRW regions, the shadow-plane vector u^a in the Schwarzschild vacuum regions does not in general track the orientation of a lens as a beam propagates past it.

4.5 Tracking the Beam Distortion

4.5.1 Initializing the Orientation

As described in Section 1.3.2, a beam propagated through the RSC model is initially unsheared and has a distortion imposed on it by the first lens that it encounters. This lens was described as having an orientation of $\beta/2$ as measured in the shadow plane (refer to Figure 1.4). In order to allow comparisons to be made between beams, this angle can be measured with respect to some fiducial direction chosen by the observer. Refer to this direction as j^a , a 4-vector expressed in the unoriented RW coordinate system describing the uppermost level of the model (the background). While j^a is arbitrary, it should not be parallel to any of the beams that are going to be propagated through the model.

In order to compute $\beta/2$, j^a must first be projected into the shadow plane, which is equivalent to making it a direction “on the sky” at the location of the observer, so that the inner product can be taken with the projected normal vector of the first lens that the beam will encounter. In the previous sections dealing with spinors, a parallel propagated basis for the shadow plane was established using the vectors u^a and v^a . As demonstrated in Section 4.4.1, in the FRW regions, the projected normal vector is parallel to u^a , so using \tilde{j}^a to represent the projection of j^a ,

$$\tilde{j}_a u^a = h_{ab} j^b u^a = (g_{ab} - k_a m_b - k_b m_a) j^b u^a, \quad (4.97)$$

where (4.72) was used to expand the projection tensor h_{ab} . Recall that the null tetrad spanning the flag and shadow planes was defined using the propagation orientation described in Section 3.1. Hence, before projecting j^a , the transformation needed to prepare the system for propagation to the first lens must also be applied to it.

Since u^a is normal to k^a and m^a , then from (4.97), $\tilde{j}_a u^a = j_a u^a$ so

$$\sqrt{|\tilde{j}_b \tilde{j}^b| |u_c u^c|} \cos\left(\frac{\beta}{2}\right) = j_a u^a.$$

Re-arranging and using the fact that $u^a u_a = -1$ yields

$$\left(\frac{\beta}{2}\right)_{\text{orig}} = \cos^{-1}\left(\frac{j_a u^a}{\sqrt{|j_b j^b - 2j_b k^b j_c m^c|}}\right). \quad (4.98)$$

Repeating this procedure, the inner product $\tilde{j}^a v_a$ can be used to find a similar expression for α_v , the angle between these two vectors. If $\alpha_v < \pi/2$, then \tilde{j}^a lies in the half of the uv plane containing the vector v^a . For a right-handed uv system, this can be interpreted as u^a being rotated by an angle of $\beta/2$ in a clockwise direction from \tilde{j}^a . This situation can be taken to imply a negative rotation angle, while $\alpha_v > \pi/2$ will imply counter-clockwise rotation and hence, a positive value of β .

4.5.2 Following the Orientation Over Multiple Encounters

In order to properly account for the effect of multiple lens encounters on a single beam, care is needed to ensure that orientations in the shadow plane are computed in a con-

sistent manner. All angles are measured with respect to the projected fiducial vector \tilde{j}^a , but this vector is not parallel propagated with the beam, so can only be used at the location of the observer. For this reason, \tilde{j}^a was expressed in the uv coordinate system in the previous section since these vectors are parallel propagated. Their form has been computed for both the FRW and Schwarzschild regions, so all that remains is to establish the connection between these quantities at the interface between the two coordinate systems. As long as the change in orientation of u^a and v^a can be computed between coordinate systems, any direction expressed using these vectors can be related back to the orientation of \tilde{j}^a in the appropriate manner.

Once β has been computed using (4.98), the beam can be propagated from the observer to the boundary of the first hole it encounters, where the system can be converted to use Schwarzschild coordinates as described in Section 3.3.1. The value of β is still valid in the vacuum coordinate system, and must also remain unchanged for the vacuum propagation. This is because the beam must propagate in a plane containing the centre of the lens, so when viewed from the shadow plane, the lens centre can only move towards or away from the origin in a radial manner (refer to Figure 1.4). The vectors u^a and v^a that were parallel propagated with the beam need to be converted to the new system in order to allow \mathcal{A} to be defined via (4.94). As described in Appendix C.1.2, the conversion of an arbitrary vector from the RW coordinate system to Schwarzschild coordinates is quite straightforward, but the conversion back needed at the end of the vacuum propagation, is not as easy. Fortunately, this procedure, along with the need to compute \mathcal{A} at all, can be avoided with a judicious choice of where the uv vectors are evaluated.

If the beam propagates past the core and reaches the boundary to the background again, then while still in the propagation orientation, the beam location and direction can be converted back to RW coordinates, and the basis vectors⁴ $u_{(1)}^a$ and $v_{(1)}^a$ computed

⁴The bracketed subscripts are simply labels.

in this system. Next, the transformations that were performed to place the system in propagation orientation can be inverted, so all quantities are expressed in terms of the unoriented background coordinates. The procedure described in Section 3.1 can then be followed to again set the system up for propagation to the next lens (or through the space to z_{\max} if no other lenses are hit), and new values of the basis vectors, $u_{(2)}^a$ and $v_{(2)}^a$, can be computed in this orientation. The transformation required to establish this new propagation orientation can then be applied to $u_{(1)}^a$ and $v_{(1)}^a$ in order to allow them to be compared to $u_{(2)}^a$ and $v_{(2)}^a$. Recall that $u_{(1)}^a$ is oriented at an angle of $\beta/2$ as defined with respect to the direction standard \tilde{j}^a , and $u_{(2)}^a$ is parallel to the normal to the next lens to be propagated to, defining a new but unknown value of β . Hence, the angle between $u_{(1)}^a$ and $u_{(2)}^a$ is the change in $\beta/2$ required to have the orientation of the next lens measured with respect to \tilde{j}^a .

If the beam does hit the core, a similar procedure can be followed to that described for the boundary. At the edge of the core, the location and direction of the beam can be converted back to RW coordinates, and u^a and v^a computed in the vacuum propagation orientation. The rotations m_5 , m_6 , and m_7 , described in Section 3.1 to ensure the beam propagates in the $\theta = \pi/2$ plane can then be inverted in order to express the system in “unoriented core coordinates”. As discussed in Section 3.3.5, the other transformations which were required to put the lens at the origin of the coordinate system do not have to be inverted as they depend only on the coordinates of the lens centre and can thus be used to define the unoriented core coordinate system.

Propagation through the core can be treated just as the propagation through the FRW background was, with the system being re-oriented as necessary and the change in β being computed as described above. As the core propagation orientation is also used to follow the beam back through the vacuum away from the core, the shadow plane vectors need to be recomputed using it once the location and direction of the beam are expressed using RW coordinates at the outer boundary of the hole. After the system

is converted back to the unoriented background coordinates, then to the propagation orientation for the next lens, new values for the shadow plane vectors can be computed. The transformations can be applied to the old vectors to express them in the new system, and the difference in the orientation of the two u^a vectors is once again the change in $\beta/2$ required.

Checking Coordinate System Handedness

While the inner product of the old and new u^a values described above will allow the magnitude of the change in β to be computed from lens to lens, it does not indicate whether this change should make β grow larger or become smaller. As mentioned when explaining the orientation of the uv system with respect to \tilde{j}^a , this involves determining whether the new u^a value lies in the upper half of the old uv system (i.e. the half containing the old v^a) or the lower one. Though it may seem that a cross product with the old value of v^a is all that is required to resolve this issue, this is not the case due to the potential handedness change of the coordinate system that can occur as a result of the transformations applied to it. Instead, the inner product can be used to compute the angle made by each of the new vectors with each of the old ones, and these four angles employed to determine which quadrant of the old system the new vectors lie in. Then, subtracting the quadrant of the new v^a from that of the new u^a will indicate whether the handedness of the two systems is the same (the quadrant difference is equal to -1 or 3) or different (corresponding to a difference of 1 or -3). Refer to Figure 4.2 for an example. Note that care is needed to properly deal with the cases where the new vectors are parallel (or anti-parallel) to the old ones. Propagation can continue as usual whenever the handedness of the coordinate system does flip, but the meaning of the new u^a lying in the upper half of the old coordinate system changes. If the situation previously implied an increase in β , it implies a decrease after the flip, while what used to indicate a decrease before the flip implies an increase afterwards.

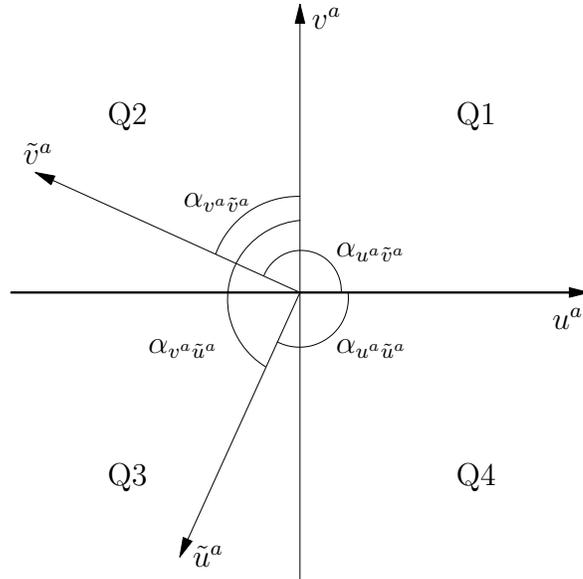


Figure 4.2: Comparing the handedness of the old (u^a, v^a) and new $(\tilde{u}^a, \tilde{v}^a)$ shadow plane vectors. With $\alpha_{u^a \tilde{u}^a} > \pi/2$ and $\alpha_{v^a \tilde{v}^a} > \pi/2$, \tilde{u}^a is in quadrant 3, and for $\alpha_{u^a \tilde{v}^a} > \pi/2$ and $\alpha_{v^a \tilde{u}^a} < \pi/2$, \tilde{v}^a is in quadrant 2. Hence, $3 - 2 = 1$ implies the coordinate systems have a different handedness as is clear from the figure.

Since changing the orientation of a lens by π has no impact on the shear, the restriction $|\beta/2| \leq \pi/2$ can be imposed. If, after a coordinate system change, β is found to have exceeded this bound, it can be reset via

$$\frac{\beta}{2} = \begin{cases} \frac{\beta}{2} - \pi & : \quad \frac{\pi}{2} < \frac{\beta}{2} \leq \frac{3\pi}{2} \\ \frac{\beta}{2} + \pi & : \quad -\frac{\pi}{2} > \frac{\beta}{2} \geq -\frac{3\pi}{2} \end{cases}, \quad (4.99)$$

then rotating the uv axes by π (multiplying the components by -1) to keep relative orientations consistent.

Chapter 5

Computational Considerations

5.1 The Packing Code

The code used to generate the packing libraries summarized in Table 2.1 was written in C. The time required for it to generate a packing varies dramatically depending on the parameters it is given, with the smallest runs listed in the table taking on the order of an hour on a P4 1-2 GHz or equivalent machine, while the larger ones involving regions with $\omega_{\max} = 2\pi$ took upward of two months to finish.

Before running the packing code itself, a utility program (also written in C) needs to be used to generate a file containing the locations of random points that properly populate the space to be packed (refer to Appendix D for details). These points are selected at random by the packing algorithm to specify the initial location of each packed sphere. The task of choosing the random locations was purposely separated from the packing process in order to allow the random distributions to be analyzed before being accepted. As the code uses a random point, or finds it to be invalid on account of it lying inside a sphere that has already been packed, it is marked by being given a negative ω coordinate (an impossible value). Once a packing has been completed, the random point file can then easily be reset before being used to generate subsequent packings.

Aside from the number of random points available to it in the data file described above, and the value of ω_{\max} for the space under consideration, the code also requires the following values be specified in a parameter file:

MAX_INIT and **MIN_INIT** : The range of radii that can be used when initially seeding the space with random spheres.

MIN_RADIUS : If set to be non-zero, will limit the range of packed sphere sizes a priori by specifying the minimum radius that they can have to be accepted.

MIN_SEARCH : If set to be non-zero, specifies the threshold value by which a trial sphere needs to change its radius while being packed in order for the optimization routine to continue working with it.

Q_THRESHOLD : The volume fraction of the space that is to be contained inside the packed spheres (i.e. the PVF).

STEP : A stepping factor which effectively determines the strength of the “force” used to move the spheres around in the effort to optimally fill a particular region of space.

While the code maintains a list of packed spheres in memory as it runs, it also periodically dumps this data to disk and updates a checkpoint record in order to allow runs to be stopped and restarted as necessary. Each line of the output file contains four numbers that describe a packed sphere, with three specifying the coordinates of its centre and the fourth indicating its radius. Once the packing is complete, another utility is available to validate it, determining the range and distribution of the packed sphere radii and ensuring that there is no overlap between them.

5.2 Using the Lensing Code

The code that propagates beams of light through an RSC universe that it constructs is also written in C, though is designed to be used in conjunction with pre and post-

processing scripts written in `Python`. It is function-based and contains on the order of ten thousand lines which are well commented in order to aid in the process of adapting it as necessary. While its runtime will obviously depend on the number of beams propagated, the time required to follow a single beam can also vary significantly based on the number of lenses potentially available (i.e. the number of spheres in the packings being used). For instance, a 1.6 GHz P4 took a fraction of a second to propagate a single beam past a single lens when only one was available, but about 10 seconds on average to do the same in a model where there were 1400 potential lenses in the background and 85 000 at each sub level.

The code requires a couple of parameter files to be passed to it when run, one containing the physical information and data that is used in the construction of the RSC model, and the other describing the observer and beam(s) to propagate. The specific contents of these files is expanded upon in Section 5.2.1 below. As the model is built, the code stores details about it in a file that is separate from the lensing output discussed in Section 5.2.2. The contents of this model file are virtually identical to the fields listed when explaining how the model is stored in memory at the end of Section 3.3.4. In this case though, rather than the pointers that are used to connect each node (lens) to the rest of the model in memory, a field of integers contains this information instead. The number of integers corresponds to the number of levels of recursion permitted in the model, with the value of each integer indicating the position of the node at the particular recursion level represented by the integer. For a model with six levels of recursion permitted, some example entries may appear as follows:

Field Entry	Description
0 0 0 0 0 0	reserved for definition of background universe
1 0 0 0 0 0	first lens defined in the background
2 3 0 0 0 0	third lens in the core of the second lens in the background

With this information, the file contains all that is needed for the lensing program to

reconstruct the model as required, and it can be passed to the program in subsequent runs to have it re-use and continue to expand on the model instead of generating a new one.

5.2.1 Initialization

Cosmological Quantities

To set up the FRW background in which the RSC model is embedded, there are four physical parameters that are specified:

- the matter density, $\Omega_m = \frac{8\pi G\rho_0}{3H_0^2} = 2\sigma_0$
- the cosmological constant density, $\Omega_\Lambda = \frac{\Lambda c^2}{3H_0^2}$
- the Hubble parameter, H_0
- the physical extent of the region under consideration, r_{universe}

Note that the deceleration parameter is given by $q_0 = 0.5\Omega_m - \Omega_\Lambda$. The curvature of the background space is then computed from the sum $\Omega_m + \Omega_\Lambda$. If it is equal to 1, the space is flat, less than one and the space is open, greater than one and the space is closed. The values of ρ_0 and Λ are computed from Ω_m and Ω_Λ respectively.

To find the initial value of the scale factor, multiply the Friedmann equation (1.12) through by $\frac{c^2}{3}$,

$$\frac{1}{R^2} \left(\frac{dR}{dT} \right)^2 = \frac{8\pi G\rho}{3} + \frac{\Lambda c^2}{3} - \frac{Kc^2}{R^2} .$$

Evaluating this at the time of observation and rearranging,

$$\frac{Kc^2}{R_0^2} = H_0^2 (\Omega_M + \Omega_\Lambda - 1) ,$$

so

$$R_0 = \sqrt{\frac{Kc^2}{H_0^2 (\Omega_M + \Omega_\Lambda - 1)}} \tag{5.1}$$

for $\Omega_M + \Omega_\Lambda \neq 1$. This implies that choosing values for H_0 , Ω_M , and Ω_Λ sets the value of R_0 that must be used for non-flat FRW backgrounds. In the situation where $K = 0$, the value of R_0 is not constrained in the same way so to choose it, the code uses the relation

$$R_0 \mathcal{S}_K(\omega_{\max}) = r_{\text{universe}} , \quad (5.2)$$

where in this flat case, $\mathcal{S}_K(\omega_{\max}) = \omega_{\max}$ is a unit-less radial measure of the extent of the background region being considered. It is randomly chosen to be equal to that of one of the packings available, and the value of R_0 is then set appropriately. When $K \neq 0$, the value of ω_{\max} is not arbitrary, but is computed using the value of R_0 from (5.1) inside (5.2). Note that the further $\Omega_M + \Omega_\Lambda$ is from 1 (or the larger the value of H_0 chosen), the larger the value of ω_{\max} will be.

Observer Description

The geodesic equation (3.1) is actually four second order differential equations for the four components of x^a , so a geodesic is uniquely determined locally by specifying a direction at a particular location. Accordingly, to start a beam off, it is necessary to specify the position 4-vector of the final observer, $x^a = (cT, \omega, \theta, \phi)$, along with the direction in which they are looking. This is given by the 4-vector $k^a = \left(\frac{dcT}{d\lambda}, \frac{d\omega}{d\lambda}, \frac{d\theta}{d\lambda}, \frac{d\phi}{d\lambda} \right)$ that is tangent to the null geodesic along which the light beam will travel. These eight components are arbitrary quantities to the degree that the observer can be placed at any location within the FRW background of the model and can look in any spatial direction desired. Without loss of generality, cT can be chosen to be zero at the beginning of the beam propagation but the spatial components x^μ must be specified in one of the parameter files. The same is true for the spatial components k^μ , with the program setting the temporal component k^0 by imposing the restriction that k^a be null. More than one direction vector can be specified, in which case the code propagates a beam along each one in sequence. These directions can either be set individually, or computed by specifying the number of beams to place in a grid with a given angular spacing about a central direction. Finally, the

spatial components of the fiducial direction vector j^a referred to in Section 4.5.1 must also be specified, with the temporal component being arbitrarily set to zero to ensure the vector is spacelike.

Limiting Values and Other Quantities

As explained in Section 2.2.1, in order to be able to truncate the RSC model, the code needs to be passed a minimum lens mass and radius that it can use. In order to know when to stop beam propagation, it needs to be given a value for z_{\max} , the redshift of the beam source. As explained in Section 4.2, the initial radial rate of change of the circular beam cross-section (which is related to the the value of \ddot{A}) is arbitrary, and thus, needs to be specified. The actual value chosen is not physically important as it scales out of the results, though it should be set to something that keeps the beam area numerically reasonable. Finally, the code must be told about the packing libraries available for it to use in generating the RSC model, so the parameter file needs to contain information on the number of these libraries available for each curvature and angular size.

5.2.2 Processing the Output

For each beam it propagates through the model, the code produces information regarding the initial direction of propagation, lenses encountered, caustics passed through, and the ultimate fate of the beam if its propagation does not end at a valid source (see Section 3.4). It also outputs the values of the beam area, distortion, and orientation at a number of different redshifts in order to allow the evolution of the beam to be followed. The specific redshift values used are determined by the step sizes taken by the integrator (refer to Section 5.3.1), but the code generates at least a minimal level of output in order to make it easier to plot the redshift evolution of the various quantities. In particular, it tries to ensure that these quantities are stored at least once for each unit interval in z for $z \leq 10$, every 10 units of z for $10 < z < 100$, and every hundred units for z up to 1000.

The output file produced can then be processed by a script which summarizes the results and provides for the interactive generation of data tables describing the various “observables” in order to allow graphs of these quantities to be generated. The redshift evolution of the distortion experienced by each beam, along with the evolution of the orientation of this distortion, and the amplification experienced by the beam can all be plotted. Furthermore, if an ensemble of beams have been propagated, the final (i.e. source) values of each of these quantities can be displayed at once in a “sky view” (refer to Section 6.2 for examples).

The distortion and orientation are generated directly by the lensing code as has been described earlier (refer to Section 4.2), while the amplification must be determined using the beam area that has been tracked. There are two types of amplification that can be computed. The first is the *geometric amplification*

$$\mathcal{A} = \frac{A_h}{A}, \quad (5.3)$$

where A is the area of the beam as computed by the lensing code, and A_h is the homogeneous beam area, which is what the beam would have if it were to propagate to the same redshift through a uniform FRW universe. This area is related to the angular size distance $D_<$ via

$$A = D_<^2 \delta\omega,$$

where $\delta\omega$ is the solid angle covered by the beam. When $\Lambda = 0$ and hence, $q_0 = \sigma_0$, $D_<$ can be computed via Mattig’s relation:

$$D_< = \frac{c}{H_0} \left(\frac{1 - q_0(1 - z) - (1 - q_0)\sqrt{1 + 2q_0z}}{q_0^2(1 + z)^2} \right) \quad (5.4)$$

(see, for instance, [Weinberg, 1972](#)).

Note that the amplification is formally infinite when the beam being followed through the RSC model goes through a caustic that the comparison beam in the homogeneous case does not experience (i.e. when $A \rightarrow 0$ while $A_h \neq 0$). The output processing code

can handle such cases by setting \mathcal{A} to be a finite number, but one significantly larger than the surrounding values to indicate the strong amplification being experienced. This is reasonable since a real beam that is not governed by the geometric optics approximation will experience interference effects that will serve to leave the amplification large, but finite (Schneider et al., 1992).

It is also possible to compute an amplification by comparing the area of a beam propagated through an RSC model with that of a beam travelling an equivalent distance through an empty universe. The relevant angular size distance can be found in this situation by integrating the Dyer-Roeder relation (Dyer and Roeder, 1973) once the intergalactic mass fraction, α , has been set to zero:

$$D_{<} = \frac{c}{H_0} \int_0^z \frac{dz'}{(1+z')^3 \sqrt{1+2q_0 z'}}.$$

Hence, when $q_0 < 0.5$,

$$D_{<} = \frac{3cq_0^2}{H_0(1-2q_0)^{\frac{5}{2}}} \left[\sinh^{-1} \sqrt{\frac{1-2q_0}{2q_0}} - \sinh^{-1} \sqrt{\frac{1-2q_0}{2q_0(1+z)}} \right] \\ - \frac{3cq_0}{2H_0(1-2q_0)^2} \left[1 - \frac{\sqrt{1+2q_0 z}}{1+z} \right] + \frac{c}{2H_0(1-2q_0)} \left[1 - \frac{\sqrt{1+2q_0 z}}{(1+z)^2} \right],$$

when $q_0 = 0.5$,

$$D_{<} = \frac{2c}{5H_0} \left[1 - \frac{1}{(1+z)^{\frac{5}{2}}} \right],$$

and when $q_0 > 0.5$,

$$D_{<} = \frac{3cq_0^2}{H_0(2q_0-1)^{\frac{5}{2}}} \left[\sin^{-1} \sqrt{\frac{2q_0-1}{2q_0}} - \sin^{-1} \sqrt{\frac{2q_0-1}{2q_0(1+z)}} \right] \\ - \frac{3cq_0}{2H_0(2q_0-1)^2} \left[1 - \frac{\sqrt{1+2q_0 z}}{1+z} \right] - \frac{c}{2H_0(2q_0-1)} \left[1 - \frac{\sqrt{1+2q_0 z}}{(1+z)^2} \right],$$

In fact, Mattig's relation (5.4) can be found in a similar way by integrating the Dyer-Roeder relation after setting $\alpha = 1$ to describe a universe with all matter distributed in the homogeneous background.

5.3 The Integrator

In order to follow the various quantities needed to propagate a beam through the RSC model, a fourth-order Runge-Kutta integration scheme with adaptive step sizes was chosen. As with any Runge-Kutta scheme, the routine provides results that are equivalent to those which would be found by expanding the relevant system of differential equations in a Taylor series of the same order. Instead of requiring higher order derivatives to be computed at each step though, this method simply evaluates the given functions at multiple points on the interval to be stepped over. A weighted sum of these evaluations is then used to advance the system over the interval.

The adaptive step size is based on the Fehlberg technique, where a judicious choice of weights and evaluation points allow both a fourth and fifth order Runge-Kutta step to be computed with only six evaluations of each ODE (which is the minimum required for any fifth order method). The routine advances the system using the fifth order calculation, but compares the results to those obtained with the fourth order method. If the two sets differ by more than some tolerance, the step is discarded and the process repeated with a smaller step size. On the other hand, if the results agree to a significantly greater degree than indicated by the tolerance, the incremented values are kept, but the step size is increased for the subsequent step. In this way, fewer steps are taken when the functions are slowly varying, while many small steps are used when the system is changing quickly. Note that while the fifth order calculation is used to advance the system, the method is still only considered to be valid to fourth order since it is to this order that the error is controlled at each step.

The routine implemented is essentially a standard Runge-Kutta-Fehlberg scheme, but with the Fehlberg parameters replaced by the Cash-Karp parameters (refer, for instance, to [Press et al., 1992](#), for more details). While the two schemes produce identical results most of the time, the one using the Cash-Karp parameters seems to be better able to handle integration close to critical points, such as when a beam nears the closest approach

to a core while propagating through a vacuum. It also seems to converge to a solution a little quicker than the other method does.

5.3.1 Setting Step Sizes

While the integration routine refines the step size it uses in order to remain within a prescribed error tolerance, it must still be told how far to step a system. Consider first, the FRW regions. As described earlier, the target value of the radial ω coordinate can be determined analytically. In order to compute the increment in the affine parameter λ that will be required to advance the system to this point, the expression for $\frac{d\omega}{d\lambda}$ in (3.5b) can be rearranged to give:

$$d\lambda = \frac{\epsilon_\omega R^2}{p \sqrt{1 - \frac{q^2}{S_K^2(\omega)}}} d\omega . \quad (5.5)$$

Since it is undesirable for the integrator to exceed the target value of ω (for instance, if propagating towards a point of closest approach, stepping to a value of ω that is too small will result in k^1 being imaginary), the code attempts to choose an underestimate for $d\lambda$ by using the larger of the initial and target values of ω in (5.5). However, the change in R over the integration interval can still lead to the step size being overestimated. Consequently, if the code finds that a quantity has been stepped beyond its reasonable limit, the integration is repeated with a scaling factor of the form 2^{-n} applied to the step estimated. The integer n is initially one, but increases until the system can be advanced successfully. If further steps are required in the region afterwards, n is reduced by one following each successful step until it reaches zero. This scaling procedure is also employed whenever the step sizes need to be reduced in order to have the code output results at some desired redshift value. When a beam has almost reached the boundary of a region, the attempts to systematically underestimate the step size can lead to a situation where the integration routine is asked to take a step smaller than what it considers reasonable for numerical precision. In such a case, a simple linear step is then employed to advance

all quantities by the small amount remaining. This typically changes the values by little more than roundoff errors and thus has little impact on the results.

Setting step sizes in the Schwarzschild regions is a little trickier, but can be accomplished in a similar fashion. The complication with the vacuum propagation is that the target radius for a beam is not necessarily known in advance. If the beam is entering a hole, the target may be the point of closest approach r_{ca} , which can be computed in advance, but it may also be the radius of the core which changes with time. To be safe, whenever a step size needs to be computed for an ingoing beam, the code chooses the larger of these two possible targets, or the event horizon if they lie within it. The estimated step interval is computed by using (3.68) to express equation (3.36b), describing the radial motion of the beam, in terms of the affine parameter ν , and isolating $d\nu$ to get:

$$d\nu = \frac{r_\gamma}{r_\gamma - 2m} \left(\frac{\epsilon_\gamma l_\gamma r_\gamma \sqrt{r_\gamma}}{\sqrt{l_\gamma^2 r_\gamma^3 - h_\gamma^2 (r_\gamma - 2m)}} + 1 \right) dr . \quad (5.6)$$

Whereas the interval $d\lambda$ was estimated by simply subtracting the target value of ω from the starting value in order to calculate $d\omega$, doing the same for dr does not provide acceptable results for $d\nu$. Since (5.6) cannot be analytically solved for general h_γ , it is integrated numerically using a 10-point Gauss-Legendre quadrature scheme. If the beam is outgoing, then using the radius of the boundary as the target will always lead to an overestimate in the step size since the boundary of a hole that does not turn around moves inward (for a beam propagated backwards). An attempt is made to compensate for this by giving the Gaussian integrator a target radius that is half way between r_γ and r_b in such a situation. As was the case in the FRW regions, the step size is further scaled using a scheme employing inverse powers of two if any of the values being numerically integrated are found to have overstepped their physical bounds.

While it appears that (5.6) will become problematic at some values of r , this is not the case. There is a potential problem for the value of r_γ that causes the root in the

denominator to go to zero, but this only happens at the point of closest approach, and as explained in Section 5.3.2 below, this never needs to be evaluated by the Runge-Kutta integrator.¹ The combination of $r_\gamma = 2m$ with $\epsilon_\gamma = 1$ is not a problem since the integrator will not need to follow a beam out from the event horizon, while an application of L'Hopital's rule shows that the expression remains finite when $\epsilon_\gamma = -1$.

Again, as was the case in the FRW regions, once the beam has come to within some small tolerance of the radius of the core or boundary, effectively having reached them, a linear step is taken to advance all the quantities the small amount to the radius required.

5.3.2 Passing Closest Approach

In situations where the beam needs to propagate past the origin of the RW coordinate system, the integration is performed in two stages; from the beam location to the point of closest approach, then from the point of closest approach onward. In the first stage, the integration stops when the ω coordinate of the beam, ω_{f_1} , agrees with the computed distance of closest approach to within some tolerance (currently 10^{-12}). Before the beam can continue away from the origin, ϵ_ω is changed to $+1$ and the value of ϕ may need to be adjusted due to the difference between ω_{f_1} and d_{ca} (refer to Figure 5.1). At one extreme, if the beam were to pass right through the origin, then ϕ would need to be changed by π . On the other hand, if $\omega_{f_1} = d_{ca}$, then no change needs to be made to ϕ . In general though, without this adjustment the beam would just propagate back in the direction from which it came.

To find $\Delta\phi$, it is assumed that the path of of the beam is straight over the region of interest² so

$$\Delta\phi = 2 \cos^{-1} \left(\frac{d_{ca}}{\omega_{f_1}} \right) .$$

¹Recall also that a Gaussian integrator using roots of Legendre polynomials as quadrature nodes does not require evaluation at the endpoints of the integration, so r_{ca} is still a valid limit to give this integrator.

²This makes the two triangles depicted in Figure 5.1 regular, right-angled triangles.

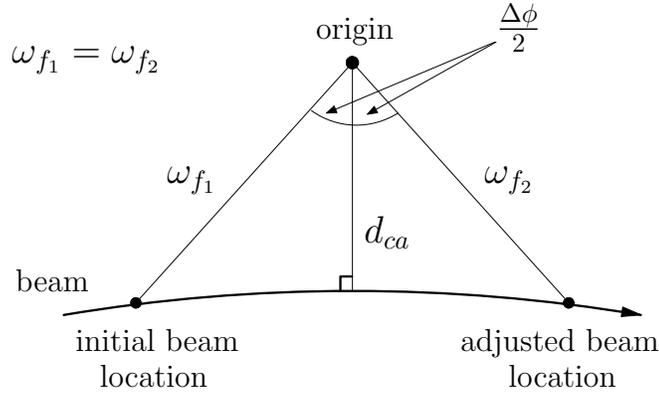


Figure 5.1: The change that must be made to the ϕ coordinate of a beam when propagating past the centre of coordinates in an FRW region.

Of course, this relationship is strictly only true for a Euclidian geometry, but given the small differences between d_{ca} and ω_{f_1} involved, the corrections for curved spaces are too small to be numerically significant.

If the point of closest approach in a vacuum region lies outside the core, then as is the case in an FRW region, the propagation has to be broken up into two parts since the sign ϵ_r changes at this point, and the code cannot otherwise easily choose when to make the change itself. Unfortunately, when trying to compute the initial change in the radial location of the beam in the second part of the integration, it is possible for the Runge-Kutta routine to try a value of r_γ smaller than r_{ca} in (3.36b), leading to a negative quantity under the square root. Since the integrator has no problems advancing any of the other quantities at this point, the problem is resolved by employing a hybrid integration routine that uses a fifth order Taylor series to advance r_γ over the critical region, while concurrently using the Runge-Kutta integrator to continue following the other quantities.

When the beam gets close enough to r_{ca} (currently set to be within the empirically determined distance of $1.0002 r_{ca}$), the standard Runge-Kutta integration is stopped and the hybrid routine is invoked. First, $\Delta\nu$, the affine parameter interval to the closest approach, is computed from (5.6) using a 100-point Gauss-Legendre integration scheme.

Next, the coefficients of a fifth order Taylor expansion of this equation are computed and used to determine r_γ at each of the Runge-Kutta “approximation” steps. These values can then be employed as appropriate to compute the derivatives required to advance the other variables by the “real” Runge-Kutta step. The coefficients are finally used to advance r_γ by the corresponding amount before being recomputed at the beginning of the next real step. This procedure is repeated until all quantities have been advanced by twice $\Delta\nu$, that is to the point of closest approach and an equal affine parameter interval away from it. At this stage, the standard Runge-Kutta integrator can be used again without running into the same problems with r_γ due to the beam being too close to the critical point.

5.4 Detecting Caustics

As a beam is propagated through the model universe, it is useful for the code to store the location of any caustics that are experienced. The least likely way for a caustic to be detected is to land directly on it. That is, the integration routine advances the system a step and finds the beam area to be exactly zero at the end. If this does occur, the beam location and optical scalar values can be dumped, and the propagation continued.

Far more likely is the situation where the beam steps over a caustic in the process of being propagated. There are two ways to detect this has happened, depending on whether a point or a line caustic has been crossed. For a beam that has started to shear, it will most probably be a line caustic, where only one dimension of the cross-section goes to zero. When this happens, there is a change in sign for the area of the beam, but not the derivative of the area. For instance, if $A > 0$ and $\dot{A} < 0$ initially, after passing through a line caustic, A will become negative while \dot{A} will not change sign since the area of the beam becomes more negative as the cross-section grows again. On the other hand, if a point caustic is encountered, both dimensions of the beam cross-section go to

zero. In this case, there is no change of sign for the area³, but the sign of \dot{A} does change. A beam with an initially positive area will still have this be true after the caustic, but would have gone from having a decreasing area to having it increasing. Hence, monitoring changes in the sign of both A and \dot{A} allows the code to determine when a caustic has been encountered, and what type of caustic it is.

When using the (ψ, η) formalism, it may seem more reasonable to track the value of the semi-minor axis b_e directly rather than its product with a_e , since this axis must go to zero when a caustic is encountered. However, recall that the values actually integrated are (x, y, z, w) and their derivatives $(\dot{x}, \dot{y}, \dot{z}, \dot{w})$.⁴ When written in terms of these quantities and expressed in units of π , the area of the beam and its derivative have the form

$$A = x^2 + y^2 - z^2 - w^2 \quad (5.7)$$

$$\dot{A} = 2(x\dot{x} + y\dot{y} - z\dot{z} - w\dot{w}) \quad (5.8)$$

while the semi-minor axis and its derivative are given by

$$b_e = \sqrt{x^2 + y^2} - \sqrt{z^2 + w^2} \quad (5.9)$$

$$\dot{b}_e = \frac{x\dot{x} + y\dot{y}}{\sqrt{x^2 + y^2}} - \frac{z\dot{z} + w\dot{w}}{\sqrt{z^2 + w^2}}. \quad (5.10)$$

The latter equations are more complicated to compute, and since the check for a caustic must be made each time the integrator advances a step, it is more effective to monitor A and \dot{A} .

Once the code detects that a caustic has been stepped over, it then uses the change in b_e to estimate where the caustic actually occurred. As the increment by which the integrator advances the beam is small, it is sufficient to use linear interpolation within the step for this task. Figure 5.2 shows the beam in profile as it passes through a caustic.

³In a sense, the change resulting from one dimension passing through zero is undone by the other dimension doing the same thing.

⁴A Runge-Kutta integration scheme requires higher order differential equations to be broken down into a system of coupled first order differential equations, so the second order equations for (x, y, z, w) are written as first order equations for $(\dot{x}, \dot{y}, \dot{z}, \dot{w})$ coupled with four trivial first order equations of the form $\frac{dx}{d\lambda} = \dot{x}$.

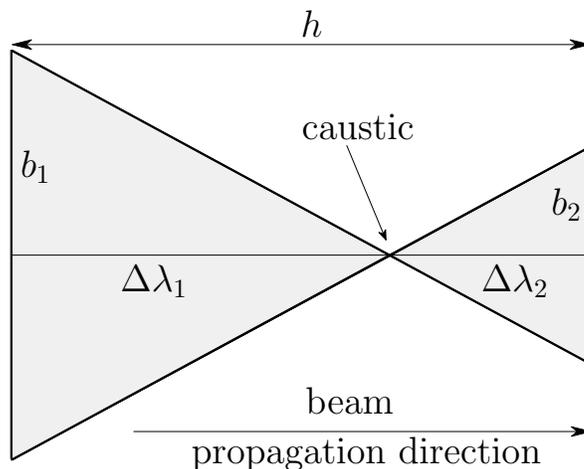


Figure 5.2: A side view of a beam cross-section propagating through a caustic.

Using simple trigonometry,

$$\Delta\lambda_1 = \frac{b_1 h}{b_1 + b_2},$$

which can be used to determine the value of λ at the caustic and subsequently, the quantities of interest can be computed at that point.

5.5 Consistency Checks

There are two general categories of checks applied to the code in order to validate the results it produces. The first category involves those that are performed in order to keep numerical errors under control, and do not necessarily have anything to do with the physics being investigated. They consist mainly of the internal monitoring that the code carries out while running, to ensure that various values remain within acceptable limits, but also include checks that are made manually through the use of debug statements. Some examples of these checks are as follows:

- Checks are made to prevent over/underflows in operations, particularly when ratios are taken.

- Checks are made to use alternate forms of some equations that could otherwise cause problems in certain limits.
- There are some cases where redundant expressions are used to provide a check on results that have been generated numerically. For instance, steps are taken to correct the tangent vector to the beam if it is found to be deviating significantly from being null.
- When critical values are computed, attempts are made to control the numerical errors so that later computations involving these quantities do not have problems with them being unphysical
- Checks are made to monitor the consistency of the coordinate rotations that are performed.
- Taylor stepping past r_{ca} was checked and found to be working properly, with the fifth order term typically being at least 8 orders of magnitude smaller than the first.
- The path that a beam takes through a packing which has not had its spheres replaced by higher density cores reveals that the correct ones are hit in the appropriate order. Furthermore, even though each sphere is used in turn to orient the propagation, the path taken is consistent with the case where the spheres are not present and the centre of coordinates of the space is the only orientation reference.
- Regardless of which form of the optical scalar equations is used (refer to Section 4.2), the variation in the beam area as it is propagated from the focus at the observer to the first caustic encountered is the same.⁵
- The ratio of beam areas between any two given redshifts remains constant regardless of the initial value chosen for the radial rate of change of the beam cross-section,

⁵The comparison is possible over this interval since it is the one where none of the forms are sensitive to the step size used by the integrator.

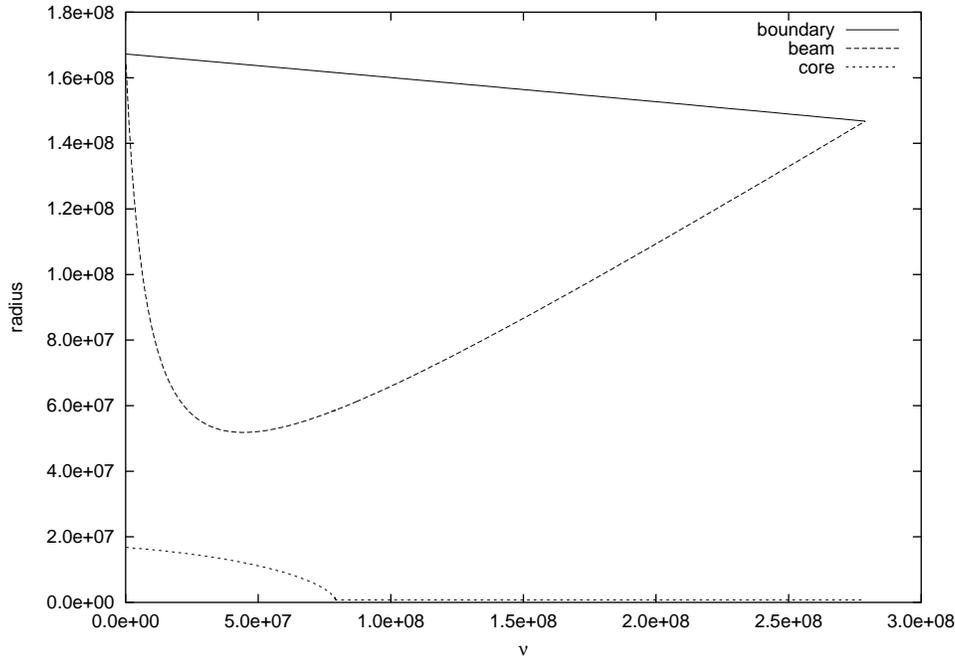


Figure 5.3: The variation with affine parameter of the radius of a hole, that of its core, and the radial position of a beam propagating through the vacuum. Note that in this case, the core has entered its event horizon before the beam has left the hole.

demonstrating that the results are not sensitive to this arbitrary choice.

- Analyzing the evolution of a beam using the (A, ξ, φ) form of the OSE, it can be seen that when $A = 0$ at a caustic, $\dot{A} = \pm 2\xi$ in agreement with equation (4.11). Further, the minimum value of \dot{A} ,

$$\dot{A} = \pm 2\sqrt{\xi^2 + \frac{4\pi G\rho_0 R_0^3 p^2 A^2}{c^2 R^5}},$$

is indeed found numerically when $\ddot{A} = 0$, as is to be expected from equation (1.28a).

- Plotting the evolution of beams, boundary observers, and core observers through a vacuum region show them to have reasonable trajectories as can be seen from the sample shown in Figure 5.3. Note that the non-symmetric nature of the graph is due to the timing “signals” that establish ν being radially infalling photons.
- When given an old model with the same initial conditions used for the beam that “discovered” it, the code propagates a new beam along the same path, verifying

that the re-use of a previously stored model works as it should.

The second category of checks involve analyzing the output produced by the code in order to make sure that the results generated for various test cases agrees with what is physically expected. These checks verify that the code does indeed generate meaningful results, and are discussed in Chapter 6.

Chapter 6

Sample Results and Discussion

This section will be used to display and analyze various data sets generated by the lensing code. It is not meant to be an exhaustive survey of all the situations that can be treated by the code, but rather, it serves as an illustration of the type of data that can be produced, and a demonstration that the results generated are reasonable and physically justified.

6.1 Trajectory Analysis

It has already been mentioned in the previous chapter that the path taken by a beam through a region of FRW space is correct, regardless of the frame of reference used to describe the propagation. Here we take a look at trajectories that involve one or more lenses, where the beam needs to travel through vacuum regions of the model.

Figure 6.1 shows the effect that varying the core size of a lens has on the trajectory of a beam. In the simplest case, the lens boundary is uniformly filled with FRW dust at the same density as the background, so the trajectory taken by the beam from the observer to point A is simply the geodesic of the FRW space. Point B indicates where the beam ends up when the lens contains a core that starts with a radius equal to 90% that of the boundary. The path does not deviate significantly from the pure FRW case because

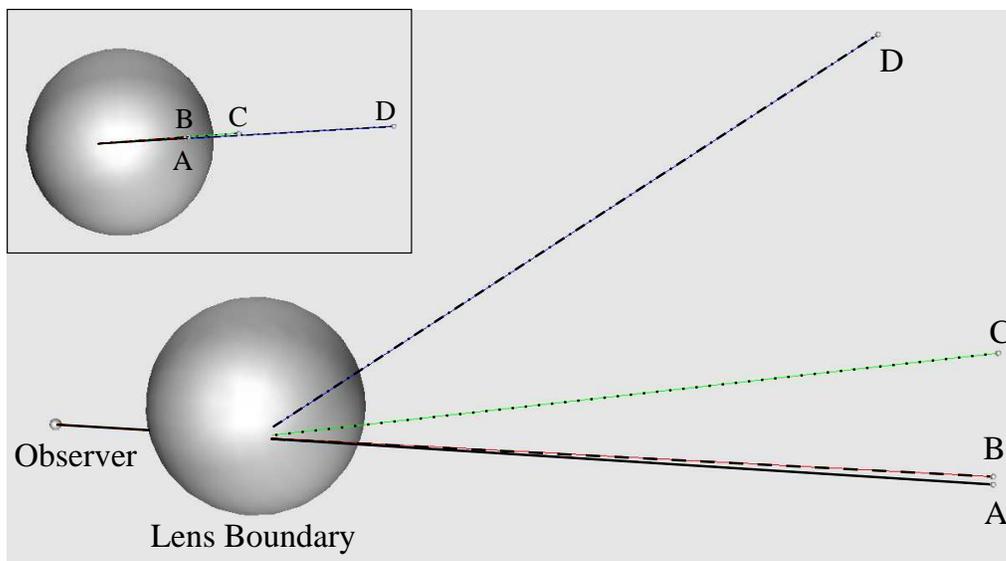


Figure 6.1: The effect that varying the core size has on the trajectory of a beam. The solid line ending at point A is the path the beam would take if there were no lens present. Propagation ends at point B if the core starts with a radius only about 10% smaller than the boundary, at point C when it starts with a radius of half the boundary, and at point D if the core is a small, dense object that the beam does not pass through. The inset image shows the system in another orientation, demonstrating that the trajectories are coplanar.

although the beam does spend a little time propagating in the vacuum region between the background and the core, most time inside the boundary is spent within the core. If the core of the lens starts with a radius that is half that of the boundary, then the beam spends more time in the vacuum and is deviated more than in the previous case to end up at point C. Finally, if the core is made small enough, the beam spends all of its time within the lens boundary in the vacuum and is affected by the entire lens mass, resulting in the greatest deviation from the original trajectory to end up at point D. Note that the trajectories are the same between the observer and the lens boundary, reflecting the fact that the direction of observation is the same in all cases. The inset image shows the system in another orientation and demonstrates that the four trajectories are coplanar as should be the case when the only difference between them is a radial variation in the size of the core they pass through.

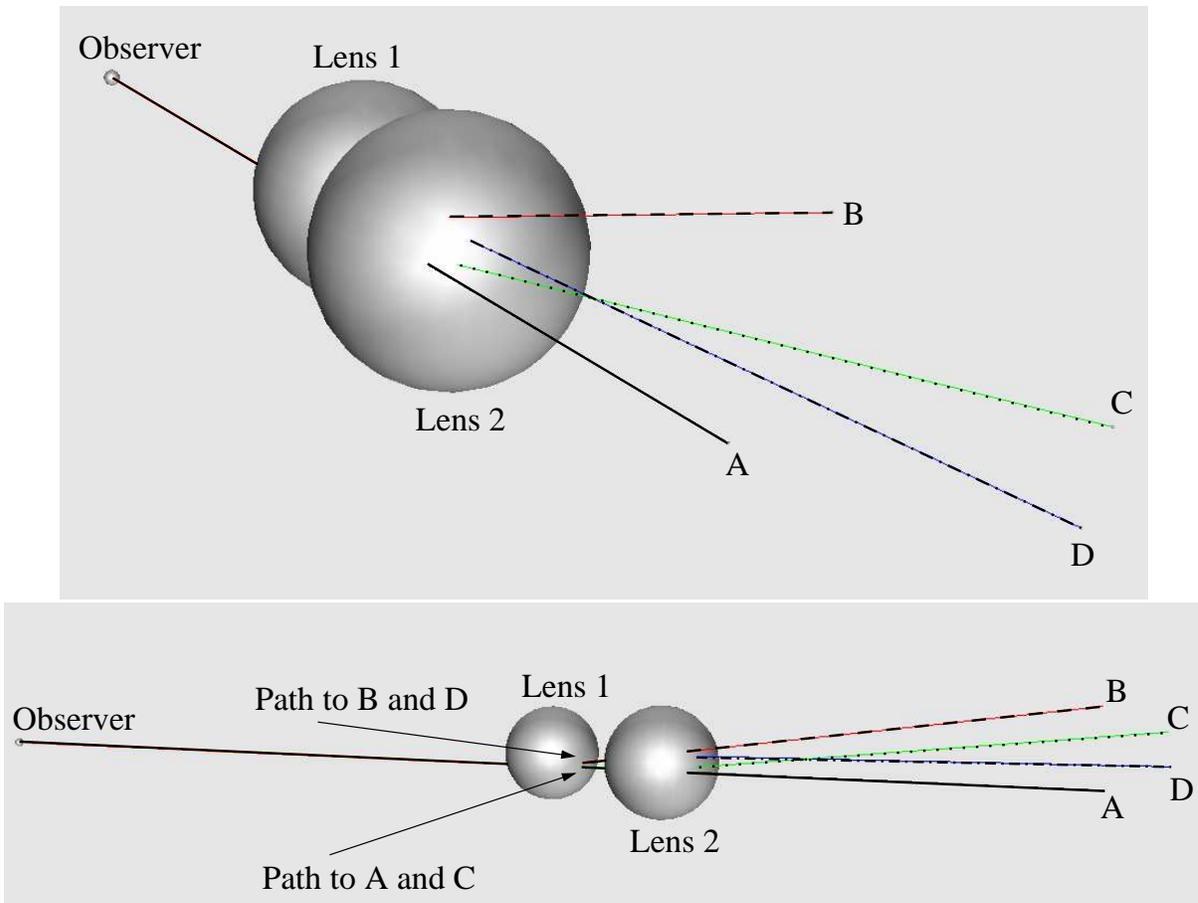


Figure 6.2: These are similar to Figure 6.1, but instead of varying the core size, change the number of lenses in the beam path. Point A again represents the location where the undeviated beam ends up. The beam stops propagating at point B if it only experiences the effect of Lens 1 (point D in Figure 6.1), at point C as a result of only passing Lens 2, and at point D after encountering both lenses. Note that in each case, the lenses contained small cores that were not hit by the beam in order to magnify the path differences.

Figure 6.2 is similar to Figure 6.1, but with the number of lenses encountered by the beam being varied instead of the core size. Two different views of the same system are shown. As was the case before, point A represents the end of the path taken by a beam when the lens boundaries are filled to the density of the background universe. When Lens 1 is given a small core and placed in the beam path, the beam ends up at point B, deviated mainly in the vertical direction as would be expected given the alignment of the undeviated beam and the centre of Lens 1. Placing only Lens 2 in the beam's path sends it to point C, a larger deviation than before which stands to reason since there is

more mass contained in this lens than in the previous one. The direction of displacement is also reasonable given the alignments. Finally, if both lenses are placed in the path of the beam, it ends up at point D. This is less of a displacement than was experienced with the second lens alone, but makes sense since the first lens deviates the trajectory of the beam further from the core of the second one, resulting in less bending taking place there.

6.2 The Optical Scalars

While it is important to ensure that the beams follow the correct trajectories as they are propagated through the RSC model, the results of most interest lie in the analysis of the amplification of the beams, along with the magnitude and orientation of the distortion they experience. As was explained in Section 5.2.2, the data produced by the code allows these quantities to be examined as they vary with redshift, and it is also possible to plot their final values for a set of beams sent through the model to see how they would vary over an observed patch of “sky”. This section will begin with a look at these values as beams are propagated through a uniform FRW region to demonstrate that the behaviour is reasonable in this simplest case. Next, trajectories that contain an encounter with a single lens will be considered, with beams first avoiding the central core, then later propagating through it. This is the situation that is easiest to assess analytically, and comparisons will be made to what can be found using the thin-lens approximation. Finally, a look will be taken at results generated using a general RSC model where multiple lenses are encountered.

6.2.1 Propagation Through the FRW Background

When propagating through a homogeneous FRW universe, the beam does not experience any distortion. Accordingly, when using the (ψ, η) form of the OSE, η remains zero,

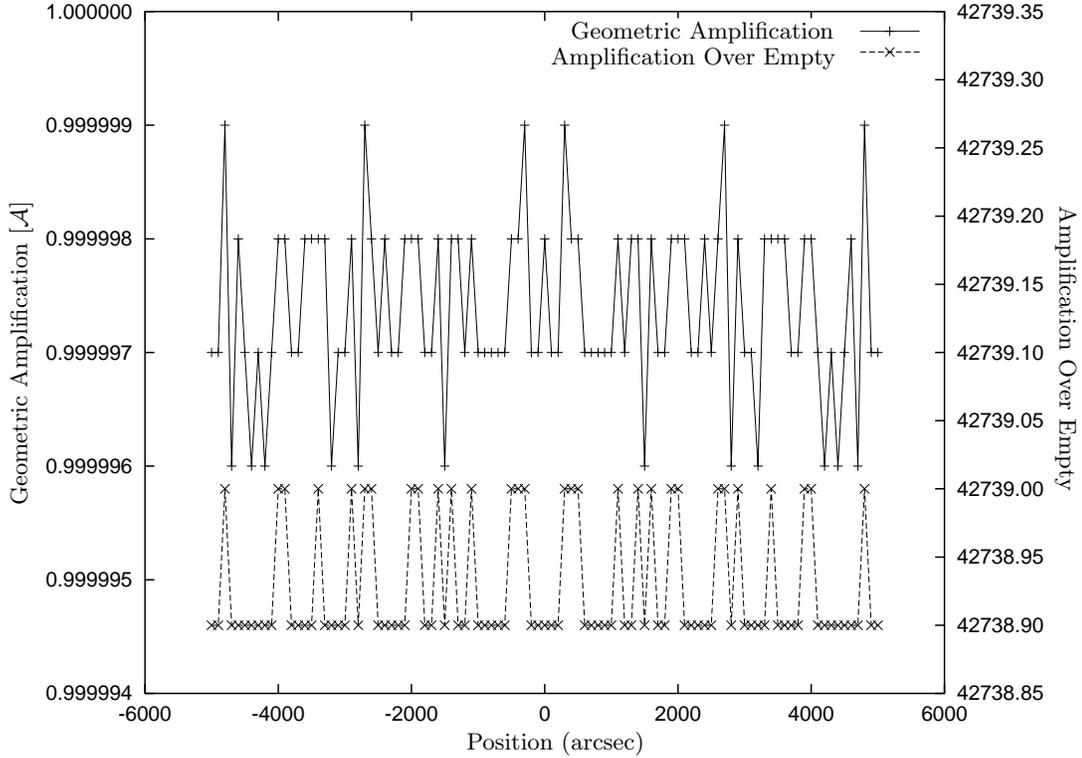


Figure 6.3: The upper line in the figure shows the geometric amplification experienced by a set of 101 beams propagated through a plain FRW region, while the lower one shows the amplification experienced when compared to propagation through an empty universe. The x-axis simply indicates the position on the “sky” each beam is propagated in, with the zero-point being directly ahead of the observer.

implying the distortion, \mathcal{D} , must also remain zero and so the orientation, α_e , is undefined.¹ Hence, in this situation, the only quantities of interest are the amplification of the beam, and the area used to compute it. The geometric amplification, \mathcal{A} , was defined in Section 5.2.2 as the ratio of the area the beam would have if computed via Mattig’s relation, to that determined by integrating the OSE. Figure 6.3 shows the value of \mathcal{A} computed for a set of 101 beams propagated through an RSC universe to a redshift of 1000 without hitting any holes (i.e. a plain FRW universe). The central beam, located at the zero-point of the x-axis, was purely radial and sent directly away from the observer, with $d\omega$ being the only non-zero component of its direction vector. Each of the

¹Equation (4.22) actually returns a value of 0 for it in this situation.

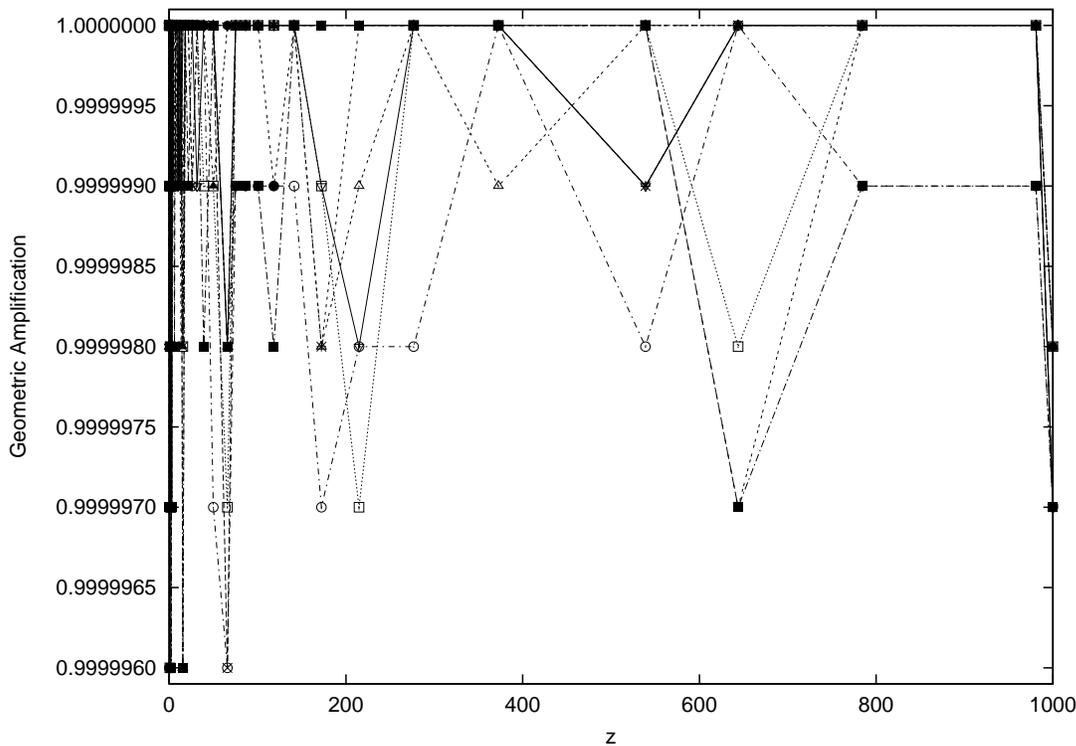


Figure 6.4: The variation in geometric amplification with redshift for a set of 10 beams propagated through an FRW region.

other beams, distributed at 100 arcsecond increments to either side of the central one, followed a geodesic which also had a non-zero $d\phi$ component to its tangent vector. The range of values for \mathcal{A} are slightly offset from the expected result of 1.0, varying about 0.9999975 instead as a consequence of the linear interpolation indicated below, but this is within an acceptable range of numerical variation. The other line plotted in the graph shows the amplification computed using the Dyer-Roeder formula for an empty universe as a reference value (refer to Section 5.2.2), and also contains small variations near the floating point precision limit.²

The evolution of \mathcal{A} with redshift for a sample of 10 of the beams is shown in Figure 6.4. While the values do fluctuate at the 10^{-6} level over the redshift interval, they are in line with the expectation that they maintain a value of unity. The fact that they all seem to

²Though double precision variables are used throughout the code, the heavy use of trigonometric function calls restricts the precision to the floating point limit for the most part.

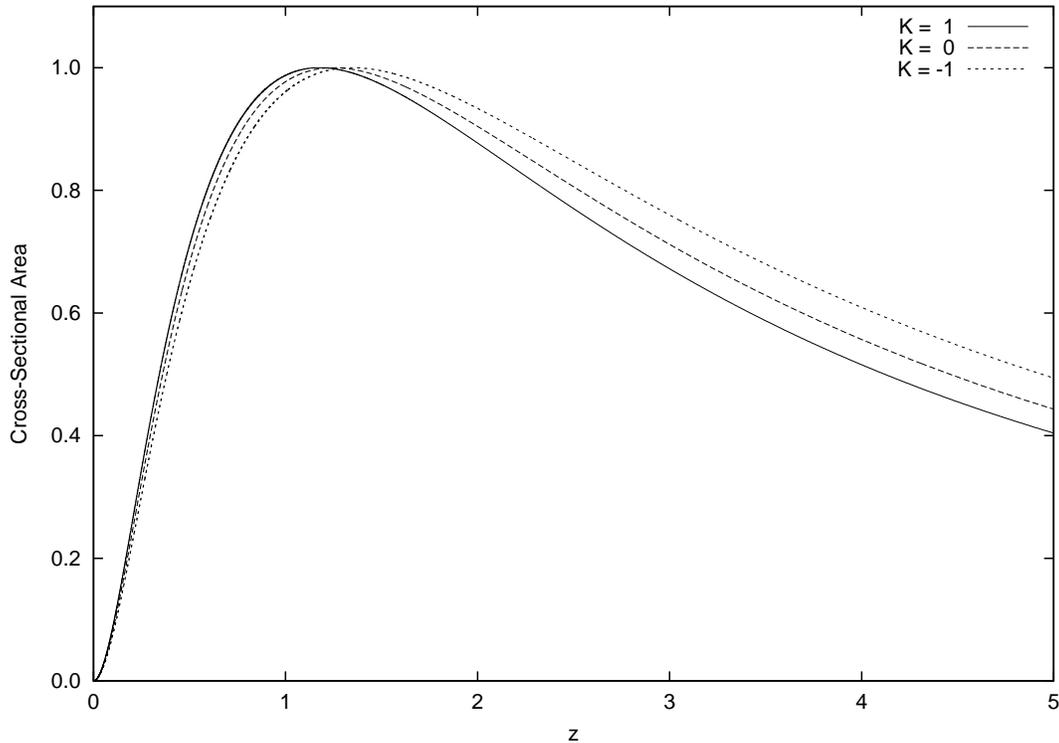


Figure 6.5: The variation in cross-sectional area for a beam propagated through FRW universes with positive, zero, and negative curvature. The vertical axis has been normalized so that the peak of each graph has a value of 1.0.

dip at the final data point is an artifact of the linear interpolation that is performed at the end of the propagation, but differences on such a small scale are unimportant.

Aside from the amplification, it is also possible to look at the variation in the cross-sectional area of the beam directly. Figure 6.5 plots the change in area experienced by a beam propagating back to a redshift of 5 through three different FRW models. The beam area increases away from the focal point at the observer until reaching a maximum value and turning around as a consequence of the Ricci focusing that it experiences. In the flat case ($K = 0$, $\Omega_m = 1$), the model describes an Einstein-de Sitter universe and the beam area turns around at a redshift of 1.25 as is expected. In the open case ($\Omega_m = 0.8$), the maximum occurs at $z \approx 1.35$, while in the closed case ($\Omega_m = 1.2$), it happens at $z \approx 1.18$. With the lower matter density of the open universe comes a smaller Ricci driving term, causing the beam to come to a focus more slowly than it does in a flat

universe, so when the beam is propagated backwards away from the observer, it reaches a maximum cross-sectional area at a higher redshift than happens in the flat case. The opposite is true in the closed case, where the higher matter density leads to a more rapid focus and hence, a turn around at a lower redshift.

6.2.2 Single Lens Without Core Encounter

With the results looking reasonable for the baseline case of a universe with no lenses, we now turn to the case of a single lens placed in the path of the beams. For the moment, we will consider a lens with a small, dense core, and the beams to be sufficiently offset from the centre of the lens that none of them actually reach the core. As in the previous section, we can first look at the variation in the geometric amplification experienced over a scan across the lens.

Aside from analyzing the results generated by using the optical scalars to ensure that they are physical and self-consistent, further reassurance that they are sensible can be obtained by comparing them to what is found using the thin-lens approximation discussed in Section 1.3.1. While this comparison can serve as a useful check though, one must bear in mind that the thick and thin lens setups are fundamentally different constructions. In the case of the thick lens, the beams must first pass through a section of the homogeneous FRW background until reaching the boundary of the lens and entering the Schwarzschild vacuum inside. It is only within this vacuum that the beams experience a non-zero Weyl driving force, causing cross-sections to distort and paths to deviate from the homogeneous case. Conversely, when a thin lens is used, its effect is felt by the beam immediately upon leaving the observer, and continues to be felt until the propagation terminates at the limiting redshift, though all bending is assumed to take place instantaneously at the point of closest approach to the lens.

The way in which a lens is defined in the RSC model has been described at length in earlier chapters. In order to set up a thin lens that is comparable to this, there are a few

parameters that need to be specified. Since the core is not hit by any beams in this case, the radius of the lens is unimportant (i.e. it can be considered to be a point mass), as is the density of the lens, with the constant geometrized mass being all that is needed in that regard. The only other values that need to be determined are the offset each beam has from the centre of the lens, and the redshift of the lens. While the geometrized mass is an unambiguous quantity, these two values are not as straightforward to set.

In the thin lens case, the offset angle is used in a simple way to determine the impact parameter of the beam. This is not the same as in the thick lens case where the beam must travel along geodesics of both the FRW and Schwarzschild regions before reaching closest approach, and only the initial observation direction needs to be specified. The redshift of the thin lens is required in order to compute the angular size distances shown in Figure 1.2. With the current setup, there is only a single lens so one need not worry about variations in angular size distances introduced due to earlier lenses. However, there is still a range of redshifts which one may choose to place the thin lens at, starting at the redshift that the boundary of the lens is found to be at when a beam enters the vacuum, to that it is at when the beam exits. Of course taking into account the evolution of the lens, varying the choice of beam used to make the measurement will lead to different limits for this range.

Figure 6.6 shows the variation in geometric amplification across a lens, analogous to Figure 6.3. Each line was generated using 101 beams separated horizontally by 300 arcseconds and propagated back to a redshift of 1000. The solid line showing the thick lens result was produced using a single SC hole with a small core. The lens is centred horizontally on the origin of the x-axis so the variation is symmetric about this point. Far from the centre, the amplification approaches a constant (refer to the inset in Figure 6.6). In the case of the thin lenses, this constant is unity, the value the amplification would have if no lensing were to take place. In the thick lens case, this value drops below one, indicating deamplified regions. This is a result of the flux conserving nature of the

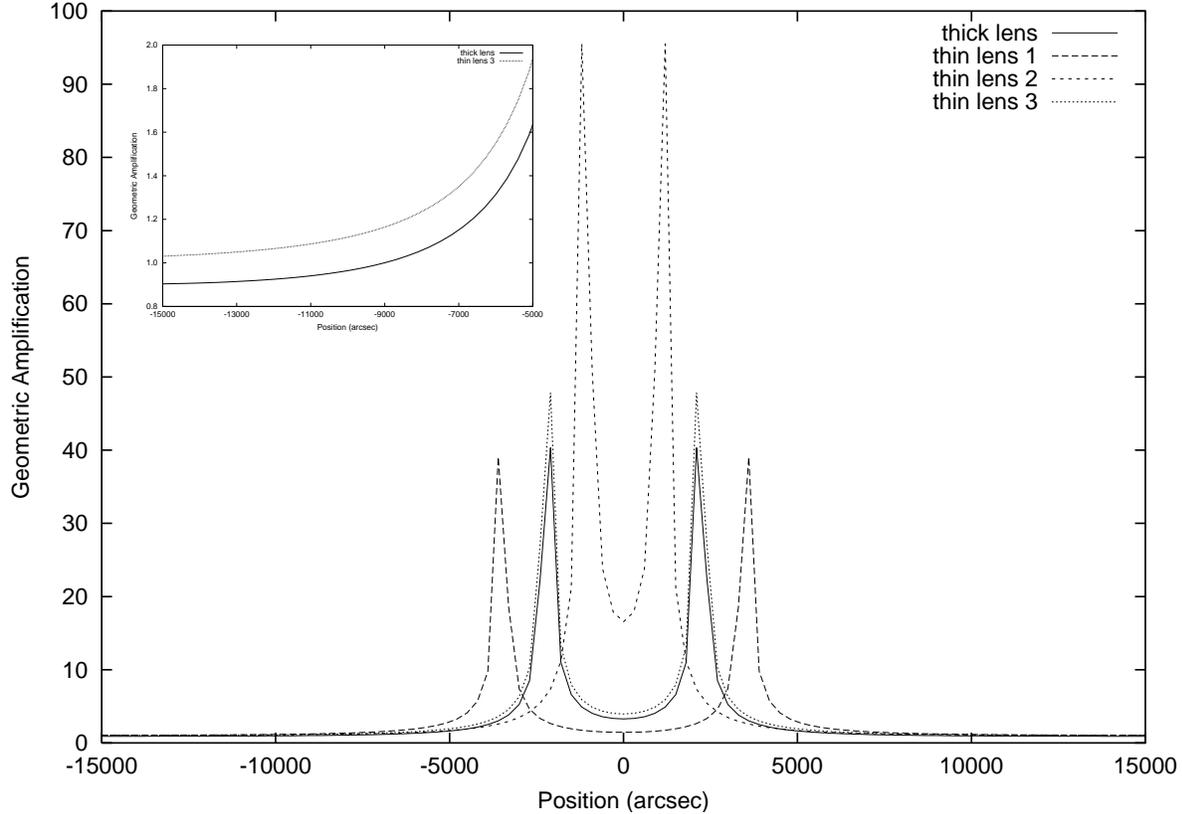


Figure 6.6: The variation in geometric amplification across a lens with no core hit. The solid line shows the result computed using a thick lens in an RSC model, while the thin lens results were generated using the method described in Section 1.3.1. The inset shows more detail of the evolution of the thick lens and third thin lens curves away from the centre. Refer to the text for more details.

optical scalar approach. Since the lens cannot add flux to a source, the amplification caused along some lines of sight through the lens is the result of null geodesics being diverted from other lines of sight, which correspond to regions of lower amplification. The thick lens curve will return to a value of one once it extends far enough from the centre of the lens that the beams used to generate it no longer hit the vacuum.

As the centre of the lens is approached, the lensing becomes stronger so the amplification increases until a maximum value is reached at a point where the beam passes through a caustic near the limiting redshift. The amplification decreases closer to the lens as those beams would have experienced stronger tidal forces, leading them to experience a caustic at a lower redshift and thus allowing their area to increase again by the time

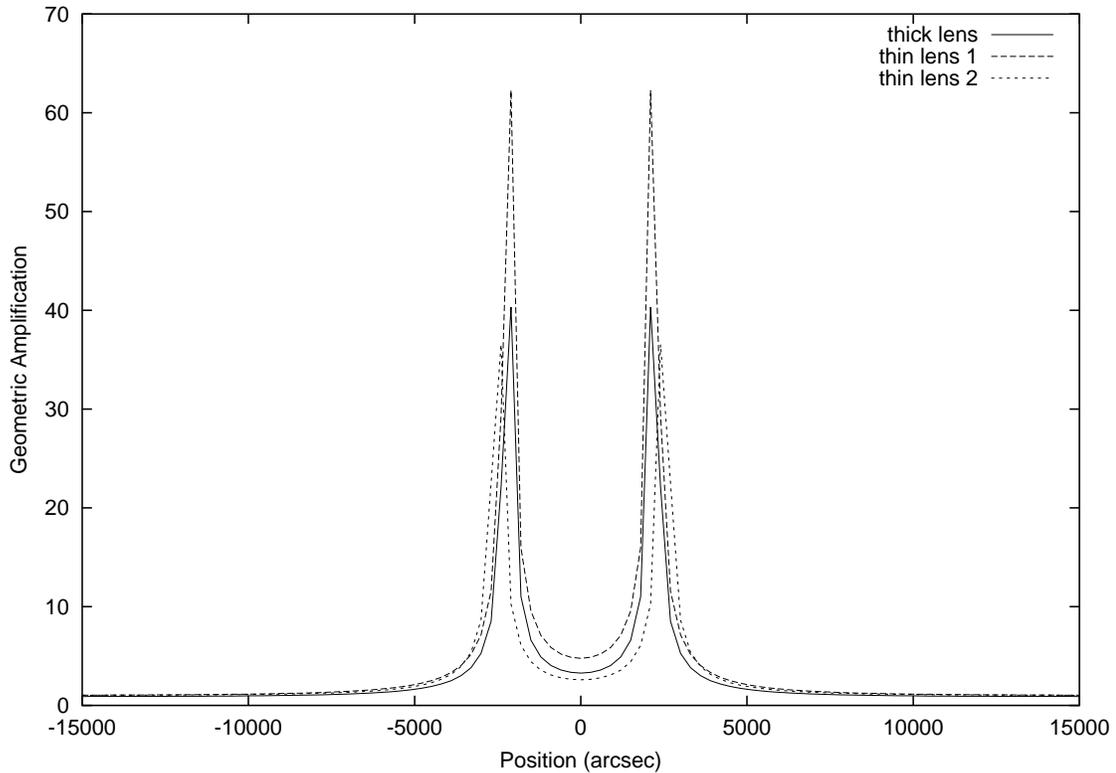


Figure 6.7: This is similar to Figure 6.6, but this time the curves for the two thin lenses were generated by varying both their redshift, as well as the offset angle of the central beam.

the limiting redshift is reached. This is illustrated in Figure 6.8.

While the general shape of the thin lensing results is similar to the thick lensing case, the details do differ which is not surprising given the difference between the two systems. The first thin lens was placed at a redshift of $z \approx 0.42$, which is about the redshift measured when the most radial beam sent through the RSC model reached the boundary of the lens. The second thin lens was placed at $z \approx 0.63$, the redshift measured when the same beam left the vacuum after passing by the lens. Placing the thin lens somewhere in between the two limits brings the agreement of the results much closer, as demonstrated by the line for the third lens which was placed at $z = 0.54$. Note that the magnitude of the peak in each case is not overly significant as it lies close to a caustic point and hence, is quite sensitive to slight changes in position. In each of these thin lens cases, the offset angle used was about 0.031 radians (that found by taking the inner

product of the direction vector of the central beam with the coordinate direction to the centre of the lens). If this angle is instead calculated by requiring the central beam to have the same impact parameter that it did when propagating past the thick lens, then a similar graph results, though the agreement between the different cases is much closer as can be seen in Figure 6.7. The thin lenses are the same as the first two used previously, with the curve for the first thin lens ($z \approx 0.42$) being generated with an offset of about 0.034 radians, while that for the second ($z \approx 0.63$) used an offset of 0.029 radians.

Figure 6.8 shows the redshift evolution of the geometric amplification for a set of four beams taken from those used to generate Figures 6.6 and 6.7. The upper plot shows the behaviour at low redshift. As expected, the amplification is equal to unity in all cases until the boundary of the hole is reached at a redshift of about 0.42. It then drops as the beams pass through the vacuum where there is no Ricci driving term, with each beam decreasing by a different amount.³ Note that there are no data points plotted for the vacuum propagation over the redshift range $0.42 \lesssim z \lesssim 0.63$ due to the lack of comoving observers within this region. The net redshift experienced by each beam across the hole is a little less than it would have been had the beam propagated the same distance through the background universe, which is expected since the beams do not experience any expansion of spacetime as they travel through this region. While their expansion rate is constant in the vacuum, their cross-sectional area does continue to change. In the case of beams 44 and 51, which pass closer to the core, when combined with the effects of an increasing shear rate, this change is enough to cause a net increase in amplification across the vacuum, with beam 51 experiencing more of an increase as one would expect. The other two beams do not experience enough shearing to compensate for the lack of Ricci driving, resulting in a decrease in amplification across the vacuum. Upon re-entering the background, the amplification begins to increase once again for all but beam 1, with the beams passing closer to the lens being more distorted (refer to Figure 6.11) and hence,

³Recall that to first order, the distortion caused by the Weyl driving generated by a Schwarzschild lens does not change the area of the beam.

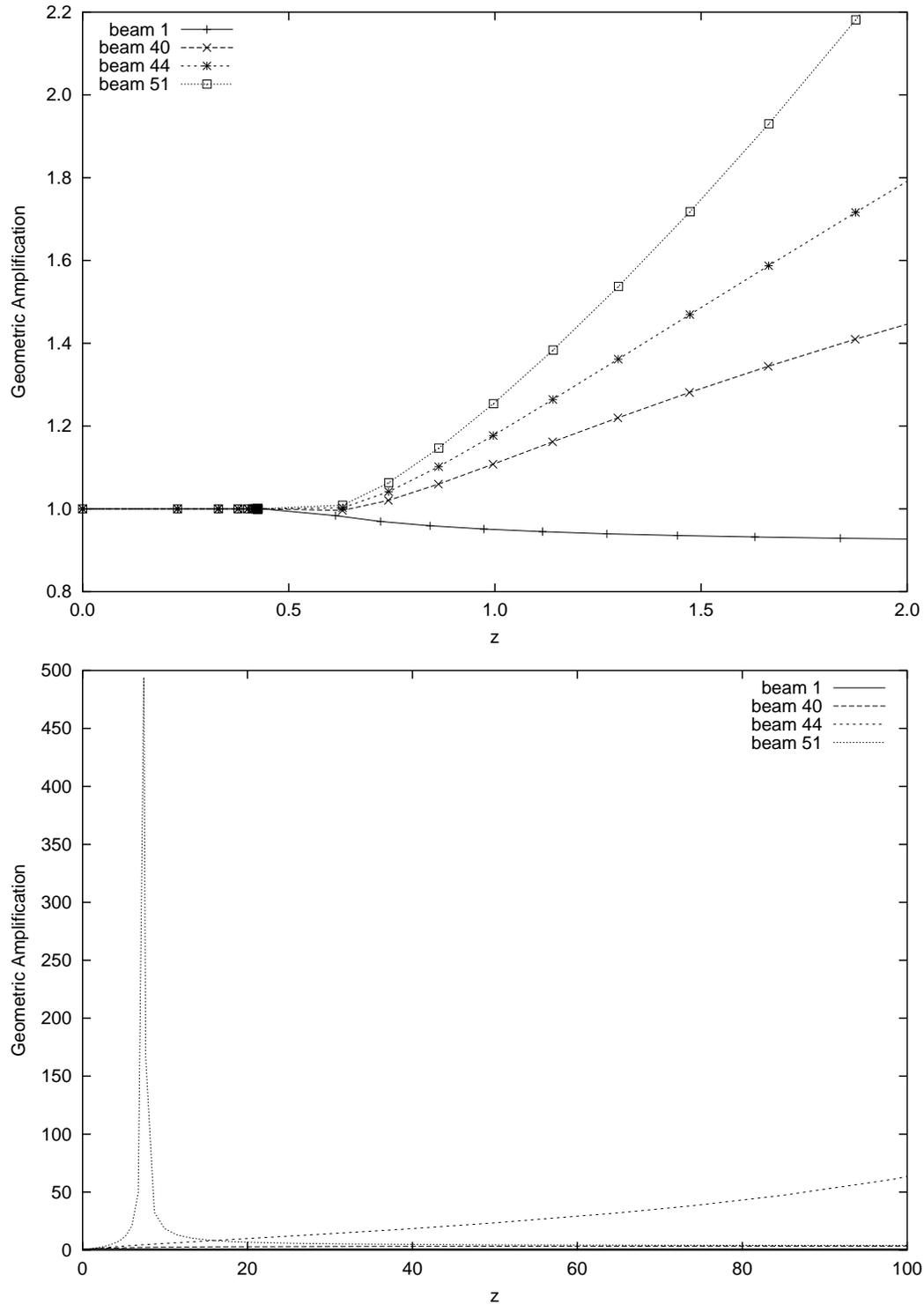


Figure 6.8: The variation in the geometric amplification with redshift for a series of four beams taken from the thick lens curve shown in Figure 6.6, with beam 1 passing furthest from the centre and beam 51 being the central one. Beam 44 had the largest final amplification, while beam 40 was offset a bit from that. The symbols in the upper graph indicate the points where the amplification was actually measured.

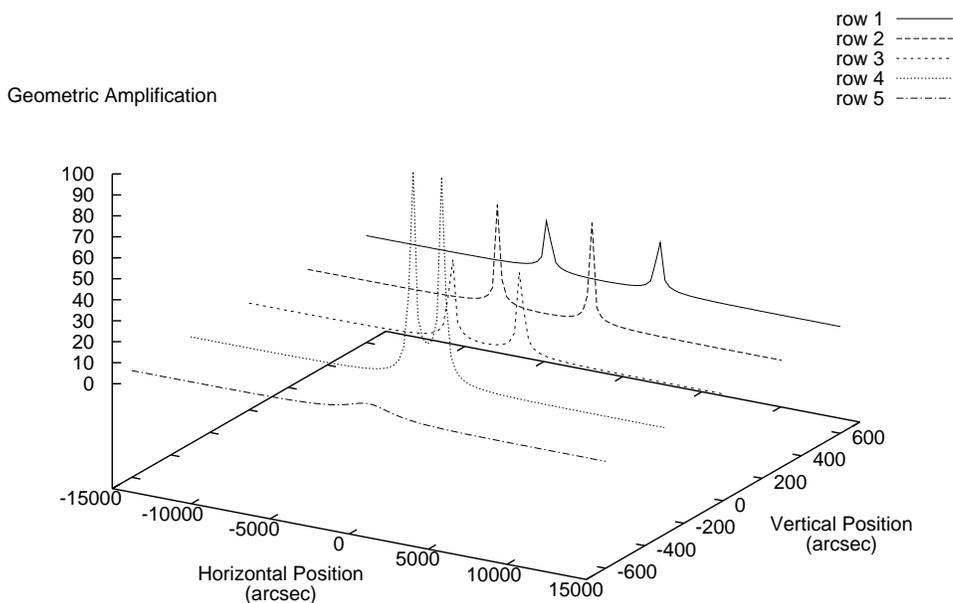


Figure 6.9: The variation in geometrical amplification across a lens for a set of five scans, each of which was generated using 101 beams. Row 3 contains the thick lens plot shown in Figure 6.6.

having their amplification grow more rapidly. While the rate of decline for beam 1 also reverses in the background due to the matter now present within it, the distortion it experiences in the vacuum is quite weak and as a result, it takes longer for the area to start changing quickly enough for the amplification to increase. The lower plot in Figure 6.8 shows the variation over a wider redshift range, with the caustic experienced by the central beam clearly evident. As the amplification of this beam drops afterwards, it is exceeded by that of beam 44 at a redshift of about 16, with the latter going on to reach a maximum near a redshift of 180, leading to the higher final amplification evident in Figure 6.6.

Figure 6.9 contains the same data as the thick lens plot of Figure 6.6, but also shows two scans offset vertically by 300 arcseconds to either side of it. The scan on the bottom left of the graph (row 5) is furthest from the centre of the lens, while that on the top right (row 1) is closest to it. As before, the amplitude of the spikes in each case are not

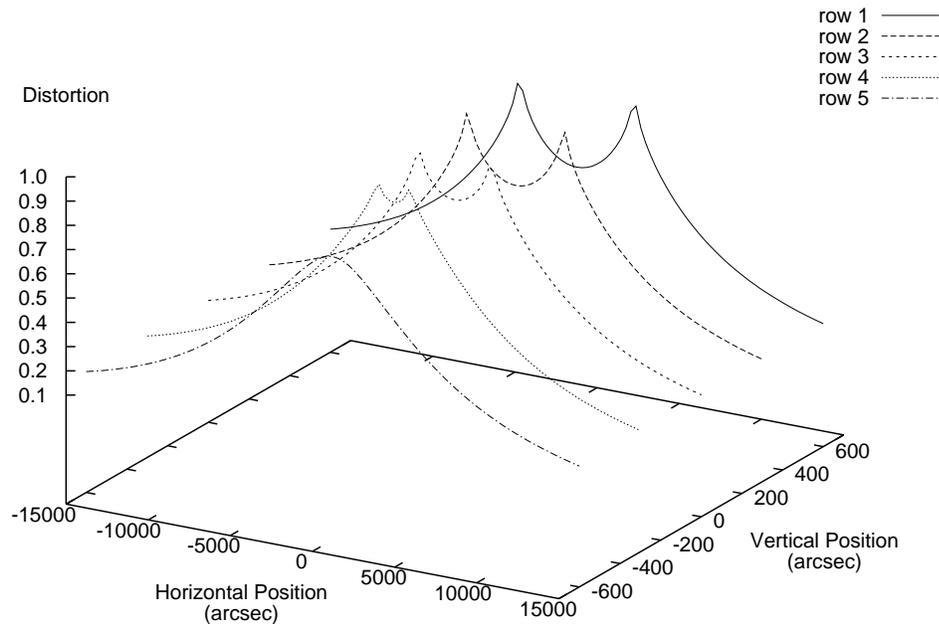


Figure 6.10: This plot is analogous to that in Figure 6.9, but shows the variation in the distortion of the beams used in the five scans instead of their amplification.

particularly significant due to their proximity to a caustic. More important is the fact that these points move further apart as one gets closer to the centre of the lens, and vanish altogether if one moves far enough away from the centre, as is to be expected for the circular caustic structure of a spherically symmetric lens.

Figures 6.10 and 6.11 are analogous to Figures 6.9 and 6.8 respectively, but deal with beam distortion as opposed to geometric amplification. Evidence for the same circular caustic structure is seen in Figure 6.10, with the distortion approaching the maximum value of 1 at separations corresponding to those of the peaks in amplification shown in Figure 6.9. In Figure 6.11, the initially circular beams have zero distortion until the boundary of the vacuum is hit, at which point the distortion grows at a rate that varies inversely with distance from the centre of the lens. The central beam has a cross-section that collapses to a line at a redshift of about 7.5, after which its distortion decreases, with beam 44 eventually becoming more distorted as it collapses down to a line later on

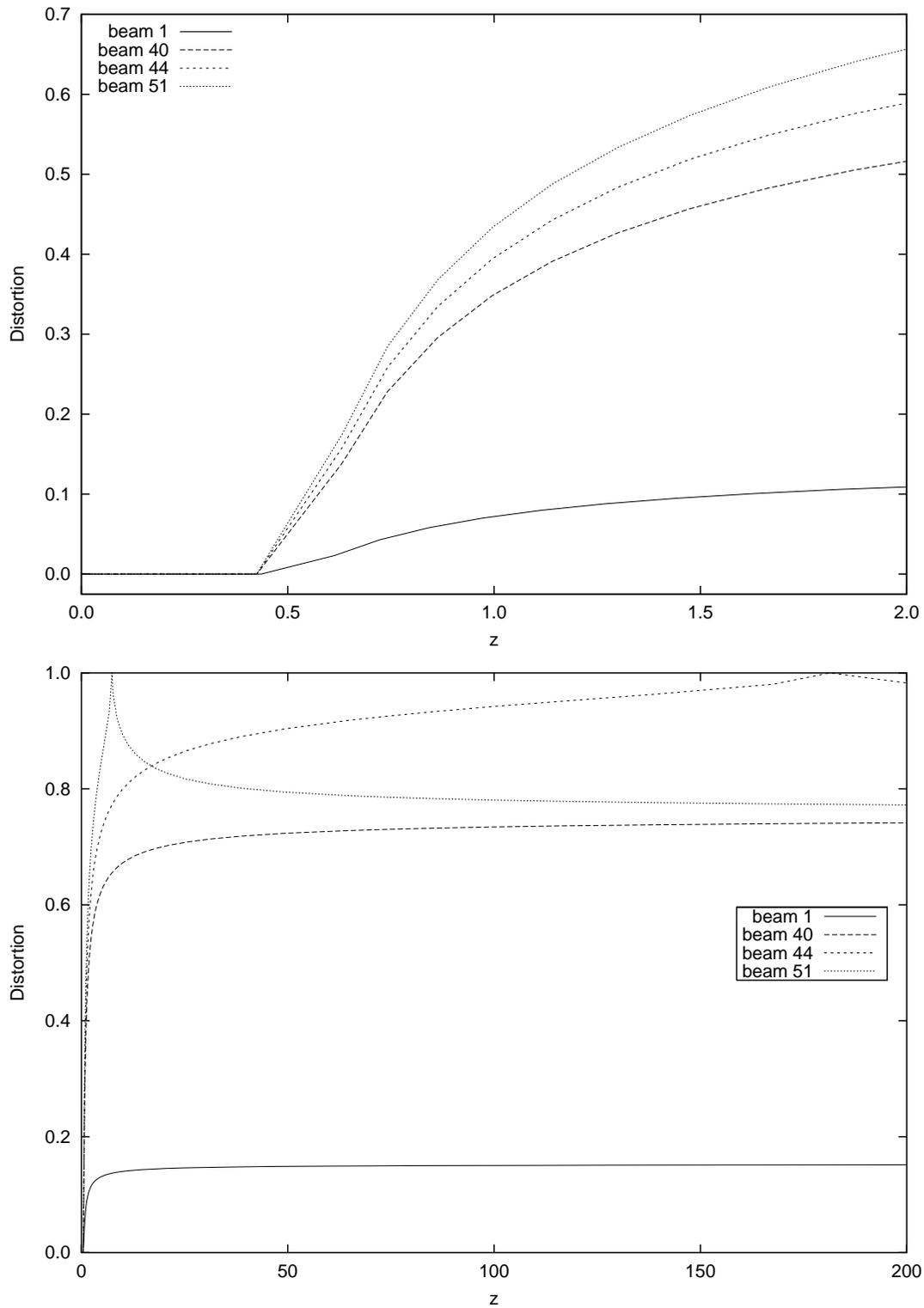


Figure 6.11: The variation in distortion with redshift for the set of four beams depicted in Figure 6.8.

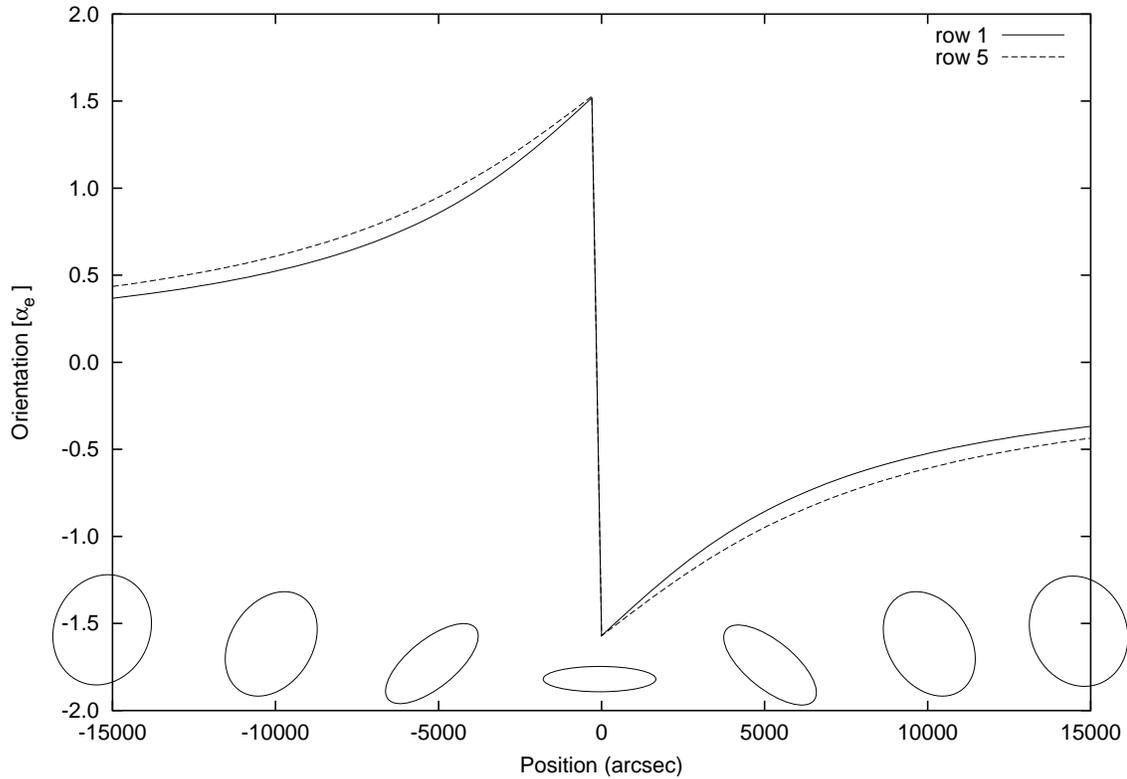


Figure 6.12: The variation in the orientation of the distortion of the beams in the first and last horizontal scans shown in Figure 6.10. The cross-sectional shape of a few beams in row 5 has been plotted along the bottom of the graph, with the semi-minor axis of each ellipse being inclined at an angle α_e from the horizontal.

in accordance with what was observed in the amplification plots.

Aside from the magnitude of the distortion, it is also possible to look at the orientation of the distorted beams. Figure 6.12 plots this quantity for the beams in the first and last scans shown in Figure 6.10. The fiducial vector to which the orientation is compared (j^a in Section 4.5.1), was chosen to lie in the horizontal direction. Since the scans are offset vertically from the centre of the hole, the outermost beams have semi-minor axes with non-zero orientation values, and these can be shown to be correct using simple trigonometry. In both cases, as one scans across the hole, the orientation increases until reaching a value of $\pi/2$ at the centre. This must be the case as in this alignment, the lens causes the beam to be compressed vertically. It then jumps to negative values due to α_e being restricted to lie between $\pm\pi/2$, and increases smoothly in a mirror image of

the behaviour approaching the centre. Row 5 is further from the centre of the lens than row 1, so accordingly, the magnitude of the orientations imposed on the beams in it are larger than those of the corresponding beams in row 1.⁴

6.2.3 Single Lens With Core Encounter

We now consider the situation where the lens used in the previous section is given a core that initially has a radius equal to a quarter that of the hole. This is large enough for some of the beams to pass through it, though it is not yet given any substructure.

Once more, it can be useful to compare the amplification computed using the optical scalars for beams passing through the thick lens, with the results found using a thin lens of the same mass. It should again be stressed that the two systems are not identical though, a fact which is even more relevant with the core being hit since the thick lens evolves while the thin one used does not. However, the comparison can still provide some reassurance that the two methods are in broad agreement. While the redshift of the lens and the offset angle to use for the central beam still need to be set as they were in the case of no core encounters, since some beams now travel through the core, its size and density are also important and need to be specified.

Though the total mass of the core remains constant as it evolves, there is a range of sizes and densities that can be selected for it, corresponding to the values found between the time the beam enters the hole and the time it exits. In order to be used in the thin lens computation, the chosen density needs to be converted to a mass profile that is projected onto the plane of the lens. This is because as indicated in Section 1.3.1, when a beam passes through the matter in the core of the lens, the only mass that affects it tidally and contributes to the distortion it experiences is that lying interior to it, often referred to as the *cylindrical mass*. For a sphere of uniform density ρ and radius a , the

⁴Other than the central beam, which must have the same orientation in both rows.

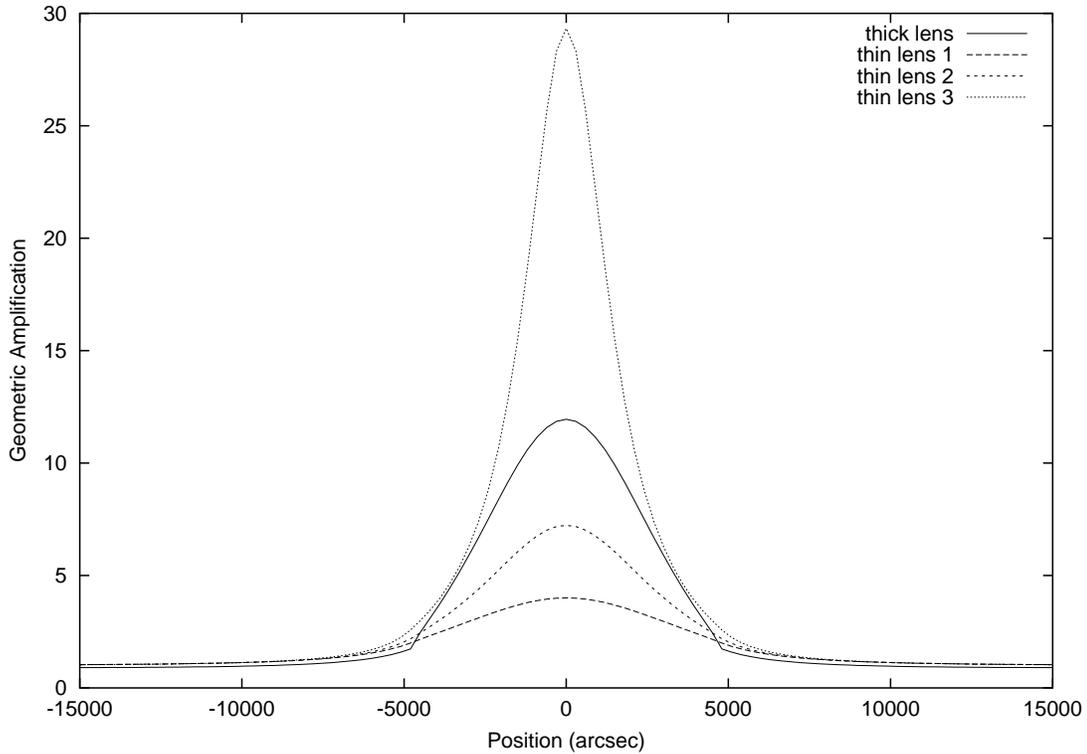


Figure 6.13: The variation in geometric amplitude across a lens with a core large enough to be encountered by some of the beams. As in Figure 6.6, the solid line was generated using a thick lens in an RSC model, and this is compared to a set of lines computed using three different thin lenses.

cylindrical mass lying within an impact parameter distance h is given by

$$m(h) = \frac{4\pi\rho}{3} \left[a^3 - (a^2 - h^2)^{\frac{3}{2}} \right],$$

which can be found by simply integrating up the mass contained within a series of concentric cylinders out to a distance h (refer to Appendix B).

Figure 6.13 shows the variation in geometric amplitude as measured by a set of 101 beams propagated past the lens in 300 arcsecond increments. As was the case when generating Figure 6.6 where none of the beams passed through the core, the central beam is still offset from the centre of the lens. The reason this time is due to the coordinate system being used not allowing for a purely radial ingoing beam to be treated, as mentioned in Section 3.3.3 when discussing the dynamics of cores. This is because it would be identical to the “timing beams” used to establish the ν coordinate, and hence,

would remain at a constant value of ν , making this variable unsuitable for use as an affine parameter.

In accordance with what was seen before, the further the beams pass from the centre of the lens, the more the amplification they experience approaches a constant value. This is undeviated from that of a pure FRW universe in the thin lens cases, while being slightly under amplified when using the thick lens. In fact, the portion of the thick lens curve generated by beams that only pass through the vacuum is identical to that shown in Figure 6.6 as must be the case. While the amplification again increases towards the centre, the peaks due to caustics seen before when the beams missed the core are not present here. This is because once the beams hit the core, the Weyl shearing stops and their distortion rate no longer increases, which means they take longer to collapse down to a caustic line, and in fact, do not do so before reaching the limiting redshift imposed at $z = 1000$. The sudden change apparent in the shape of the thick lens curve occurs between beams 35 and 36, the last one to pass through the vacuum without striking the core, and the first to spend some time propagating through it respectively. Thin lens 1 was defined using the density, radius, and redshift that the core had in the thick lens case when the central beam first encountered it, while thin lens 3 used the values found when the beam was close to exiting it.⁵ The second thin lens used the intermediate values defined when the central beam passed the point of closest approach to the middle of the core. In all three cases, the central beam offset was chosen so that the same number of beams (31 out of the 101) hit the mass distribution. Note that while there is general agreement between the forms of the thick and thin lens curves, the details of the shape do vary as is to be expected when comparing an evolving core to a static one.

Figure 6.14 shows the variation in beam distortion for a series of scans across the lens used to generate Figure 6.13. As with Figure 6.10, row 1 is closest to the centre of the lens, while row 5 is the most distant. In each scan, the two peaks indicating the greatest

⁵Use of the values exactly at beam exit results in large amplifications which compress the relevant range of interest for the other three curves.

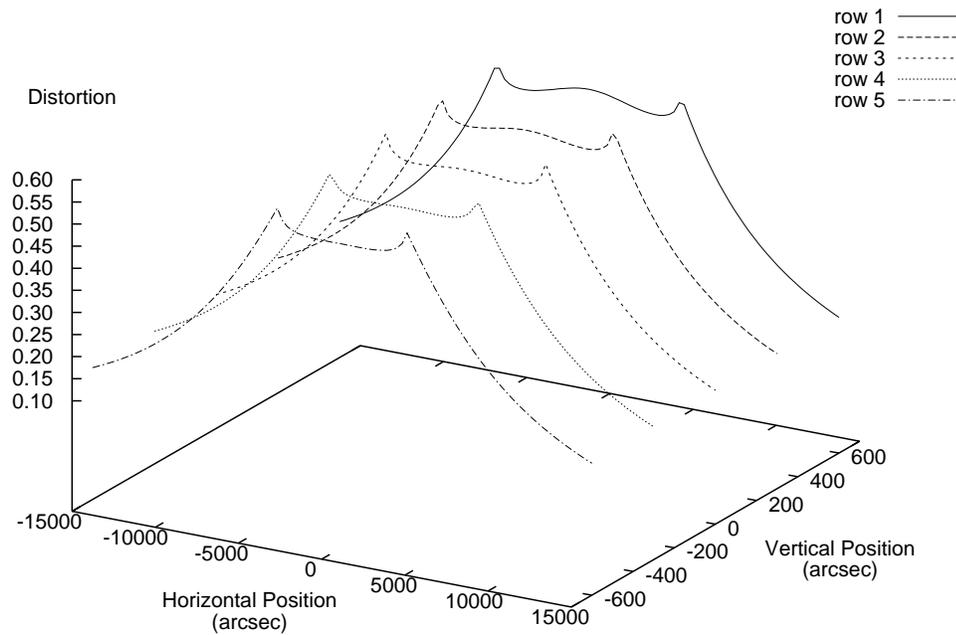


Figure 6.14: This plot is analogous to Figure 6.10 and shows the variation in the beam distortion for a series of scans across the lens used to generate Figure 6.13.

distortion were produced by the last beams to propagate through the vacuum without hitting the core. This makes sense as these beams spend longest in the vacuum and pass closer to the central mass than the others that do not actually hit the core, hence being exposed to the greatest amount of shear. As expected, these peaks get closer together further from the centre of the core, with 37 beams passing through the core in row 1, but only 25 doing so in row 5. Note that as the scans get closer to the centre of the core, a broader peak starts to form in between the two maxima in distortion. While this may at first seem counter-intuitive since the most central beams spend the least time in the vacuum, one must also consider the strength of the tidal force experienced during the time they do propagate through it. With the core evolving, the most central beam must climb out of the potential well with the deepest bottom, and as a result experiences the strongest tidal force of all beams when initially emerging from the core. The interplay between this factor and the time spent in the vacuum can lead to the non-monotonic

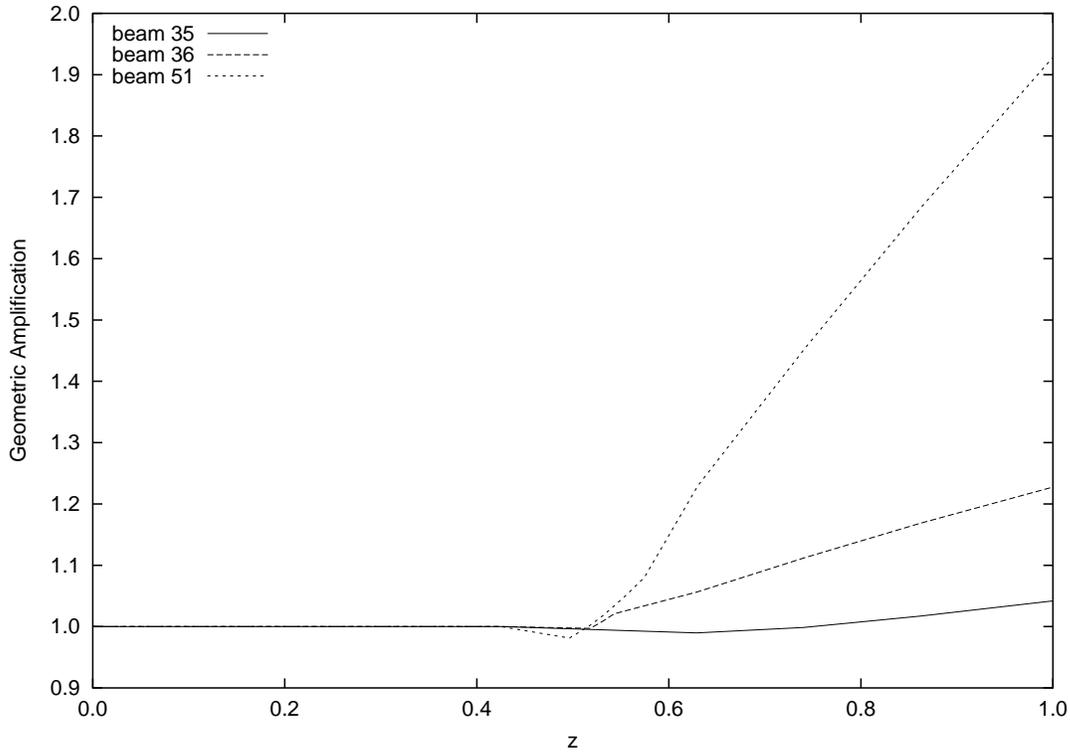


Figure 6.15: The variation in \mathcal{A} with redshift for a set of three beams which were used to generate the previous figures in this section. Beam 35 was the last to pass through the lens without hitting the core, while beam 36 was the first to pass through a portion of it. Beam 51 had the most radial trajectory of all the beams and hence, was the one that spent the longest time propagating inside the core.

behaviour observed in the plot.

The evolution of the geometric amplification out to a redshift of 1 is shown for a set of beams in Figure 6.15. The curve shown for beam 35 is in line with those seen for some of the beams in the upper plot of Figure 6.8, with a jump taking place across the vacuum, from $\mathcal{A} = 1$ for the initial propagation through the FRW background, to a lower value at the point of re-entry. There is a similar jump shown for beams 36 and 51, the first to hit the core on one side, and the central beam respectively. However, in these cases, there is also a subsequent jump when the beam leaves the core to propagate through the vacuum back to the boundary. More details of this behaviour can be seen in Figure 6.16, which looks at a narrower redshift range and shows the individual data points used to define the amplification and distortion curves for beam 51. The amplification remains at

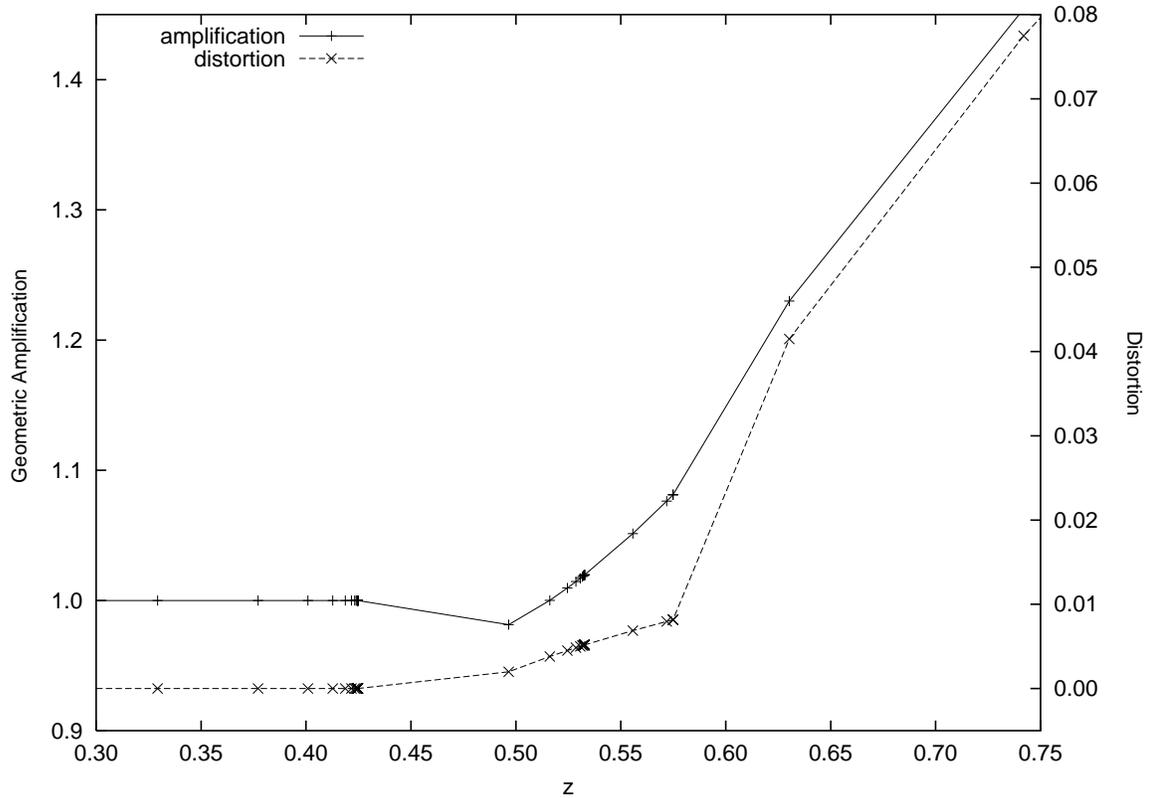


Figure 6.16: The variation in geometric amplification and distortion with redshift for beam 51 as it propagates from the FRW background, through a hole and its central core, then back to the background.

unity and the beam undistorted until it strikes the boundary of the hole at a redshift of about 0.42. It then propagates through the vacuum until encountering the edge of the core, which is at $z \approx 0.50$ when measured with respect to an observer comoving with this edge.⁶ As the beam passes through the core, its amplification and distortion increase until it emerges at a redshift of about 0.57. There is another jump in the data as the beam again travels through the vacuum before re-entering the background at the other end of the hole with a redshift of about 0.63, as measured by a comoving observer at this boundary. The evolution then continues in the background, with the amplification changing at roughly the same rate as it was in the core, but the distortion increasing significantly more rapidly.

⁶Recall that no points are shown for the vacuum portions of the trajectory.

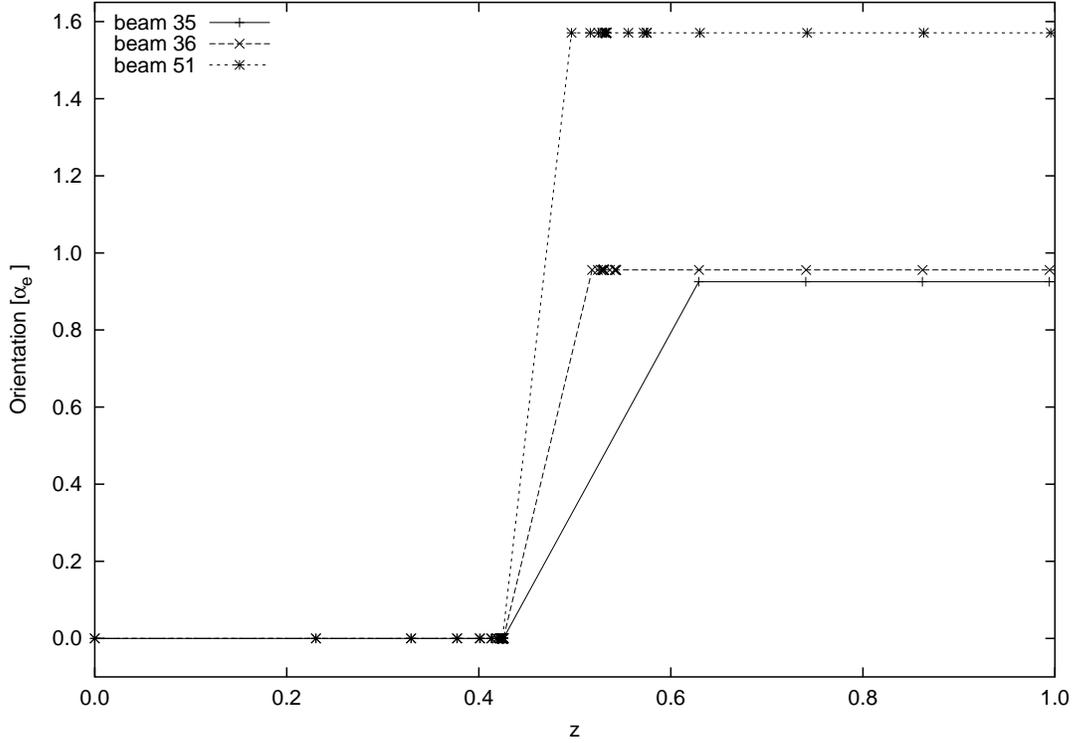


Figure 6.17: The imposition of an orientation on the cross-section of the three beams used in Figure 6.15. α_e is set for each beam when it starts to distort, and remains constant for the rest of the propagation.

The evolution of the orientation of the same three beams shown in Figure 6.15 is plotted in Figure 6.17. Each of the beams initially has a circular cross-section and hence, an undefined orientation⁷ until the boundary of the lens is reached around a redshift of 0.42 (the exact value varies for each beam). When beam 35 emerges from the vacuum at a redshift of about 0.63 without having hit the core, its elliptical cross-section is seen to have a semi-minor axis making an angle of about 0.93 radians with the projected fiducial direction \tilde{j}^a (see Section 4.5.1). This orientation remains fixed for the remainder of the beam's propagation since this is the only lens that it encounters. The curves for both beams that pass through the core display the same behaviour, though in these instances, the orientation has been imposed by the time they have hit the core at $z \approx 0.52$ and 0.50 for beam 36 and 51 respectively. This must be the case since it is actually set as

⁷Recall that a value of $\alpha_e = 0$ is returned by equation (4.22) in such a case.

soon as the vacuum is entered and the beams start to distort. It is constant during the propagation through the core and the second trip across the vacuum, and remains the same as the beams propagate through the background again. Note that beam 51, which passes closest to the centre of the lens, has the largest final orientation, while beam 35, which is the furthest of the three, has the smallest one, consistent with what is expected from Figure 6.12. A similar scan of orientations in this case displays the same behaviour shown there.

6.2.4 Multiple Lenses

We finish this section with a look at the lensing results generated by a full RSC model containing multiple lenses, each of which can potentially have substructure of its own. The background universe selected for the run was an open FRW model with $\Omega_m = 0.28$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The packing used to fill the background contained 190 078 spheres, while the two made available to describe substructure in the cores contained a total of 187 007 spheres. There were 121 beams propagated from the observer to a redshift of 1000 in an 11x11 grid with a 50 arcsecond spacing. The run took about three seconds per beam on average on a P4 1.6 GHz machine, and resulted in a model where 10 lenses were found, 5 of which had cores that were hit.

Figure 6.18 shows the variation in the geometric amplification “observed” over the patch described. The result is not as simple to interpret as it was for the specially constructed single lens systems depicted earlier. However, one can see how the structure encountered seems to have caused an area of deamplification, with the decrease being more prominent along a diagonal strip running across the graph. Figure 6.19 shows the corresponding variation in distortion over the patch, indicating that the beams experiencing the least amplification have actually been distorted the most from their circular cross-section. This suggests that they have not spent much time propagating through high-density regions where they would be subject to more intense focusing, but may have

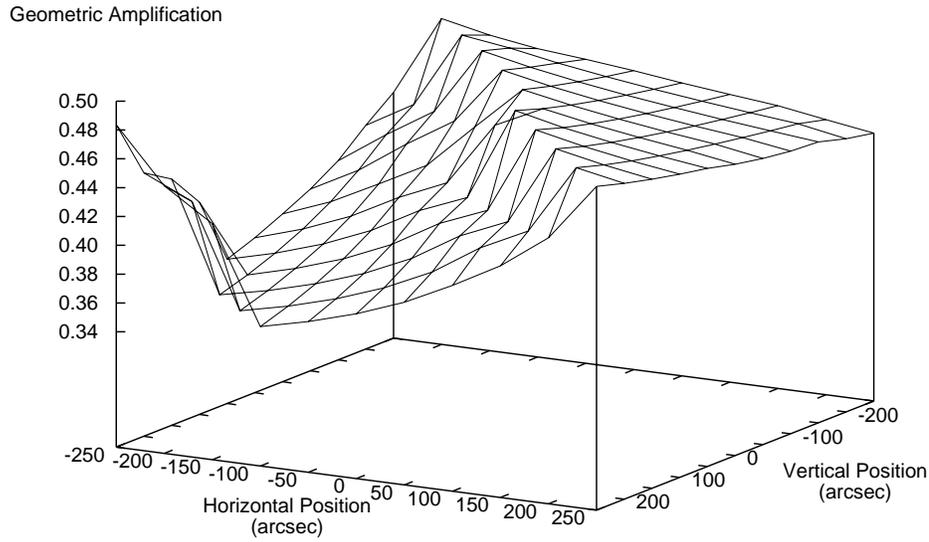


Figure 6.18: The variation in geometric amplification over a 500x500 arcsecond patch of “sky” in an RSC model. An 11x11 grid of beams was propagated back to a redshift of 1000 to generate the values used to produce the surface shown.

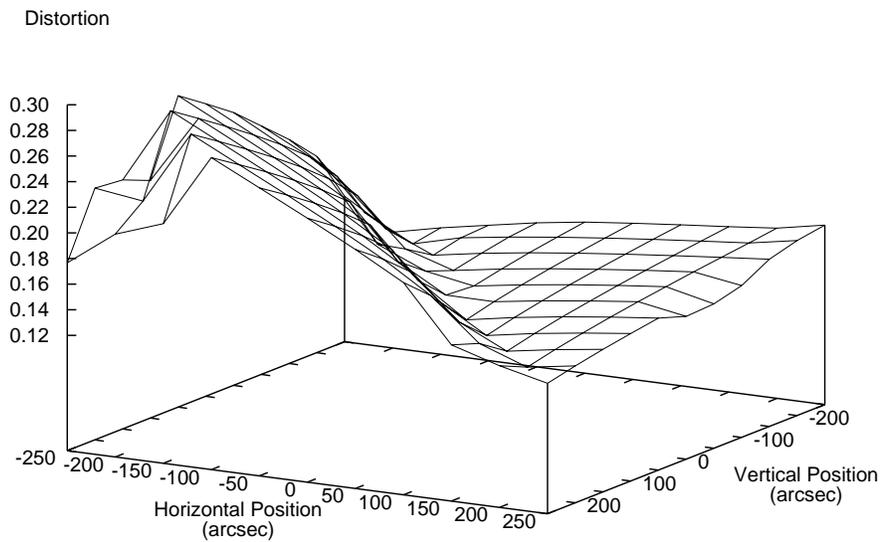


Figure 6.19: The changes in beam distortion over the patch depicted in Figure 6.18.

passed close enough to a mass concentration to experience significant tidal forces.

The low redshift evolution of the geometric amplification that the central beam undergoes is shown in the upper plot of Figure 6.20. As expected, this evolution is not as clean as it was in the single lens cases. In order to help interpret it, the lower plot displays the variation in the normalized cross-sectional area of the beam as a function of redshift, alongside the variation in area that the beam would have experienced were it to propagate the same distance through the homogeneous background. Note that the smoothness of the curves in the lower plot is a result of re-integrating the system over the redshift interval while forcing the integration routine to use a small and constant step size in order for it to compute the state of the system more frequently.

The beam initially propagates through the FRW background until reaching the boundary of the first lens at a redshift of about 0.64. It then propagates through the vacuum and hits Core 1 at $z \approx 0.70$ (as measured by an observer comoving with the edge of the core) with an amplification larger than it would otherwise have had. This is reasonable since the lack of Ricci driving in the vacuum means that the area does not increase as quickly over the interval as it does in the homogeneous case. While it propagates inside this core though, the amplification drops as the increased Ricci driving from the higher density matter causes the beam area to expand more quickly than in the homogeneous case. Accordingly, the propagated beam area in the lower plot gets closer to that of the homogeneous case until substructure is encountered within the core, with the boundary of this second lens being reached at a redshift of about 0.78. The subsequent vacuum leads to another jump up in amplification, though a smaller one since Core 1.1 is encountered soon afterwards at a redshift of about 0.79. The beam propagates through this FRW region without encountering any additional substructure and the amplification initially continues to decrease. However, the increased density in the core means that the beam area starts to turn around more quickly than in the background, leading to a divergence of the two area measurements and an increasing amplification by the time

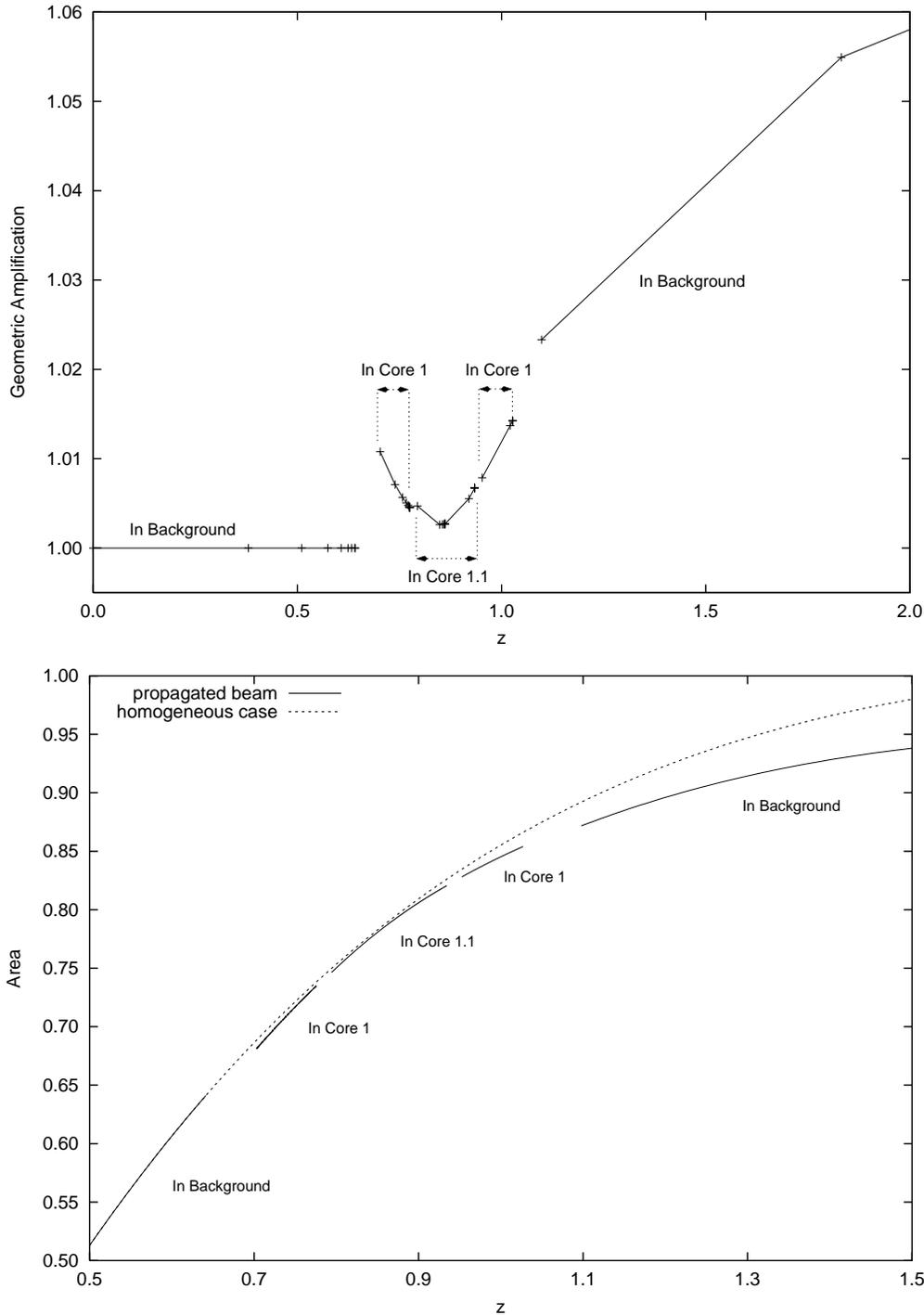


Figure 6.20: The upper plot shows the variation in \mathcal{A} for the central beam propagated through the model described earlier over a redshift range containing an encounter with two lenses, one of which (Core 1.1) lies within the other (Core 1). The lower plot compares the change in the normalized cross-sectional area of the beam over a narrower range in redshift, to that it would experience if propagated through the homogeneous background universe as computed via Mattig's relation.

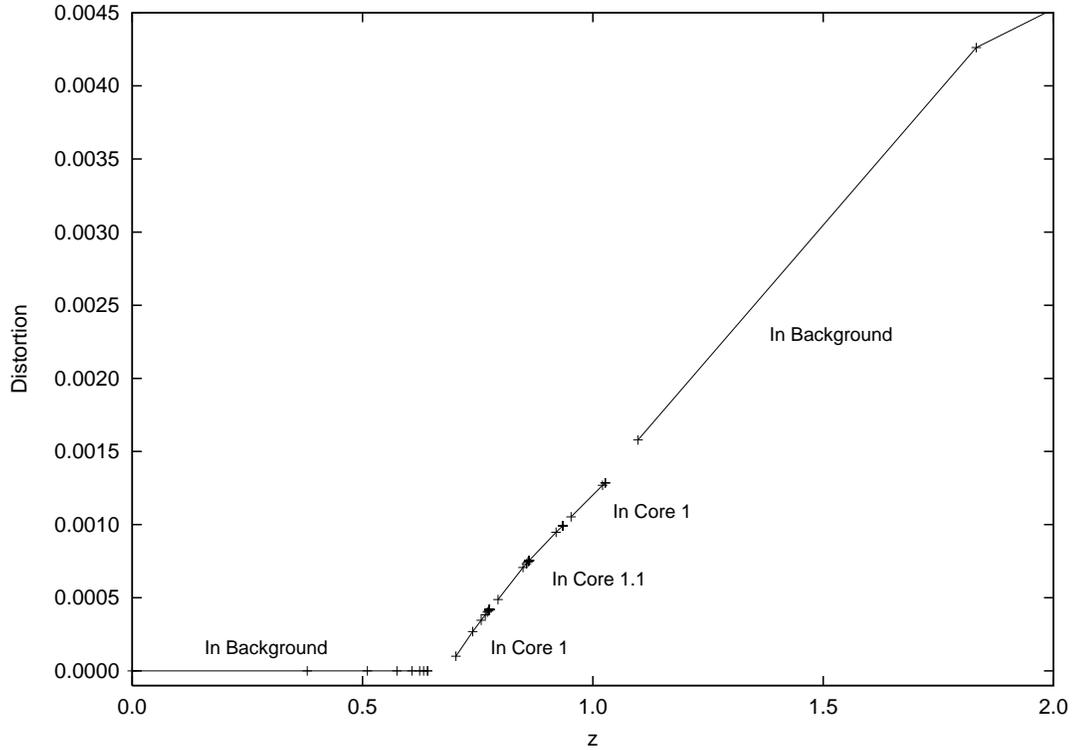


Figure 6.21: The distortion experienced by beam 51 over the redshift range that was used to illustrate the evolution of its amplification in the upper plot of Figure 6.20. Once again, the gaps in the curve correspond to periods of vacuum propagation.

the beam leaves the core at a redshift of about 0.93. The trend continues as it returns to Core 1 at $z \approx 0.95$, propagating through until leaving at $z \approx 1.03$ and re-entering the FRW background at a redshift of about 1.09.

One can see the corresponding changes in distortion for this beam plotted in Figure 6.21. The jumps in redshift are the same, but the distortion of the beam increases whenever it passes through a vacuum region. This must be the case since, when measured in the shadow plane, the first lens has an orientation of -47° while the second is at -38° , so each reinforces the impact that the other has on the shape of the beam. The contour lines given in Figure 6.22 indicate how the orientation of the distortion experienced by the beams used to probe the patch of “sky” depicted in Figures 6.18 and 6.19 is distributed.

Aside from the final distribution of orientations, it is also interesting to look at the

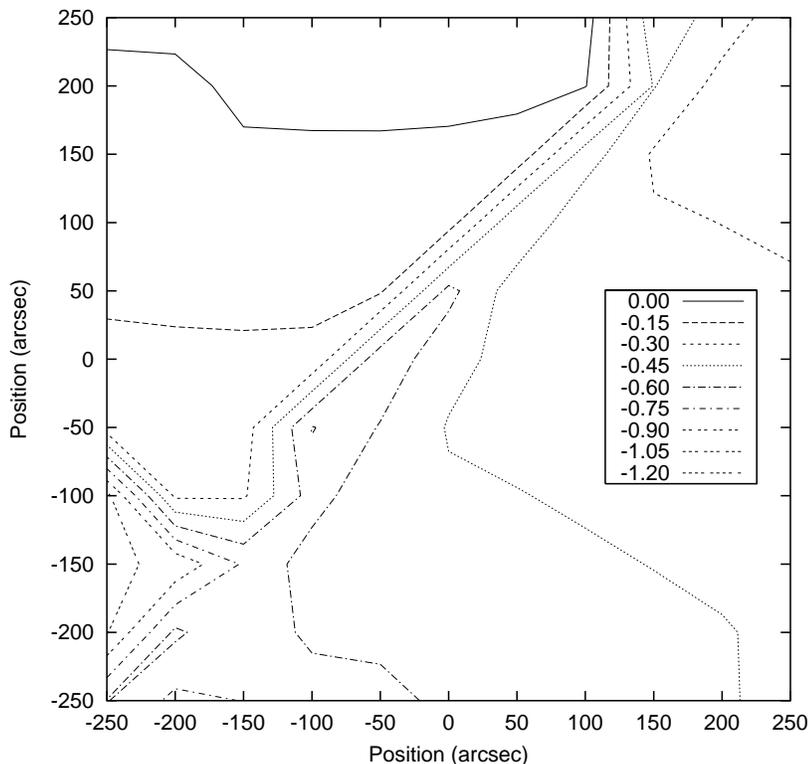


Figure 6.22: Contour lines depicting the variation in the orientation of the beam distortion over the patch of “sky” shown in Figure 6.19.

evolution of the orientation for a single beam as it passes through an RSC model. For the sake of clarity, a beam propagating past two lenses without hitting their cores was chosen.⁸ The boundary of the first lens was reached at a redshift of about 0.43 and it had an orientation of about 0.107 radians, while the second was oriented at an angle of about 0.572 radians and the vacuum surrounding it was first entered by the beam at a redshift of about 0.68. The results are plotted in Figure 6.23. Note that while the affine parameter scale was manipulated such that the propagation in each region could be suitably ordered in a continuous manner, in reality, the affine parameter used for the vacuum propagation was not the same as that used inside the FRW regions. Further, bear in mind that redshift intervals do not scale in exactly the same way as the affine parameter intervals do so the process of patching the evolution in the different regimes

⁸This was from a different RSC model than that used to generate the previous graphs since all beams followed in that case hit the core of the first lens they encountered.

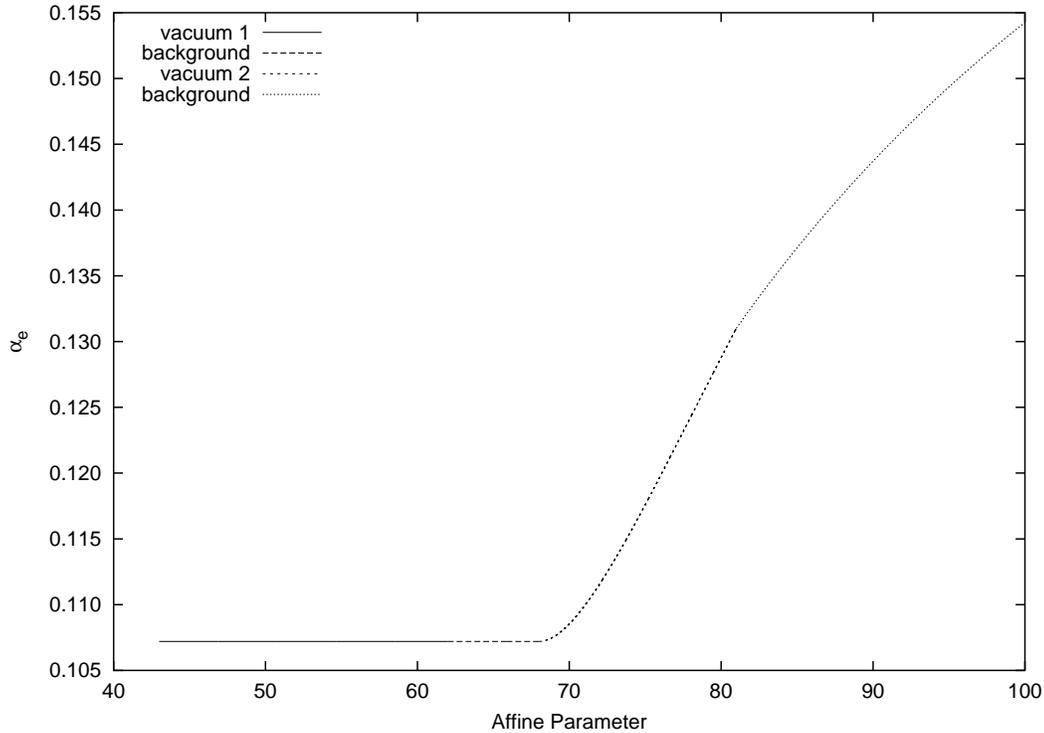


Figure 6.23: The variation in the orientation of the distortion experienced by a beam as it passes through the vacuum of two lenses. The affine parameter making up the horizontal axis has been scaled so that the length of each of the four intervals shown is proportional to the redshift range it covers.

together, while useful, is not perfect.

Once the beam enters the vacuum of the first lens at an affine parameter value of about 43 in the graph, the orientation of that lens is immediately imposed on it as expected. This orientation is maintained for the duration of the propagation through the vacuum, as well as for the subsequent propagation through the FRW background. The vacuum of the second lens is entered at an affine parameter value of about 68, at which point the mass causes the orientation of the distortion to start rotating towards it. Since the angular separation between the two lenses as measured in the shadow plane is less than $\pi/2$, and with the orientation of the second lens being larger than that of the first, it stands to reason that the value of α_e must increase here. The orientation continues to change once the vacuum propagation is complete and the beam re-enters the FRW background at an affine parameter value of about 81. However, the rate of

change decreases as it must since the rotation of the beam is no longer being driven. In other words, the first lens starts the beam shearing, giving it an elliptical cross-section with an orientation that is maintained until it experiences the tidal force of the second lens. This then changes the shear rate and sets the ellipse rotating, a process which slows but continues even after the Weyl driving stops. If the second lens is instead put at an orientation of 2.784 or -0.358 radians,⁹ the same distance away from the first one but in the opposite direction, then the results are the same but the variation is in the opposite sense, with the value of α_e starting to decrease as soon as the second vacuum is entered. This decrease is also seen as expected if the orientation of the second lens is set to a positive quantity slightly smaller than that of the first.

6.3 Comments On Thick Versus Thin Lenses

When investigating gravitational lensing, there is no question that the most relativistically correct approach is to treat matter distributions as the full multidimensional objects that they are. As was discussed earlier in Section 1.3.1, using the thin lens approximation, which considers only the projection of these objects onto a plane lying between the observer and source of light, limits the analysis to working in the limit of a weak gravitational field causing small deflection angles, and requires the assumption of an angular size distance relation a priori. We have already discussed how the thick and thin lens setups are fundamentally different constructions, with beams passing through and experiencing mass distributions in a different way in each case, and have already seen how there can be variations in their computed effect as a result. Nonetheless, it is often argued that when dealing with typical observed data, the thin lens approximation is sufficient even when working in the strong lensing regime (see, for instance, [Kochaneck et al., 2004](#)).

Of course there are certainly a number of situations where the use of a single thin

⁹These are equivalent due to the degeneracy that arises with the restriction that $-\pi/2 \leq \alpha_e \leq \pi/2$.

lens is not appropriate. The behaviour of a beam of light encountering two or more mass distributions that are well spaced in redshift, a significant factor for high redshift sources, is not well described using just one mass sheet (Wambsganss, 2004). Even if the line of sight separation of the multiple distributions is not very large, a single lens plane will not allow the rotation of images to be looked at. The linearized field equations adopted for the weak field limit also cannot describe the giant arcs that are sometimes observed, and the Born approximation used in this limit breaks down if the extent of the mass distribution responsible for the lensing is not much smaller than the lens-source and lens-observer distances (see, for instance, Schneider, 2003). This latter point is problematic when one wishes to consider the impact of large scale structure, such as to compute the cosmic shear, a topic of current interest in cosmology as it is used to map the distribution of dark matter.

In situations such as these, the usual approach is to use multiple thin lens planes. For instance, when trying to model the impact of large scale structure, n-body simulations are used to construct a model of the structure, then this structure is divided up and projected onto a number of different planes lying between the source and observer (see Bartelmann, 2003, for a description). While this process may seem to be the obvious thing to do, computing the properties of light paths that are altered by the presence of multiple lenses is not a straightforward procedure. Unlike classical lenses, the focal length of a gravitational lens is not unique, but varies depending on the impact parameter of the light path considered. It should not come as a surprise that even in theory, using multiple thin lenses does not necessarily produce the same result as a full thick lens calculation would, and is not as generally applicable.¹⁰

Ultimately, the question on whether to use one or many thin lenses or a proper thick lens should be addressed on a problem by problem basis as the answer depends on the

¹⁰For instance, multiple lens plane theory requires multiple lenses to be far enough apart for the beam propagation to reach an asymptotic “straight line” regime between them (Schneider et al., 1992), which is not the case for thick lenses.

type of information one wishes to extract from the system.¹¹ If one simply wants a rough idea of the impact of a single, small, slowly evolving lens system on a background source, then a single thin lens is likely appropriate. As the number of lenses gets larger, the size of the lenses increases, the lenses acquire substructure, the lenses evolve more rapidly, or if one is interested in greater levels of detail, then the projection of the matter distribution onto one or even more planes can become increasingly restrictive. Certainly, any work that plans to use lensing observations to test and constrain general relativity should take particular care to consider the impact of the assumptions that are implicit when carrying out an analysis based on mass sheets.

Even if one is fairly confident that the information one is interested in can be obtained using a single thin lens or a set of them, it does not mean that carrying out a full thick lens calculation is futile. The RSC model and associated lensing code is one example that demonstrates the non-linear form of the field equations does not have to be impractical or unreasonably cumbersome to work with numerically. While the linearization of these equations may make them easier to work with,¹² physics is full of examples where the most interesting behaviour of a system arises from the non-linearities that it contains. Always starting automatically with a linearized set of the equations can inadvertently lead to something significant being overlooked.

6.4 Directions For Future Work

There are many aspects of the RSC model that provide the opportunity for further investigation and development.

A more careful quantitative analysis can be performed on the SC packings which are used to generate the RSC model in order to get a better measure of the structure they

¹¹This should be obvious, though one sometimes gets the impression that the thin lens approximation is just assumed to be relevant.

¹²Though some attempts at working within a linearized theory but trying to correct for its limitations can lead one to wonder how appropriate the linearization was in the first place.

contain, and to see what impact the choice of different spatial curvatures and angular extents has on the packing process. This analysis, which can be based on multi-point correlation function or multifractal methods, could then be applied to a larger RSC model to reveal how the recursive construction affects the structure. This in turn might indicate ways in which the construction paradigm can be adapted to provide structure that is more closely associated with particular sets of observational data or the results of specific n-body simulations.

In addition to examining the placement of objects inside the model, it would also be a good idea to allow more flexibility in the way that they can evolve. As described in Section 3.3.4, the cores of lenses are currently being restricted to experiencing an expansion after a possible delay following the initial “bang” of the background universe. Consequently, it is quite possible for the delay chosen when a beam first encounters a particular lens to be too short to allow the radius of its core to be defined when the lens is encountered by a subsequent beam whose trajectory is different enough from the initial one to cause it to hit the lens boundary significantly sooner.¹³ This has at times proven to be an issue in practice when considering an array of beams being sent through a region of the model. As a first step, it may make sense to lift the restriction that the cores not be allowed to turn around since allowing them to recollapse, or even to expand and contract in a cyclic fashion, will mean that the size of a core can always be defined, even if it is only a point mass. Further, while it will require more computations to be carried out numerically instead of analytically, allowing non-zero values of Λ to be used to describe the evolution of the structure will also add flexibility. Ultimately though, it may be worth lifting the restriction that the cores be constrained to evolve as regions of a dynamic FRW universe. This does introduce complications into the model since care would need to be taken to ensure, for instance, that as the universe is evolved backwards,

¹³That is, later in the evolution of the lens since the beams are propagated backwards through the model.

the boundaries of such holes do not “pass through” their cores in the case where these are essentially static mass distributions. However, removing this constraint will allow for a broader range of structure to be modelled and make it easier to tune the structure to behave in a particular way, rather than have its evolution be determined in a random fashion.

One can also look into generalizing the coordinate system used to describe the propagation of the beam and evolution of the model in vacuum regions. As was explained earlier, the present form, using radially infalling beams as a means of synchronizing observers on the two sides of the vacuum, cannot be used to track an ingoing beam travelling in a purely radial fashion. Further, it can potentially lead to numerical difficulties when trying to follow one with a trajectory that is very nearly radial since the almost constant value of ν for such a beam would render it ineffective for use as an affine parameter. Allowing timing beams that travel radially in each direction, similar to the dual-null Kruskal-Szekeres form of the Schwarzschild metric (see for instance, [Misner et al., 1973](#)), is one possible solution, though the details of keeping the system synchronized when switching from one set of beams to the other remain to be worked out.

On a more technical side, it may be worth going through the code and eliminating all instances where a linear step is used to advance the system being integrated. While this simplification has not proven to be a problem thus far, it is possible that it can start to provide a limitation for more detailed analyses. Since many of the linear routines are carried out rather infrequently overall, using a more sophisticated method should not have a significant performance impact.

Once an RSC model has been constructed, it should be quite straightforward to parallelize the code in order to allow multiple beams to be propagated through it at once, something that can be of use when considering a large grid of beams with small inter-beam separations. Since the trajectory of each beam is independent of any other, dividing the work load over a series of processors should be very straightforward to accomplish. Each

would simply need access to the packings and instructions necessary for it to construct the parts of the model that are relevant to the beam it is propagating. The process becomes less simple to do while the code is still allowed to randomly “discover” additional structure in the model, as care must then be taken in order to avoid two or more processors simultaneously discovering different structure in the same location. While this type of issue commonly arises when trying to parallelize code, the communication overhead it would add suggests that parallelization would not yield large benefits in the initial stages of model construction while the “known structure” is increasing rapidly, unless the code is modified to add the structure in a deterministic fashion. Alternatively, one might choose to separate the process of model generation from the beam propagation entirely.

Even without implementing any modifications to the code, there are a number of investigations that can be carried out with the capabilities it already has. One can propagate beams through a number of models generated with the same background cosmology and model construction parameters, in order to find a statistical measure of some lensing effect their structure produces. These input parameters can then be varied to see what impact, if any, this has on the lensing. It is also possible to bypass the packing algorithm and construct a “packing” containing sphere positions and sizes that are determined by observation or n-body simulation, perhaps representing a cluster of galaxies, then using the code to investigate the lensing of a background source, such as the CMB, that results. Being able to quantify the amount of distortion that beams from this source typically experience as they pass through the universe can have interesting implications when interpreting the extent of causally connected regions in the observed background.

Beyond its value as a tool for gravitational lensing research, the RSC model can also be used to examine the more general problem of trying to determine cosmological quantities by making measurements in a universe containing inhomogeneities. A scientific measurement is typically valid only over some range of scales, and implies an averaging

of properties at scales smaller than this range. One can imagine that astrophysical measurements, which can involve length scales that are non-trivial fractions of the Hubble scale, might be particularly sensitive to the averaging process. Due to the non-linearity of the gravitational field equations, averaging over inhomogeneities may cause the interpretation of measurements used to establish observational and dynamical relations to differ significantly from what would be understood when considered from within the framework of a uniform FRW model (see [Ellis and Buchert, 2005](#), for a recent review and references). This smoothing effect can produce what looks like an extra curvature term and/or contribution to the energy-momentum tensor when calculating the field equations. Using an inhomogeneous Lemaître-Tolman-Bondi model, [Moffat \(2005\)](#) showed that by accounting for the effect of averaging when trying to measure the deceleration parameter, it is possible to find a region of accelerated expansion without a negative pressure dark energy or cosmological constant. By following the detailed geodesic trajectories of beams through the well defined structure contained in an RSC model, the software developed can already be used to investigate the variation in the angular size-redshift relation in a more comprehensive way than was possible before. It would be a natural extension to add to the measurements that can be made by observers placed within the model, taking advantage of the ease with which multiple observations can be made. This would enhance the tool, providing a computational ability that may shed light on the implications of averaging on the measurement of cosmological quantities and contribute to the current debate over this issue.

6.5 Concluding Remarks

A tool has been developed to assemble and explore a Recursive Swiss Cheese cosmological model. Though a rather simple construct, the RSC model is an exact solution to Einstein's field equations that allows significant density inhomogeneities to be placed within

an evolving universe in a relativistically consistent way. It is built by making recursive use of Swiss Cheese sphere packings that are created in a random yet efficient process for spaces of arbitrary constant curvature. It is stored in a compact and extensible form that makes it suitable for use with a wide range of computing hardware.

When used in conjunction with the optical scalar equations, the RSC model allows for the investigation of gravitational lensing in a manner that is quite different from the ray tracing techniques commonly used. It provides a powerful means of examining lensing with an approach that is automatically flux conserving, and that does not require an assumed angular size distance relationship to apply over the patch of “sky” being observed. These advantages are particularly significant when the lensing effect becomes stronger and the small angle approximation that is often relied upon breaks down. The fact that the proper null geodesics are used to follow the beam trajectories past and through the thick lenses present in the model implies that this technique also provides an unambiguous way of considering the impact of multiple lenses on a single beam, something that is not as straightforward to accomplish when considering thin lenses and the instantaneous bending of beams.

The parameters describing the structure in the RSC model are chosen in a way that allows this structure to experience significant evolution. By allowing the model’s structure to evolve as beams are propagated through it, large or rapidly evolving lenses can be treated properly. The coordinate system selected to map the vacuum regions of the model is numerically stable across Schwarzschild singularities, and led to the development of a means of tracking spatially separate events in a synchronized fashion when following beams through these regions. As a result, the model can be observed in a rational way from multiple locations within it, with the appearance of the lenses it contains being consistent over any time scale, even when they are observed from a variety of different directions.

The form of the optical scalar equations that was developed makes them stable

through the caustics beams can experience as they propagate through the RSC model. When used in conjunction with the parallel propagated basis that was established to describe the shadow plane of the beams, these equations permit the area, shape, and orientation of a distorted beam to be tracked in a coherent manner over encounters with multiple lenses, through and between all regions of the model. As a result, it is possible to directly compare the quantities describing the evolution of a large set of propagated beams in a consistent fashion.

The software produced over the course of this work was designed to run efficiently, and given the ability to track and control various sources of numerical error. The source code is thoroughly commented and will be made freely available for those who are interested in using it to further their own research. While the examination of the structure generated by the packing algorithm and the gravitational lensing effects it produces are interesting problems in and of themselves, the applicability of the methods developed go beyond this. The intention is for the code to be modified and adapted as necessary, being combined with input derived from various other sources and contrasted with results generated by other means in order to help lend insight into a broader range of gravitational lensing and cosmological research.

Appendix A

RW Volume Computation

From the Robertson-Walker line element (1.7), the differential volume element is given by

$$dV = R^3 \mathcal{S}_K^2(\omega) \sin(\theta) d\omega d\theta d\phi ,$$

where strictly there should be a negative sign out front due to the signature chosen. Hence, a region of the FRW model defined out to some radius ω_{\max} has a volume given by

$$\begin{aligned} V &= \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \int_{\omega=0}^{\omega_{\max}} R^3 \mathcal{S}_K^2(\omega) \sin(\theta) d\omega d\theta d\phi \\ \therefore V &= 4\pi R^3 \int_{\omega=0}^{\omega_{\max}} \mathcal{S}_K^2(\omega) d\omega . \end{aligned} \quad (\text{A.1})$$

The comoving volume can be computed by dropping the scale factor R and using (1.8) to replace $\mathcal{S}_K(\omega)$, resulting in

$$V = \begin{cases} 2\pi \left[\sinh(\omega_{\max}) \cosh(\omega_{\max}) - \omega_{\max} \right] & : \quad K = -1 \\ \frac{4\pi}{3} \omega_{\max}^3 & : \quad K = 0 \quad , \\ 2\pi \left[\omega_{\max} - \sin(\omega_{\max}) \cos(\omega_{\max}) \right] & : \quad K = +1 \end{cases} \quad (\text{A.2})$$

with the flat case having the familiar Euclidian form as is to be expected.

Appendix B

Computing the Cylindrical Mass

As explained in Section 1.3.1 when discussing thin lenses, a ray of light passing through a spherical mass distribution has a deflection angle that is determined by the mass contained within a cylinder of radius h , the distance of closest approach that the ray has to the centre of the projected lens. In order to derive an expression for this mass applicable to the uniform density spherical cores at the centre of the lenses in an RSC model, consider the ball of uniform density ρ and radius a shown in Figure B.1. Inscribe within it, a thin cylindrical shell of radius r and thickness dr as shown in the diagram. If this shell has height $2z$, then its mass, given by the product of its volume and the density of

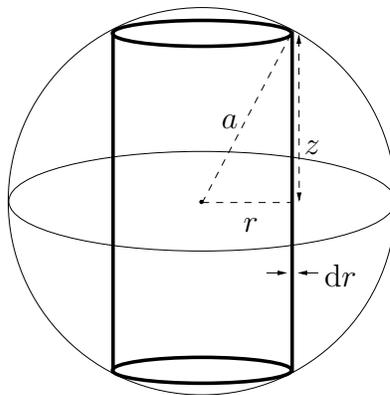


Figure B.1: Quantities used for cylindrical mass derivation.

the sphere, is

$$M_{\text{shell}} = \rho 2\pi r(2z) dr ,$$

so

$$M_{\text{shell}}(r) = 4\pi r \rho \sqrt{a^2 - r^2} dr . \quad (\text{B.1})$$

Hence, the total mass enclosed within a similar cylinder of radius h can be found by integrating the mass given by (B.1) for all such cylinders with $r \leq h$,

$$\begin{aligned} M_c &= \int_0^h 4\pi r \rho \sqrt{a^2 - r^2} dr \\ &= -\frac{4\pi\rho}{3} (a^2 - r^2)^{\frac{3}{2}} \Big|_0^h \\ \therefore M_c(h) &= \frac{4\pi\rho}{3} \left[a^3 - (a^2 - h^2)^{\frac{3}{2}} \right] . \end{aligned} \quad (\text{B.2})$$

This expression can be used to compute the cylindrical mass that affects a ray with an impact parameter h , as it propagates past a thin lens representing an RSC core of density ρ and radius a .

Appendix C

Coordinate Systems

C.1 Coordinate Conversions

When working with the RSC model, it is frequently necessary to convert between various coordinate systems. If one can write the components of a point expressed in one coordinate system in terms of another, i.e.

$$\tilde{x}^a = \tilde{x}^a(x^b) ,$$

then the components of a rank-1 tensorial quantity in one system can be expressed in terms of the other via the relationships

$$\tilde{v}^a = v^b \frac{\partial \tilde{x}^a}{\partial x^b} \tag{C.1a}$$

$$\tilde{v}_a = v_b \frac{\partial \tilde{x}^b}{\partial x^a} . \tag{C.1b}$$

These are applied below to a couple of coordinate systems of interest.

C.1.1 Angular and Cartesian Coordinates

In order to generate some of the projections used in Section 2.1.2, or to orient the system for propagation in the FRW regions as described in Section 3.1.1, it is necessary

to convert from the comoving angular RW coordinates used to a Cartesian system. In the case of the packings generated in flat space, this is a simple task as the (ω, θ, ϕ) coordinates can be converted to and from the usual (x, y, z) Cartesian coordinates via:

$$\begin{aligned}
 x &= \omega \sin(\theta) \cos(\phi) & \omega &= \sqrt{x^2 + y^2 + z^2} \\
 y &= \omega \sin(\theta) \sin(\phi) & \theta &= \cos^{-1} \left(\frac{z}{\omega} \right) \\
 z &= \omega \cos(\theta) & \phi &= \tan^{-1} \left(\frac{y}{x} \right)
 \end{aligned} \tag{C.2}$$

recalling that in this situation, ω is a linear (though unit-less) radial measure. When the curvature of the packings is non-zero, the curved 3-space can be embedded into a flat 4-dimensional Euclidian space spanned by the four mutually orthogonal axes x , y , z , and w . In the case of positive curvature, the conversions are given by

$$\begin{aligned}
 x &= r \sin(\omega) \sin(\theta) \cos(\phi) & \omega &= \cos^{-1} \left(\frac{w}{r} \right) \\
 y &= r \sin(\omega) \sin(\theta) \sin(\phi) & \theta &= \cos^{-1} \left(\frac{z}{r \sin(\omega)} \right) \\
 z &= r \sin(\omega) \cos(\theta) & \phi &= \tan^{-1} \left(\frac{y}{x} \right) \\
 w &= r \cos(\omega) & r &= \sqrt{x^2 + y^2 + z^2 + w^2}
 \end{aligned} \tag{C.3}$$

which is very similar to the case of negative curvature where

$$\begin{aligned}
 x &= r \sinh(\omega) \sin(\theta) \cos(\phi) & \omega &= \cosh^{-1} \left(\frac{w}{r} \right) \\
 y &= r \sinh(\omega) \sin(\theta) \sin(\phi) & \theta &= \cos^{-1} \left(\frac{z}{r \sinh(\omega)} \right) \\
 z &= r \sinh(\omega) \cos(\theta) & \phi &= \tan^{-1} \left(\frac{y}{x} \right) \\
 w &= r \cosh(\omega) & r &= \sqrt{w^2 - x^2 - y^2 - z^2}
 \end{aligned} \tag{C.4}$$

Note that in both curved cases, the linear radial variable r has been used for the conversions to be formally correct. In the case of the packings, this is simply set to unity as it is the scale factor R that gives a physical dimension to the coordinates, just as it does in the flat case.

Using the above expressions in equations (C.1), with x^a referring to angular coordinates and \tilde{x}^a representing Cartesian coordinates, the conversions of rank-1 tensors are given below for each curvature value. Note that the indices on v and \tilde{v} in this section run from 1 to 4 (1 to 3 in the flat case), covering the non-indexed variables (x, y, z, w) or $(\omega, \theta, \phi, r)$ as appropriate.

K = 0

Angular to Cartesian:

$$\begin{aligned}\tilde{v}^1 &= v^1 \sin(\theta) \cos(\phi) + v^2 \omega \cos(\theta) \cos(\phi) - v^3 \omega \sin(\theta) \sin(\phi) \\ \tilde{v}^2 &= v^1 \sin(\theta) \sin(\phi) + v^2 \omega \cos(\theta) \sin(\phi) + v^3 \omega \sin(\theta) \cos(\phi) \\ \tilde{v}^3 &= v^1 \cos(\theta) - v^2 \omega \sin(\theta)\end{aligned}\tag{C.5}$$

$$\begin{aligned}\tilde{v}_1 &= v_1 \sin(\theta) \cos(\phi) + \frac{v_2}{\omega} \cos(\theta) \cos(\phi) - \frac{v_3 \sin(\phi)}{\omega \sin(\theta)} \\ \tilde{v}_2 &= v_1 \sin(\theta) \sin(\phi) + \frac{v_2}{\omega} \cos(\theta) \sin(\phi) + \frac{v_3 \cos(\phi)}{\omega \sin(\theta)} \\ \tilde{v}_3 &= v_1 \cos(\theta) - \frac{v_2}{\omega} \sin(\theta)\end{aligned}$$

Cartesian to Angular:

$$\begin{aligned}v^1 &= \frac{1}{\omega} (\tilde{v}^1 x + \tilde{v}^2 y + \tilde{v}^3 z) & v_1 &= \frac{1}{\omega} (\tilde{v}_1 x + \tilde{v}_2 y + \tilde{v}_3 z) \\ v^2 &= \frac{v^1 \frac{z}{\omega} - \tilde{v}^3}{\sqrt{\omega^2 - z^2}} & v_2 &= \frac{v_1 z \omega - \tilde{v}_3 \omega^2}{\sqrt{x^2 + y^2}} \\ v^3 &= \frac{\tilde{v}^2 x - \tilde{v}^1 y}{x^2 + y^2} & v_3 &= \tilde{v}_2 x - \tilde{v}_1 y\end{aligned}\tag{C.6}$$

K = +1

When applying these conversions to the embeddings in the RSC model, note that any vector of interest must have $v^4 = v_4 = 0$ in order to remain within the curved space under consideration.

Angular to Cartesian:

$$\begin{aligned}
\tilde{v}^1 &= v^1 r \cos(\omega) \sin(\theta) \cos(\phi) + v^2 r \sin(\omega) \cos(\theta) \cos(\phi) - v^3 r \sin(\omega) \sin(\theta) \sin(\phi) + v^4 \frac{x}{r} \\
\tilde{v}^2 &= v^1 r \cos(\omega) \sin(\theta) \sin(\phi) + v^2 r \sin(\omega) \cos(\theta) \sin(\phi) + v^3 r \sin(\omega) \sin(\theta) \cos(\phi) + v^4 \frac{y}{r} \\
\tilde{v}^3 &= v^1 r \cos(\omega) \cos(\theta) - v^2 r \sin(\omega) \sin(\theta) + v^4 \frac{z}{r} \\
\tilde{v}^4 &= -v^1 r \sin(\omega) + v^4 \frac{w}{r}
\end{aligned} \tag{C.7}$$

$$\begin{aligned}
\tilde{v}_1 &= \frac{v_1}{r} \cos(\omega) \sin(\theta) \cos(\phi) + \frac{v_2 \cos(\theta) \cos(\phi)}{r \sin(\omega)} - \frac{v_3 \sin(\phi)}{r \sin(\omega) \sin(\theta)} + v_4 \frac{x}{r} \\
\tilde{v}_2 &= \frac{v_1}{r} \cos(\omega) \sin(\theta) \sin(\phi) + \frac{v_2 \cos(\theta) \sin(\phi)}{r \sin(\omega)} + \frac{v_3 \cos(\phi)}{r \sin(\omega) \sin(\theta)} + v_4 \frac{y}{r} \\
\tilde{v}_3 &= \frac{v_1}{r} \cos(\omega) \cos(\theta) - \frac{v_2 \sin(\theta)}{r \sin(\omega)} + v_4 \frac{z}{r} \\
\tilde{v}_4 &= -\frac{v_1}{r} \sin(\omega) + v_4 \frac{w}{r}
\end{aligned}$$

Cartesian to Angular:

$$\begin{aligned}
v^1 &= \frac{1}{\sqrt{r^2 - w^2}} \left(\frac{v^4 w}{r} - \tilde{v}^4 \right) & v_1 &= \frac{v_4 r w - \tilde{v}_4 r^2}{\sqrt{r^2 - w^2}} \\
v^2 &= \frac{1}{\sqrt{x^2 + y^2}} \left(\frac{v^4 r z - \tilde{v}^4 z w}{r^2 - w^2} - \tilde{v}^3 \right) & v_2 &= \frac{z(\tilde{v}_1 x + \tilde{v}_2 y)}{\sqrt{x^2 + y^2}} - \tilde{v}_3 \sqrt{x^2 + y^2} \\
v^3 &= \frac{\tilde{v}^2 x - \tilde{v}^1 y}{x^2 + y^2} & v_3 &= \tilde{v}_2 x - \tilde{v}_1 y \\
v^4 &= \frac{1}{r} (\tilde{v}^1 x + \tilde{v}^2 y + \tilde{v}^3 z + \tilde{v}^4 w) & v_4 &= \frac{1}{r} (\tilde{v}_1 x + \tilde{v}_2 y + \tilde{v}_3 z + \tilde{v}_4 w)
\end{aligned} \tag{C.8}$$

K = -1

As in the positively curved case, any vector in the RSC model to which these conversions may be applied must have $v^4 = v_4 = 0$ in order to remain within the curved space under consideration.

Angular to Cartesian:

$$\begin{aligned}
\tilde{v}^1 &= v^1 r \cosh(\omega) \sin(\theta) \cos(\phi) + v^2 r \sinh(\omega) \cos(\theta) \cos(\phi) - v^3 r \sinh(\omega) \sin(\theta) \sin(\phi) + v^4 \frac{x}{r} \\
\tilde{v}^2 &= v^1 r \cosh(\omega) \sin(\theta) \sin(\phi) + v^2 r \sinh(\omega) \cos(\theta) \sin(\phi) + v^3 r \sinh(\omega) \sin(\theta) \cos(\phi) + v^4 \frac{y}{r} \\
\tilde{v}^3 &= v^1 r \cosh(\omega) \cos(\theta) - v^2 r \sinh(\omega) \sin(\theta) + v^4 \frac{z}{r} \\
\tilde{v}^4 &= v^1 r \sinh(\omega) + v^4 \frac{w}{r}
\end{aligned} \tag{C.9}$$

$$\begin{aligned}
\tilde{v}_1 &= \frac{v_1}{r} \cosh(\omega) \sin(\theta) \cos(\phi) + \frac{v_2 \cos(\theta) \cos(\phi)}{r \sinh(\omega)} - \frac{v_3 \sin(\phi)}{r \sinh(\omega) \sin(\theta)} - v_4 \frac{x}{r} \\
\tilde{v}_2 &= \frac{v_1}{r} \cosh(\omega) \sin(\theta) \sin(\phi) + \frac{v_2 \cos(\theta) \sin(\phi)}{r \sinh(\omega)} + \frac{v_3 \cos(\phi)}{r \sinh(\omega) \sin(\theta)} - v_4 \frac{y}{r} \\
\tilde{v}_3 &= \frac{v_1}{r} \cosh(\omega) \cos(\theta) - \frac{v_2 \sin(\theta)}{r \sinh(\omega)} - v_4 \frac{z}{r} \\
\tilde{v}_4 &= -\frac{v_1}{r} \sinh(\omega) + v_4 \frac{w}{r}
\end{aligned}$$

Cartesian to Angular:

$$\begin{aligned}
v^1 &= \frac{\tilde{v}^4 r - v^4 w}{r \sqrt{w^2 - r^2}} & v_1 &= \frac{v_4 r w - \tilde{v}_4 r^2}{\sqrt{w^2 - r^2}} \\
v^2 &= \frac{1}{\sqrt{x^2 + y^2}} \left(\frac{\tilde{v}^4 z w - v^4 r z}{w^2 - r^2} - \tilde{v}^3 \right) & v_2 &= \frac{z(\tilde{v}_1 x + \tilde{v}_2 y)}{\sqrt{x^2 + y^2}} - \tilde{v}_3 \sqrt{x^2 + y^2} \\
v^3 &= \frac{\tilde{v}^2 x - \tilde{v}^1 y}{x^2 + y^2} & v_3 &= \tilde{v}_2 x - \tilde{v}_1 y \\
v^4 &= \frac{1}{r} (\tilde{v}^4 w - \tilde{v}^1 x - \tilde{v}^2 y - \tilde{v}^3 z) & v_4 &= \frac{1}{r} (\tilde{v}_1 x + \tilde{v}_2 y + \tilde{v}_3 z + \tilde{v}_4 w)
\end{aligned} \tag{C.10}$$

C.1.2 RW and Schwarzschild Coordinates

When following a beam from a matter-filled region to a vacuum and back, it is necessary to convert the components of its position, along with those of those of the various vectors being propagated along with it, between the RW and Schwarzschild coordinate systems used. Consider first the transition from propagation in an FRW region to a vacuum. Comparing the RW metric (1.7) to that of the exterior Schwarzschild solution (1.2), it

is apparent that at the point of contact between the two regions, the following relations must hold:

$$r = R(cT)\mathcal{S}_K(\omega) \quad (\text{C.11a})$$

$$\theta_S = \theta_R \quad (\text{C.11b})$$

$$\phi_S = \phi_R, \quad (\text{C.11c})$$

where the subscript S and R refer to a Schwarzschild and RW coordinate respectively. There is no simple way to write the Schwarzschild time in terms of RW parameters, but since its absolute value is arbitrary, this is not significant at this point.

Using equation (C.1a), it is easy to show that

$$v_S^1 = v_R^0 \mathcal{S}_K(\omega) \frac{\partial R}{\partial(cT)} + v_R^1 R \frac{\partial \mathcal{S}_K(\omega)}{\partial \omega} \quad (\text{C.12a})$$

$$v_S^2 = v_R^2 \quad (\text{C.12b})$$

$$v_S^3 = v_R^3. \quad (\text{C.12c})$$

Due to the lack of an expression for the conversion of the temporal components, the same procedure cannot be followed to find a similar equation for v_S^0 . However, the coordinate conversion cannot change the inner product of the vector with itself, so if

$$v^a v_a = g_{ab} v^a v^b = N$$

in both coordinate systems, then expanding the metric and isolating for the relevant term,

$$v_S^0 = \pm \sqrt{\frac{N + \frac{v_S^1 v_S^1}{f(r)} + r^2 (v_S^2 v_S^2 + v_S^3 v_S^3)}{f(r)}}, \quad (\text{C.13})$$

where use has been made of the fact that $\theta = \pi/2$ whenever the coordinate conversion needs to be made due to the propagation orientation employed. Technically, this still leaves a sign uncertainty, but one that can be resolved in practice since it is either k^a , the direction vector of the beam, that needs to be converted, or a vector that lies in the

shadow plane. As explained in Section 3.3, matching k^a across the boundary does not require the formal transformation to be carried out so there is no ambiguity there. The fact that k^a must be normal to the shadow plane means that its inner product with any vector lying in this plane must be zero, a constraint that can be used to set the sign of v_S^0 accordingly when dealing with such vectors.

When returning to an FRW region from a vacuum, (C.11a) can be inverted to yield

$$\omega = \mathcal{S}_K^{-1} \left(\frac{r}{R(cT)} \right),$$

so the spatial location of the beam being followed can be determined. However, the lack of an expression for the RW time in terms of Schwarzschild parameters means that the conversion of the components of a vector from Schwarzschild to RW coordinates is a more complicated process. One could potentially derive the relevant equations by inverting (C.12) and (C.13), but as explained in Section 4.5.2, it turns out that these conversions are not necessary for the propagation of the beam.

C.2 Alternate Coordinates for Vacuum Regions

When developing the (ν, μ, θ, ϕ) vacuum coordinate system described in Section 3.3.3, the original intent was to be able to follow a beam beyond the Schwarzschild coordinate singularity for an arbitrary value of the cosmological constant Λ . This turned out to be less than straightforward to implement in practice. A scan through a reference, such as Misner et al. (1973) or Stephani (1994), will quickly reveal a variety of alternate coordinate systems that have been developed to try and address the complications arising with conventional Schwarzschild coordinates.

The Lemaître metric adapts the coordinate system to that of freely falling timelike observers. Using T to represent the proper time of these observers, and R as a label for

each of them,¹ the metric can be written as

$$ds^2 = d(cT)^2 - \frac{2m}{r} dR^2 - r^2 d\Omega^2$$

when $\Lambda = 0$, where

$$r = \left[(R - cT) \frac{3\sqrt{2m}}{2} \right]^{\frac{2}{3}}$$

is now simply a parameter, and $d\Omega$ is the line element on a spherical surface given in (1.4).

Eddington-Finkelstein coordinates introduce an advanced or delayed time, u or v respectively, to produce the metrics

$$ds^2 = f(r) d(cu)^2 + 2 dcu dr - r^2 d\Omega^2$$

and

$$ds^2 = f(r) d(cv)^2 - 2 dcv dr - r^2 d\Omega^2 ,$$

where $f(r)$ is the same as that used in the Schwarzschild metric and defined in (1.3). The transformation is valid regardless of the value of Λ , with the change effectively replacing the Schwarzschild time coordinate by a label describing radially outgoing (advanced) or ingoing (delayed) null observers. These two coordinate systems can be combined into the dual-null Kruskal-Szekeres system, which has a metric of the form

$$ds^2 = \frac{32m^3}{r} e^{-r/2m} dU dV - r^2 d\Omega^2$$

when $\Lambda = 0$, where U and V are functions of u and v respectively, and $r = r(U, V)$ is no longer an independent variable.

None of these coordinate systems has a problem at $r = 2m$ (or equivalently, at $f(r) = 0$ for $\Lambda \neq 0$). In fact, the (ν, μ, θ, ϕ) system derived is in some sense a combination of the Lemaître and delayed Eddington-Finkelstein systems. However, in all cases, even with the cosmological constant set to zero, trying to solve the Euler-Lagrange equations

¹In other words, a freely-falling timelike observer has $dR = 0$.

of motion directly is no simple task, and leaving them expressed as a set of coupled differential equations makes working with them in the RSC model significantly more complicated. It is of course possible to use equations (C.1) to derive expressions for the various quantities of interest from their Schwarzschild form, but this process will typically require Schwarzschild values to be computed, at least as an intermediate step, leading to potential difficulties with the coordinate singularity that led to the investigation of alternate systems for the vacuum to begin with.

Appendix D

Random Point Generation

As discussed in Sections 2.1.1 and 5.1, in order to generate the SC packings with which the RSC model is constructed, a means of choosing points that are randomly distributed within the space being packed is needed. Simply selecting random values of ω , θ , and ϕ does not lead to a suitable random volume distribution of points as they would tend to cluster near the poles of the coordinate system. The Cartesian embedding described in Appendix C.1.1 can be used to get around this problem.

A flat region with $K = 0$ is the most straightforward space to deal with. Imagine a cube of side length $2\omega_{\max}$ surrounding the spherical region of radius ω_{\max} that needs to be packed. Choosing three random values between $-\omega_{\max}$ and ω_{\max} allows a location inside this cube to be specified using Cartesian (x, y, z) coordinates. If this point lies outside the inscribed sphere, it is rejected. Otherwise, equations (C.2) are used to convert it to the appropriate angular coordinate system and it is saved. The process is repeated until a sufficient number of points have been stored, with these now having an equal probability of lying anywhere¹ within the region to be packed as required.

The situation is similar in the case where $K = +1$, though this time equations (C.3) are used for the embedding, and a unit hypersphere defined by $r = 1$ is inscribed within

¹To within floating point tolerance.

a four-dimensional hypercube of side length 2. Sets of three random numbers in the interval $[-1,1]$ are selected to represent x , y , and z coordinates, while values in the range $[0,1]$ are chosen randomly for w , defining points effectively populating the “upper” half of the hypercube. Any points lying outside the hypersphere are rejected, while those within it are projected onto its three-dimensional surface by dividing each coordinate component by r . These points are then converted to angular coordinates, and any found to be lying outside the region defined by the boundary at $\omega = \omega_{\max}$ are also rejected. The remaining points can be used as suitable random locations within the portion of the hypersurface to be packed, and the process is repeated until a sufficient number of positions have been accumulated.

The hyperbolic shape of a $K = -1$ space makes random point selection a bit more complicated for it. If the limiting value of ω_{\max} is small, one can take an approach that is analogous to the $K = +1$ case. That is, using equations (C.4) for the embedding, choose random values of (x, y, z, w) to populate the “upper” half of a four-dimensional hypercube of side length $2 \cosh(\omega_{\max})$, and discard any points with $r > 1$. Scale the coordinate components of any location satisfying this criteria by r to place it onto the unit three-dimensional hyperbolic surface, then convert to angular coordinates and retain any point remaining within the boundary defined by $\omega = \omega_{\max}$. Unfortunately, the high degree of rejection inherent in this method makes it quite impractical to use as ω_{\max} grows. As a result, an alternative technique was employed to generate most of the packings that have $\omega_{\max} = 2\pi$. Under this new scheme, θ and ϕ are chosen in the same way that they are in the $K = 0$ case, from points populating the interior of a unit sphere. However, the value of ω is determined by picking a random location along the length of the radial hyperbolic curve describing the extent of the surface of interest. Considering the curve resulting from the projection of this surface onto the $w - z$ plane when $r = 1$ and $\theta = 0$, the parametric equations of the hyperbola are

$$z = \sinh(\omega) \quad \text{and} \quad w = \cosh(\omega) ,$$

and its arc-length over the range of interest in ω is given by the integral

$$l(\omega) = \int_0^\omega \sqrt{(z')^2 + (w')^2} d\tilde{\omega} = \int_0^\omega \sqrt{\cosh(2\tilde{\omega})} d\tilde{\omega} ,$$

where the prime indicates a derivative with respect to $\tilde{\omega}$. This is an elliptic integral of the second kind which has no general analytic solution, but for larger values of ω , $\cosh(\omega) \approx 0.5e^\omega$, so for the curves of interest in this case,

$$l(\omega) \approx \frac{1}{\sqrt{2}} (e^\omega - 1) .$$

This equation can be inverted so that a randomly selected value of l lying between 0 and $l(\omega_{\max})$ can be converted to a corresponding value of ω , yielding the final component of the random point required. With no rejection involved when choosing the ω coordinate, this method is orders of magnitude quicker than the previously described one, and can be effectively used when ω_{\max} is large and the other becomes impractical.

The random point population schemes described above were used to generate random data sets that were analyzed separately before being used to generate packings. Refer to [Attard \(1997\)](#) for more details on how these random point files are actually used by the packing code.

Before concluding this section, it is worth emphasizing the importance of ensuring that the random number generator being used by the code is indeed valid. Initial efforts at generating packings used `rand()`, the default C random number generator. While this worked properly on some architecture/OS/compiler combinations, others produced results that were decidedly non-random, as illustrated in [Figure D.1](#). Further analysis revealed that the generator used in these cases had a small dynamic range, which lead to the points chosen in the Cartesian embedding space being distributed in widely spaced planes, an effect that translates into the sinusoidal gaps apparent in the two distributions shown. Use of a more appropriate random number generator showed no such structure in either the random points chosen or the packings produced with them.

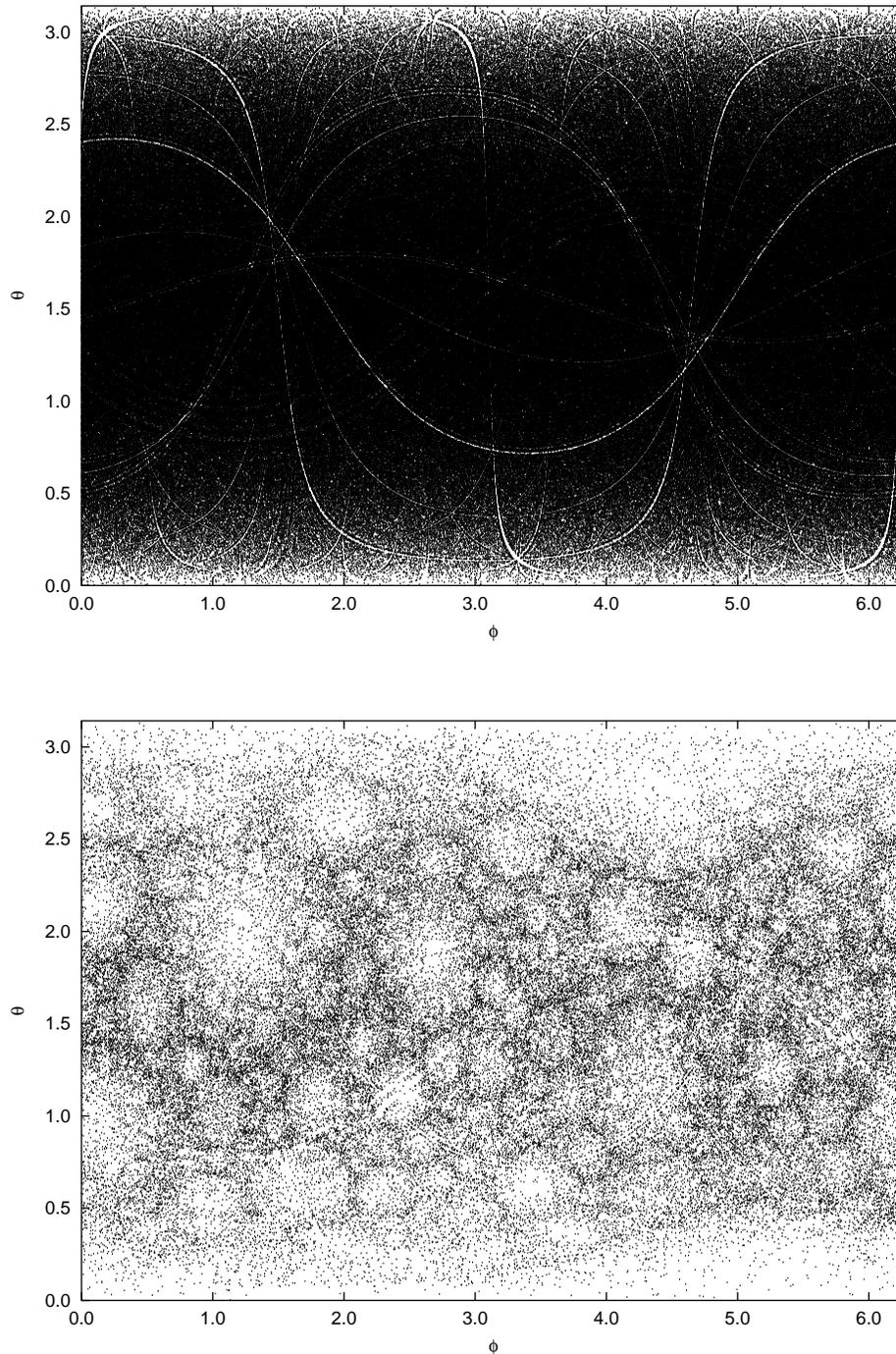


Figure D.1: The upper plot shows the “sky” projection of a million “random” points chosen using an ineffective random number generator to populate a positively curved space with $\omega_{\max} = \pi/10$. The lower one shows the distribution of 77 682 sphere centres from a packing generated using these points.

Bibliography

- Allen Attard. Simulating a Swiss-Cheese Cosmological Model. Master's report, University Of Toronto, September 1997.
- Matthias Bartelmann. Numerical Methods in Gravitational Lensing. arXiv:astro-ph/0304162, April 2003.
- M. Davis and P. J. E. Peebles. A Survey of Galaxy Redshifts. V - The Two-Point Position and Velocity Correlations. *ApJ*, 267:465–482, April 1983.
- C. C. Dyer. Gravitational Lenses and the Inhomogeneous Universe. *Canadian Journal of Physics*, 64:160–164, February 1986.
- C. C. Dyer. The Gravitational Perturbation of the Cosmic Background Radiation by Density Concentrations. *M.N.R.A.S.*, 175:429–447, May 1976.
- C. C. Dyer. Optical Scalars and the Spherical Gravitational Lens. *M.N.R.A.S.*, 180:231–242, July 1977.
- C. C. Dyer and L. M. Oattes. On the Effects of Inhomogeneity on the Luminosity Structure of the Past Null Cone. *ApJ*, 326:50–62, March 1988.
- C. C. Dyer and R. C. Roeder. Toward a Realistic Nebular Gravitational Lens. *ApJ*, 249:290–296, October 1981.
- C. C. Dyer and R. C. Roeder. Distance-Redshift Relations for Universes with Some Intergalactic Medium. *ApJ*, 180:L31–L34, February 1973.
- C. C. Dyer and R. C. Roeder. Observations in Locally Inhomogeneous Cosmological Models. *ApJ*, 189:167–176, April 1974.
- Charles C. Dyer. *Observational Aspects of Locally Inhomogeneous Cosmological Models*. PhD thesis, University Of Toronto, 1973.
- A. Einstein and E. G. Straus. The Influence of the Expansion of Space on the Gravitation Fields Surrounding the Individual Stars. *Reviews of Modern Physics*, 17:120–124, April 1945.
- George F. R. Ellis and Thomas Buchert. The Universe Seen at Different Scales. arXiv:gr-qc/0506106, July 2005.

- A. J. S. Hamilton. Toward Better Ways to Measure the Galaxy Correlation Function. *ApJ*, 417:19–35, November 1993.
- John Francis Harper. *The Luminosity Structure of the Past Null Cone in Inhomogeneous Cosmologies*. PhD thesis, University Of Toronto, 1991.
- H. G. E. Hentschel and I. Procaccia. The Infinite Number of Generalized Dimensions of Fractals and Strange Attractors. *Physica D Nonlinear Phenomena*, 8:435–444, September 1983.
- M. H. Jensen, L. P. Kadanoff, and I. Procaccia. Scaling Structure and Thermodynamics of Strange Sets. *Phys. Rev. A*, 36:1409–1420, August 1987.
- R. Kantowski. Corrections in the Luminosity-Redshift Relations of the Homogeneous Fried-Mann Models. *ApJ*, 155:89–103, January 1969.
- R. Kantowski. Another Interpretation of the Optical Scalars. *Journal of Mathematical Physics*, 9(2):336–338, 1968.
- C. S. Kochanek, P. Schneider, and J. Wambsganss. Gravitational Lensing: Strong, Weak & Micro. In G. Meylan, P. Jetzer, and P. North, editors, *Proceedings of the 33rd Saas-Fee Advanced Course*. Springer-Verlag, 2004.
- S. D. Landy and A. S. Szalay. Bias and Variance of Angular Correlation Functions. *ApJ*, 412:64–71, July 1993.
- Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation*. W. H. Freeman and Company, 1973.
- J. W. Moffat. Late-time Inhomogeneity and Acceleration Without Dark Energy. arXiv:astro-ph/0505326, July 2005.
- Jonas R. Mureika. *Topics in Multifractal Analysis of Two- and Three-Dimensional Structures in Spaces of Constant Curvature*. PhD thesis, University Of Toronto, 2001.
- J. V. Narlikar. *Introduction to Cosmology*. Cambridge University Press, 1993.
- L. M. Oattes. *The Luminosity Function of Standard Candles in an Inhomogeneous Cosmological Model*. PhD thesis, University Of Toronto, 1987.
- P. J. E. Peebles. *Principles of Physical Cosmology*. Princeton University Press, 1993.
- R. Penrose. Structure of Space-Time. In Cecile M. DeWitt and John A. Wheeler, editors, *Battelle Rencontres*, pages 156–171. W. A. Benjamin, Inc, 1968.
- Serge Pineault. *Geometrical Optics in Axially Symmetric Spacetimes*. PhD thesis, University Of Toronto, 1975.
- F. A. E. Pirani. Introduction to Gravitational Radiation Theory. In *Brandeis Lectures on General Relativity*, volume 1, pages 249–373. Prentice-Hall, Inc., 1964.

- William H. Press et al. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, second edition, 1992.
- R. Sachs. Gravitational Waves in General Relativity. VI. The Outgoing Radiation Condition. *Royal Society of London Proceedings Series A*, 264:309–338, November 1961.
- P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. Springer-Verlag, 1992.
- Peter Schneider. Gravitational Lensing as a Probe of Structure. arXiv:astro-ph/0306465, June 2003.
- Hans Stephani. *General Relativity*. Cambridge University Press, second edition, 1994.
- R. C. Tolman. *Relativity, Thermodynamics and Cosmology*. Clarendon Press, 1934.
- Joachim Wambsganss. Gravitational Lensing in a Concordance LCDM Universe: The importance of secondary matter along the line of sight. arXiv:astro-ph/0405147, May 2004.
- S. Weinberg. *Gravitation and Cosmology*. Wiley, 1972.

