3-D Reconstruction from Single Projections, with Applications to Astronomical Images

by

Michael Cormier

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Computer Science

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

A variety of techniques exist for three-dimensional reconstruction when multiple views are available, but less attention has been given to reconstruction when only a single view is available. Such a situation is normal in astronomy, when a galaxy (for example) is so distant that it is impossible to obtain views from significantly different angles. In this thesis I examine the problem of reconstructing the three-dimensional structure of a galaxy from this single viewpoint. I accomplish this by taking advantage of the image formation process, symmetry relationships, and other structural assumptions that may be made about galaxies.

Most galaxies are approximately symmetric in some way. Frequently, this symmetry corresponds to symmetry about an axis of rotation, which allows strong statements to be made about the relationships between luminosity at each point in the galaxy. It is through these relationships that the number of unknown values needed to describe the structure of the galaxy can be reduced to the number of constraints provided by the image so the optimal reconstruction is well-defined. Other structural properties can also be described under this framework.

I provide a mathematical framework and analyses that prove the uniqueness of solutions under certain conditions and to show how uncertainty may be precisely and explicitly expressed. Empirical results are shown using real and synthetic data. I also show a comparison to a state-of-the-art two-dimensional modelling technique to demonstrate the contrasts between the two frameworks and show the important advantages of the three-dimensional approach. In combination, the theoretical and experimental aspects of this thesis demonstrate that the proposed framework is versatile, practical, and novel—a contribution to both computer science and astronomy.

Acknowledgements

Dedication

This thesis is dedicated to my parents, Rose Cormier and Prof. James Cormier, for encouraging my interest in all fields of science. I also dedicate this work to my late grandfather, Prof. Randall Cormier, and my late uncle Mike Cormier.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

I consider the problem of reconstructing the three-dimensional structure of a galaxy, using only one image, and using this three-dimensional reconstruction to explain the two-dimensional image of the galaxy. Figure 1.1 shows the relationship between the viewer and the galaxy for two different viewing angles. Since the viewers would need to be so far apart to obtain these views (on the order of 100 million light-years or more for many galaxies), it is impossible for astronomers to take photographs of distant galaxies from more than one viewing angle; this is why it is necessary to use a single view for all modelling of a galaxy.
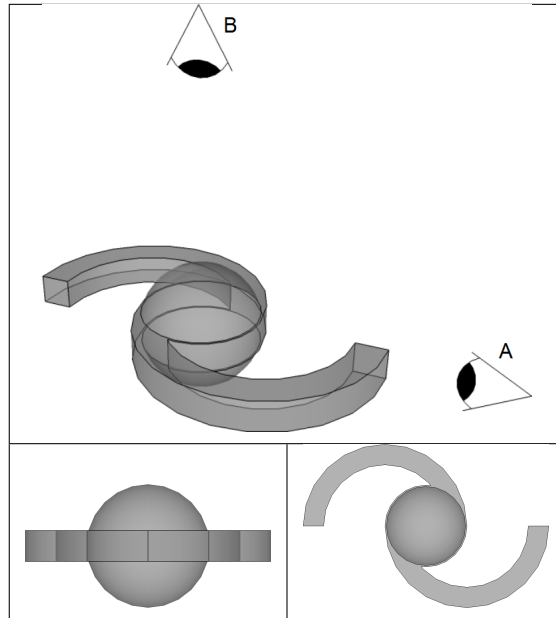
Figure 1.1: A three-dimensional schematic view of a spiral galaxy with two views. On the left is the image seen by viewer A, and on the right is the image seen by viewer B.

Over the course of this thesis, I show that working with three-dimensional reconstruction has significant advantages over using two-dimensional image models, even when only explaining properties of the two-dimensional image. This process is shown in Table 1.1: the original image (of the galaxy NGC 4452) is shown on the left, the three-dimensional reconstruction is shown (from a different viewing angle) at centre, and the reconstructed image is shown at right.
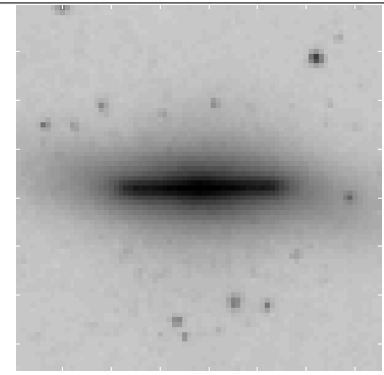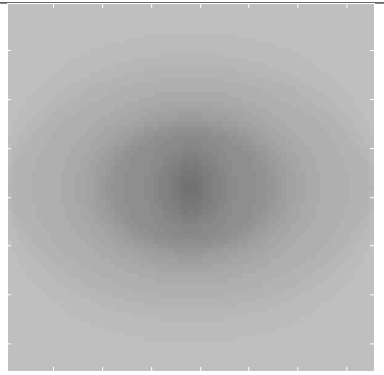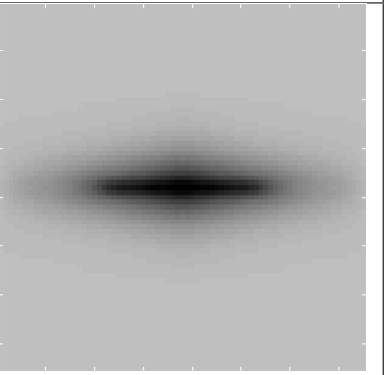
Original Image	Reconstruction (45°-view)	Reconstructed Image

Table 1.1: Example of a real galaxy (specifically, NGC 4452) and a reconstruction which is plausible in three dimensions. A view of the reconstruction from another angle is shown in the center.
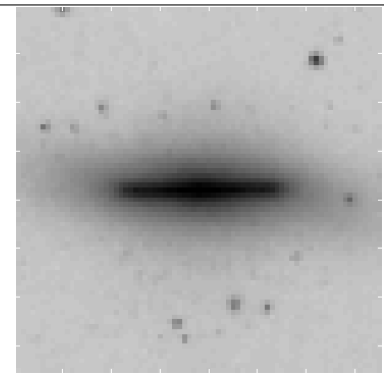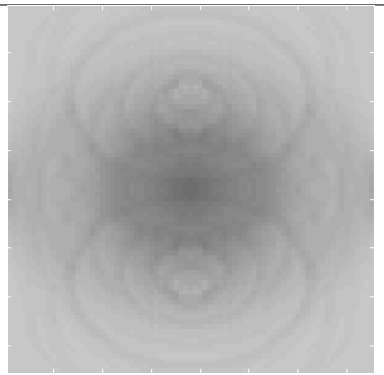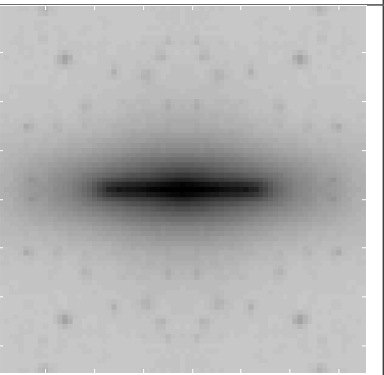
Original Image	3-D Reconstruction	Reconstructed Image

Table 1.2: NGC 4452, with a 3-D reconstruction that explains the two-dimensional image but is implausible when viewed from another angle.

Even reconstructions which reproduce the original two-dimensional image may not be reasonable in three dimensions. An example of this problem, using the same image as above, is shown in Table 1.2. Although the reconstructed *image* is close to the original, the *three-dimensional* reconstruction shows strange ring structures far from the rest of the galaxy. If this were true, one would expect to see galaxies like this from all angles,

3
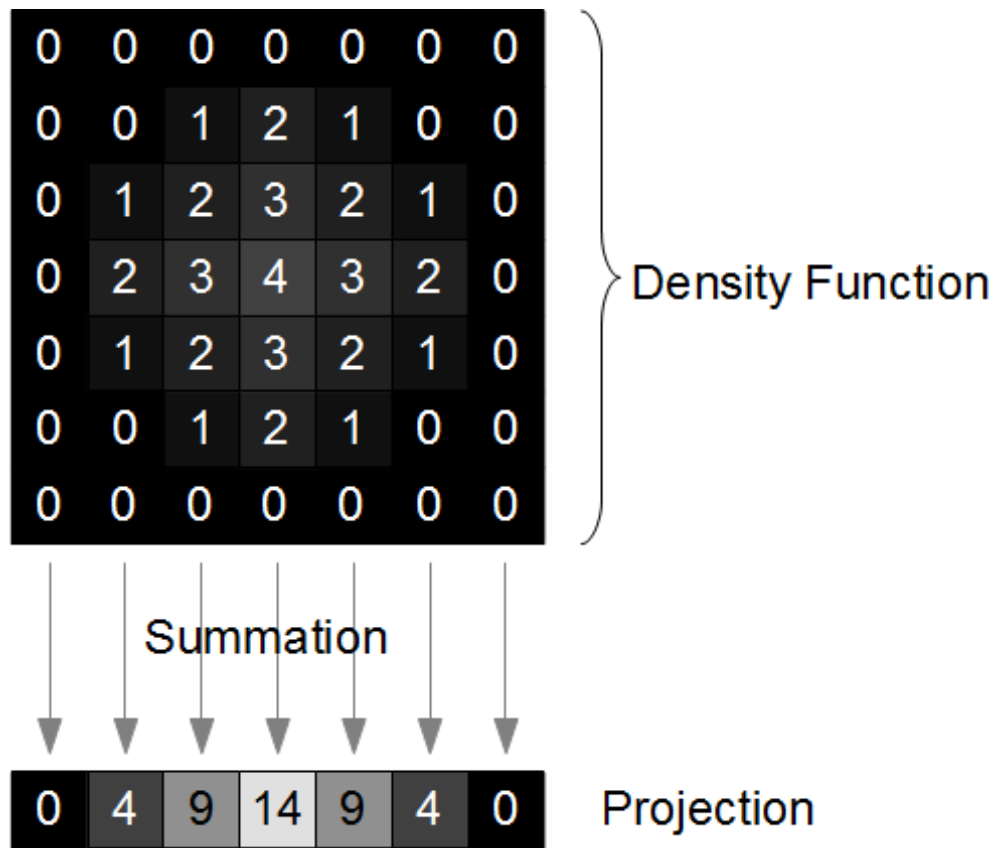
Figure 1.2: Diagrammatic representation of projection of a discrete two-dimensional function to a discrete one-dimensional function

but that is not the case; this reconstruction is simply not plausible. A significant portion of the thesis is dedicated to describing the differences between reconstructions which are plausible (under an application-dependent model) in two dimensions and those which are also plausible in three dimensions.

In a real image of a galaxy, the original image is formed by a process called projection, in which the image is formed by summation of luminosity density along lines of sight. Figure 1.2 shows an example of the projection process in graphical and numerical form. Although this diagram shows a two-dimensional luminosity density distribution, it serves to illustrate the principles that apply to the three-dimensional case.

This projection process should not be confused with perspective or orthographic projec-

tion, which determine the directions of lines of sight. In describing an image, it is necessary to specify whether or not the image is formed by projection and whether the projective geometry is orthographic or perspective. From this point, I assume orthographic projection *geometry*, in which lines of sight are parallel. The term *projection* refers to either the process of summation along a line of sight or the image formed by the projection process, unless otherwise noted.

I present new techniques for reconstruction which use only a single projection. Since much less data is available for reconstruction, structural constraints must be imposed to obtain a good solution. Possible structural constraints include symmetry, smoothness, and compactness; I focus on symmetry constraints and propose a compactness constraint expressed as a regularization term. I also provide a framework for developing new constraints and relate some of these constraints to existing regularization techniques.

The reconstruction problem falls into a very broad class of mathematical problems called inverse problems; this class of problem is described in Section 5.1. In the past, computer vision approaches to three-dimensional reconstruction from a single image have assumed that the image of of an opaque surface which reflects light, generally using the silhouette of an object, shading, or similar cues. This problem is called shape-from-X; "X" refers to the specific cue used in each problem, such as shape-from-shading [22]. Computer vision work with projections has typically assumed that many projections are available, such as in computed tomography [7], which is used in medical imaging, among other areas, and requires projections from a range of angles. Most existing work on reconstruction from a single projection has been done in astronomy, but this work is less versatile than the framework I present here. Related work in computer science is discussed in Chapter 5, including one single-projection method developed for but not inextricably linked to astronomical problems (Section 5.5). Chapter 6 describes related work which is specific to astronomical applications.

This framework will be of use to astronomers in that it will allow nonparametric three-dimensional reconstruction with a wide variety of models. My intention in developing this framework is that the reconstructions will be used for generating sophisticated three-dimensional structural features for large-scale, automated processing of thousands of galaxy images. For example, information about which parts of the image can be explained by three-dimensional symmetric functions could be used to classify galaxies by their structure. Images in fields other than astronomy could also be analysed using this framework, provided that they are formed by projection and appropriate models can be created.

In this thesis I provide both theoretical and experimental results. The theoretical aspects of this thesis focus on the application of linear algebra techniques to the recon-

struction and proofs that solutions are or are not unique in specific significant cases. The algorithms used for reconstruction are based on these proofs. Practical concerns about implementation, including discretization techniques and efficiency issues, are discussed as well as theoretical considerations. When possible, alternative methods are shown.

The experimental aspects of this thesis are based largely on synthetic data. Since quantitative ground truth data is not available for the aspects of galaxy structure under investigation, I use synthetic data to evaluate the performance of the system. Fortunately, detailed models of the imaging process exist for astronomical images, including optical effects and noise characteristics. These models can be used to generate very realistic synthetic images of any three-dimensional luminosity density distribution. Furthermore, empirically-derived functions can be used to create a three-dimensional structure that results in images very similar to real images of galaxies. In addition to the synthetic data, some tests are performed on real data.

This thesis is arranged into seven chapters, including this introduction. Chapter 2 briefly summarizes the broader context of the galaxy reconstruction problem. The theoretical and practical aspects of the new reconstruction method are described in Chapter 3. Experimental results are presented in Chapter 4. Chapter 5 describes related work in the field of computer science and shows how the new reconstruction system differs from previous techniques. Chapter 6 describes related work in the field of astronomy, and compares the capabilities of the new reconstruction system with established techniques. Finally, Chapter 7 consists of concluding remarks.

Chapter 2

Astronomical Background

My work is motivated by the problem of reconstructing the three-dimensional structures of galaxies. A galaxy is a large-scale concentration of matter, consisting of a gravitationally-bound collection of stars, gas, dust, and other bodies. The luminosity of a galaxy ranges from as little as 1000 times the luminosity of the Sun to as much as 10^{12} times the luminosity of the Sun. Luminosity is related to the mass of a galaxy, but is also influenced by other factors such as the age of the stars in the galaxy [15].

One of the most fundamental measurements that astronomers use in studying galaxies is brightness. There are two measurement systems used by astronomers to express brightness. The simplest system uses power per unit of area or unit of solid angle. The other system used is the magnitude system; this system uses a logarithmic scale, where lower numbers represent greater brightness. This system can be further divided into relative magnitude (based on the brightness as seen from Earth) and absolute magnitude (based on the intrinsic luminosity of each object). For my purposes, I use the simpler, linear system based on standard units, rather than the more complex magnitude system.

Surface brightness is defined as the brightness of the projected image of the galaxy per unit of area [15]. The surface brightness profile is the distribution of surface brightness with increasing distance from the center of the galaxy.

Galaxies are classified morphologically—that is, based on shape—by astronomers. Such classifications are important in astronomy and astrophysics for a number of reasons. Since morphological classifications reflect the physical structures of galaxies, they allow astronomers to determine the relationship between the structure of a galaxy and the processes occurring within it. Furthermore, morphology is a characteristic useful in associating

7
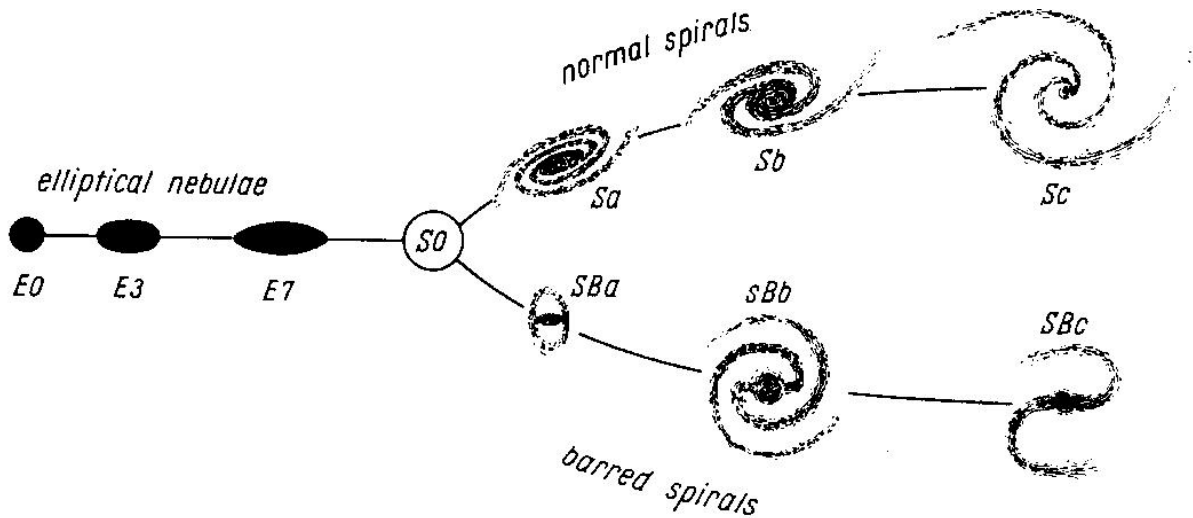
Figure 2.1: A "tuning fork" diagram of Hubble's morphological classification system, which shows different types of galaxies on a continuum. Image from [8]

nearby galaxies with more distant ones (which are observed at an earlier stage in their evolution due to the finite speed of light). Finally, shape is a natural point of division given the variety of shapes of galaxy that are observed. The fact that galaxies can be divided into several categories based on shape is an important influence on the framework I present; the specific properties of these shapes determines the symmetry relationships and other structural assumptions used in testing the framework.

Galaxy morphology has long been studied. The first system of classification to come into common use was devised by Hubble in 1926; Hubble extended the system in 1936 [24]. This system consists of three broad categories: elliptical (S1–S7), spiral (Sa–Sc), and barred spiral (SBa–SBc) galaxies. It also includes lenticular galaxies (S0), similar to a spiral in that they have a disk structure and similar to an elliptical in that they have no spiral arms. Since Hubble, irregular galaxies (with and without bar-like structures) have been added (Im and IBm). These categories are generally represented by the "tuning fork" diagram shown in Figure 2.1 [15, 24].

2.1 Elliptical Galaxies

In general, elliptical galaxies have approximately elliptical *isophotes* (level curves of the image function) and a much simpler morphology than disk galaxies [15]. In some cases, the ellipticity of the isophotes can change with the isophote level [15]. Elliptical galaxies can be divided into "boxy" and "disky" ellipticals [15, 24]; disky ellipticals have an internal disk-like substructure detectable in the isophote curves, whereas boxy ellipticals have isophotes with a rounded rectangular appearance [24].

The surface brightness of an elliptical galaxy can be described effectively by a Sérsic model, as shown in Equation 2.1 [15, 24]. In this equation, $I(r)$ represents surface brightness at the reduced radius $r = \sqrt{r_{min}r_{max}}$, I_0 represents the surface brightness at the center of the galaxy, and r_0 is a scale length for the galaxy [24].

$$I(r) = I_0 \exp\left(-\left(\frac{r}{r_0}\right)^{\frac{1}{n_{sersic}}}\right) \tag{2.1}$$

This model is a generalization of de Vaucouleur's Law, in which $n_{sersic} = 4$. This relationship is effective for higher-luminosity elliptical galaxies, but a more general Sérsic model is needed to obtain a good fit for lower-luminosity ellipticals [15, 24].

2.2 Disk Galaxies

Disk galaxies are morphologically complex. The category includes lenticular and spiral galaxies; in both cases, a bar may be present. In addition to these features of the disk, there is frequently a central bulge; this feature is dominant in early-type disk galaxies and small or absent in late-type galaxies [15].

The surface brightness of a disk in this type of galaxy can be modelled by an exponential profile; this is equivalent to the Sérsic model described in Equation 2.1 with $n = 1$ [15]. The core is generally fit by the Sérsic model with n left as a free parameter (as in the case of an elliptical galaxy), and the surface brightness profile is modelled by the sum of these two components [15]. In some cases, the fitting procedure for the bulge assumes $n_{sersic} = 4$ (corresponding to de Vaucouleur's law) [9]. This modelling can be difficult even with the use of full 2D information because the isophotes of the disk may not have the same center, and dust can block significant light. Blurring of the galaxy image can also interfere with this fitting [15]. One advantage in using two-dimensional information rather than trying

to create a one-dimensional profile is that while the flat disk is foreshortened along the minor axis of the image, the approximately spherical bulge is not; it is difficult to account for this when reducing the dimensionality of the profile before fitting, but much easier to account for while fitting two-dimensional functions.

2.2.1 Lenticular Galaxies

Lenticular galaxies (S0 in the Hubble classification system) are disk galaxies having no spiral arms. In some respects they look similar to elliptical galaxies, and it is believed by some, including van den Bergh [24], that some S0 galaxies have been misclassified as ellipticals and vice versa. Disks with an exponential profile are observed more frequently in S0 galaxies than in ellipticals. A related type of galaxy is the barred lenticular (SB0); these are similar to S0 galaxies, but contain an oblong bar structure in the center.

2.2.2 Spiral Galaxies

Spiral galaxies (Hubble classifications Sa, Sb, and Sc) are probably the most familiar of all galaxies. The most famous images of galaxies are of these types. Spiral galaxies are characterized by the presence of spiralling structures—called spiral arms—in the galaxy disk. The classification of the galaxy is determined by, among other features, the relative prominence of the bulge and arms and the tightness of the spiral [15, 24].

In some spiral galaxies, the arms originate not from the bulge but from the ends of bar structures. These galaxies are called barred spiral galaxies, and are classified as SBa, SBb, or SBc, according to criteria similar to those for ordinary spiral galaxies. Our own Milky Way galaxy is a barred spiral [15, 24].

2.3 Image Modelling and Reconstruction

Several approaches have been take in attempts to model images of galaxies or to reconstruct the three-dimensional structure of a galaxy from an image. I divide these approaches into three categories: two-dimensional parametric models, three-dimensional parametric models, and three-dimensional Fourier methods. Two-dimensional parametric models (discussed in detail in Section 6.2.2) fit specific functions to the image; in these cases, the optimization process is non-convex and the use of specific two-dimensional functions limits the

expressiveness of these models. Two-dimensional models also ignore the three-dimensional structure of the galaxy. Three-dimensional parametric models (Section 6.3) are similar, but use three-dimensional functions and model the projection process; these models are still limited in expressiveness and require non-convex optimization. The three-dimensional Fourier method (Section 5.5) is similar to the method I will present, but has been derived for a specific type of symmetry (which makes it less adaptable than my method) and struggles to handle uncertainty.

Chapter 3

Reconstruction from a Single Projection

My goals are to estimate a three-dimensional density function using a single two-dimensional projection of that function, and to use that estimate in modelling the two-dimensional image in a way consistent with assumptions about three-dimensional structure. The result is a three-dimensional model of the density function. Section 3.1 describes the problem and approaches to a solution in general. Next, Section 3.2 provides a more detailed mathematical analysis of the problem and its solutions. Finally, I discuss implementation issues in Section 3.3.

3.1   Overview

In this section, I describe in general terms the various aspects of the reconstruction problem and the techniques used to solve it. First, I elaborate upon the projection problem. Second, I describe the reconstruction problem in detail. Third, I define and describe the ambiguity in the reconstruction problem. Finally, I describe the use of symmetry to reduce or eliminate ambiguity.

3.1.1   Projection

When a telescope acquires an image of a galaxy, light from each star falls on a portion of the sensor corresponding to the projection of the star's location in world coordinates

onto the image plane. If the sensor resolution were high enough, individual stars would be visible in the image. In real instruments, each star is effectively a point source; even stars in our own galaxy are almost always too small (in terms of angular diameter) to be resolved as a disk by any telescope yet built.

In addition to being point sources, stars are very densely packed in the image; many stars along the line of sight to a pixel in the sensor contribute light to the pixel value. Specifically, the pixel value is determined by the sum of light emitted by all of these stars. Because the individual stars are so small and so numerous, it is better to model the galaxy as a continuous light-emitting medium rather than as a collection of point sources. This is analogous to treating a gas as a continuous substance with pressure and temperature rather than as a collection of discrete molecules. In the case of the galaxy, the important property of the continuous medium is its "luminosity density"—the amount of light emitted per unit of volume. Note that luminosity can vary dramatically between different stars, so the same luminosity density can be achieved with many fewer red giant stars than red dwarf stars.

The value of a pixel in the continuous model is the integral of the luminosity density along the line of sight through the pixel. The integration of a density function along a line of sight is called projection. Figure 3.1 shows a two-dimensional example of projection along two different directions.

Consider a projected image I of a luminosity density distribution L. For notational simplicity, assume that the viewing direction is parallel to the z axis of the coordinate system. Furthermore, since a galaxy is so distant compared to its size, assume that all of the lines of sight are parallel. A line of sight through the point (x, y) in the projection therefore passes through the point (x, y, z) in the luminosity distribution for every z, and does not pass through any other points. Based on the definition of projection as "integration along a line of sight", it is easy to see that I can be related to L by the following equation:

$$I(x, y) = \int_{z=-\infty}^{\infty} L(x, y, z)\mathrm{d}z. \tag{3.1}$$

The distance to the galaxy along the line of sight does not change the projection. Consider the projection of a distribution $L'(x, y, z) = L(x, y, z + k)$. The projection of L', then, is

$$
\begin{aligned}
I'(x, y) &= \int_{z=-\infty}^{\infty} L'(x, y, z)\mathrm{d}z \\
&= \int_{z=-\infty}^{\infty} L(x, y, z + k)\mathrm{d}z. \tag{3.2}
\end{aligned}
$$

13
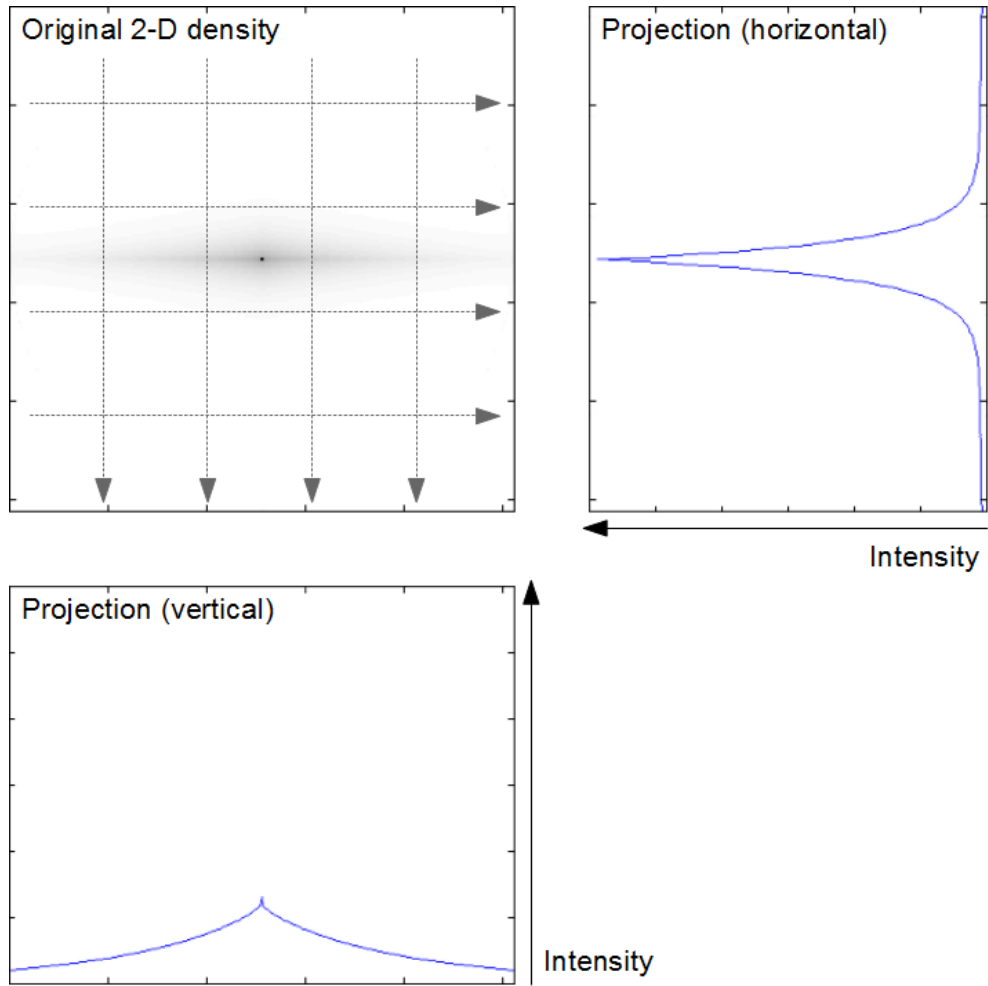
Figure 3.1: Projections of a 2-D distribution—based on a radial slice through a galaxy—along two different lines of sight. In the image of the distribution, white corresponds to a value of 0.

Since the integration limits are infinite, it follows that $I'(x, y) = I(x, y)$. This ambiguity is shown in Figure 3.2. Note that the x and y position of the origin is arbitrary: it affects the position of the origin in the projection, but does not change the structure of the projected function. It is often convenient to use the position of the viewer to determine the direction of the axes of the coordinate system, but to fix the origin at a logical position within the luminosity density distribution. For simplicity, I assume that the luminosity density distribution is defined in a coordinate system where the axis directions are determined by the viewing direction. It is possible to define L in a different coordinate system and perform rotation for the projection process; I will discuss this in section 3.2.2.

Since the projection process is fundamentally a process of addition, it is simple to show that it is linear. Assuming that $I_1(x, y) = \int_{z=-\infty}^{\infty} L_1(x, y, z) \mathrm{d}z$ and I_2 is similarly defined in terms of L_2, linearity can be shown as follows:

$$
\begin{aligned}
I'(x, y) &= \int_{z=-\infty}^{\infty} aL_1(x, y, z)\mathrm{d}z & (3.3)\\
&= a \int_{z=-\infty}^{\infty} L_1(x, y, z)\mathrm{d}z \\
&= aI_1(x, y)
\end{aligned}
$$

$$
\begin{aligned}
I''(x, y) &= \int_{z=-\infty}^{\infty} \left(L_1(x, y, z) + L_2(x, y, z)\right)\mathrm{d}z & (3.4)\\
&= \int_{z=-\infty}^{\infty} L_1(x, y, z)\mathrm{d}z + \int_{z=-\infty}^{\infty} L_2(x, y, z)\mathrm{d}z \\
&= I_1(x, y) + I_2(x, y)
\end{aligned}
$$

Discretization brings implementation challenges which will be discussed in Section 3.3. For now, I assume that L can be modelled by samples of the luminosity density at a finite number of points, arranged in a rectangular grid. I also assume that if the sampling resolution is high enough, L is approximately linear between any two sampling points. Sampling points of this type are often called "voxels", as they may be treated as the volume analogue of pixels. The integration process is converted to simple summation, resulting in the following equations of projection (in which x, y, and z are integers):

$$
I(x, y) = \sum_{z=-z_{min}}^{z_{max}} L(x, y, z) \tag{3.5}
$$

It is important to note that, if a finite number of points are used, the range along each axis must be finite. In practice, this means that there is a local region where the luminosity

15
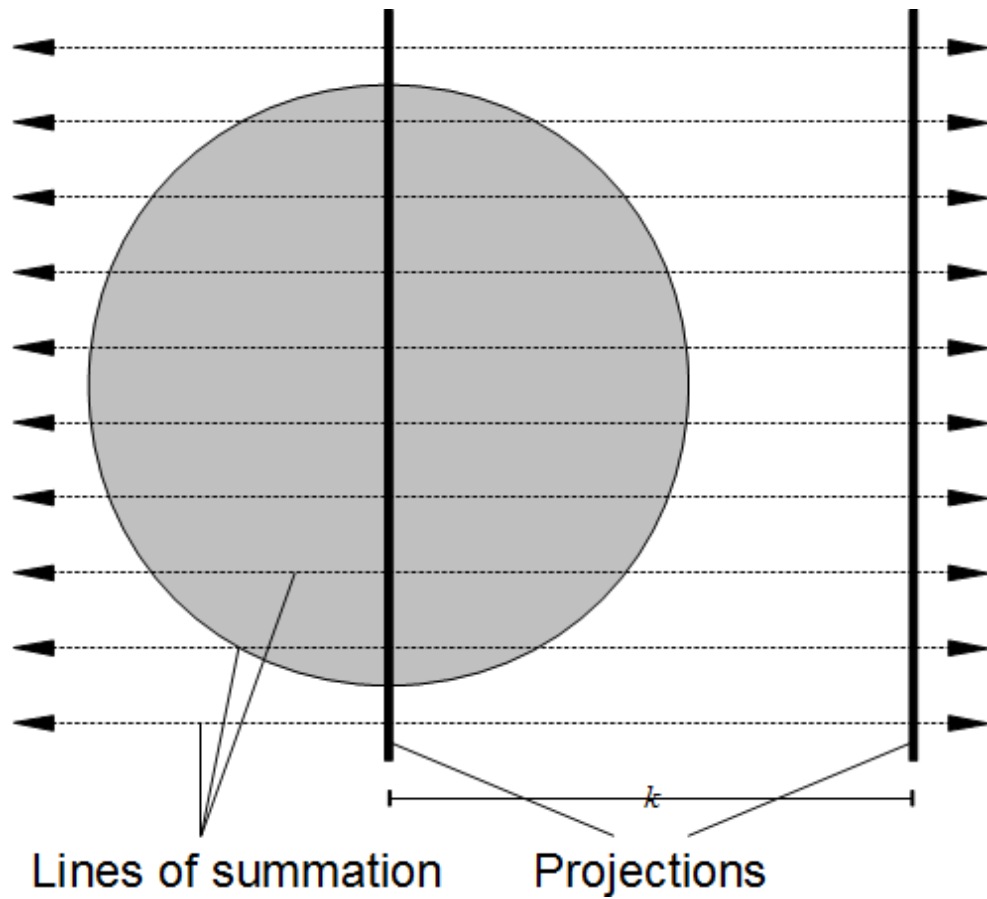
Figure 3.2: Ambiguity along the z-axis (horizontal in the diagram shown). The two projections of the luminosity density distribution are identical, even though they are at different locations along the direction of projection.
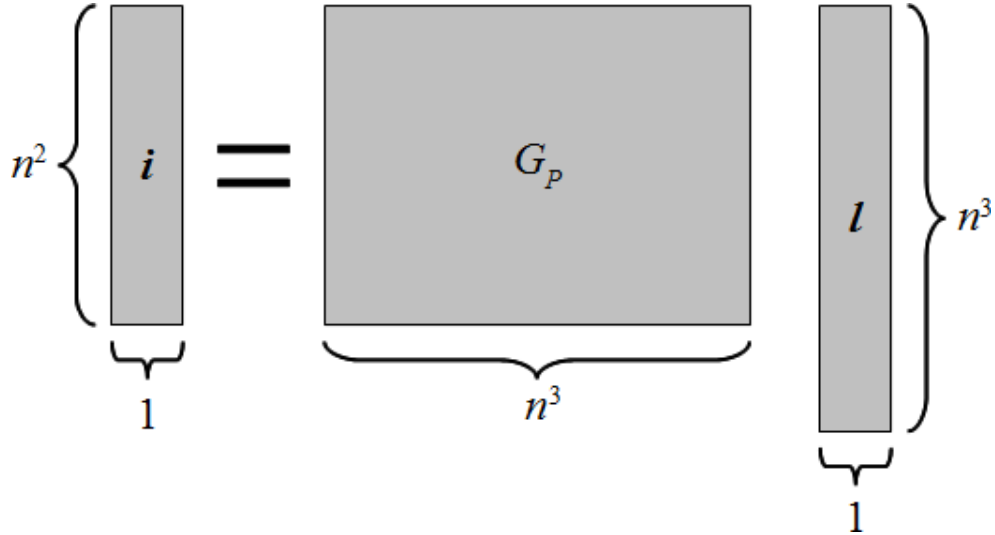
Figure 3.3: Matrices and their dimensions for the discrete projection operation.

density is significant, embedded in a volume where it is negligible. Such an assumption is generally reasonable for galaxies.

The discrete formulation of projection is a logical approach given the nature of astronomical data. The CCD sensors used in modern telescopes have a finite pixel density, and even film has distinct grain. If the resolution were infinite, it would not be necessary to use the luminosity density model; it would be possible to use two telescopes to resolve individual stars in stereo. Since such high resolution will never be achieved, the discrete formulation is the most natural for the problem of reconstructing a galaxy.

In the discrete formulation, the projection process can be seen as a matrix operation. If the projection and the luminosity density distribution are "packed" (typically in raster-scan order) as the vectors $\mathbf{i}$ and $\mathbf{l}$, respectively, it is possible to define a matrix G_P such that

$$\mathbf{i} = G_P \mathbf{l} \tag{3.6}$$

For an image with a resolution of $n \times n$ pixels, and a luminosity density distribution with a resolution of $n \times n \times n$ voxels, the relative dimensions of the matrices are shown in graphical form in Figure 3.3. Each row of the the projection matrix G_P defines the combination of voxels that sum to give the value of the corresponding pixel. This matrix is quite sparse, since only voxels along the line of sight through the correct pixel have a nonzero value in a given row; this reduces memory requirements (see Section 3.3.3). The structure of G_P is

17

$$
\begin{bmatrix} i_{11} \\ i_{12} \\ i_{21} \\ i_{22} \end{bmatrix}
=
\underbrace{\begin{bmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1
\end{bmatrix}}_{G_P}
\begin{bmatrix} l_{111} \\ l_{112} \\ l_{121} \\ l_{122} \\ l_{211} \\ l_{212} \\ l_{221} \\ l_{222} \end{bmatrix}
$$
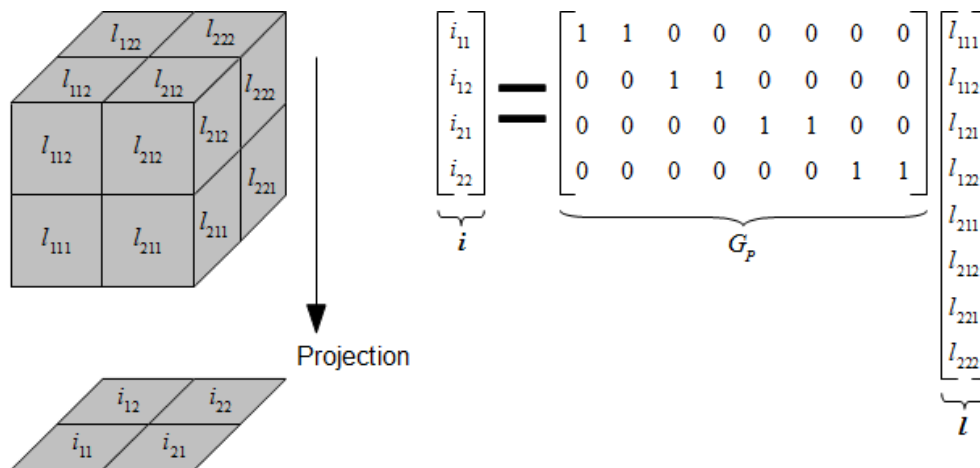
Projection

Figure 3.4: Formation of a projection matrix for a very small 3-D density distribution.

dependent upon implementation; a typical case is shown in Figure 3.4.

The continuous and discrete formulations of the problem each have advantages and disadvantages in terms of theoretical elegance and practical realizability. In the remainder of the discussion, I will use whichever formulation which is clearest and most useful.

3.1.2 Reconstruction

Reconstruction from multiple projections is a well-established problem in computer vision which has been studied extensively using spatial-domain [13, 12] and Fourier [7] approaches. In the case of galaxies, however, only one view is available. For a galaxy 1.0×10^8 light-years from Earth, the camera would need to travel over 1.7×10^6 light-years to obtain a $1°$ difference in viewing angle. Single-projection reconstruction has not been studied as extensively. It has been examined (from a Fourier-domain perspective) by Zaroubi *et al.* [25], which I discuss in Chapter 5; the approach they present is inspired by astronomical problems, but would be applicable to other domains as well.

To determine the best way to reconstruct L from I, it is necessary to define a measurement of reconstruction error to minimize. Since the three-dimensional world is inaccessible, the error function must depend only on the two-dimensional projection. Given an estimate $\hat{L}$ of the three-dimensional luminosity density, it is simple to define a projection $\hat{I}$. A

18

sum-of-squared-errors measurement can be defined in the discrete case:

$$E = \sum_x \sum_y \left(I(x, y) - \hat{I}(x, y) \right)^2.$$ (3.7)

The same error function can also be defined in the continuous case:

$$E = \int_x \int_y \left(I(x, y) - \hat{I}(x, y) \right)^2 \mathrm{d}y \mathrm{d}x.$$ (3.8)

The sum-of-squared-errors error function is a common method which is easy to analyse.

The matrix formulation of the discrete case of projection suggests a way to find an explicit formula for minimizing the sum of squared errors. If the projection matrix could be inverted, the optimal solution could be found using the equation $\hat{\mathbf{l}} = G_P^{-1} \mathbf{i}$. Unfortunately, G_P is an $n^2 \times n^3$ matrix. It is not invertible, and has infinitely many solutions; that is, it is underdetermined. Fortunately, there are established methods for solving an underdetermined system of equations (see, for example, Strang's *Linear Algebra and its Applications* [21]).

3.1.3 Ambiguity

The fact that the projection matrix is not invertible implies that there are multiple solutions which reproduce the projection equally accurately; the problem is *ambiguous*. For example, Figure 3.5 shows voxels representing a "checkerboard" pattern separated into groups with a constant z-value, which I call z-layers. Because addition is commutative, transposing any pair of z-layers does not change the projection formed by summation along the z direction. This shows quite clearly how additional solutions may be constructed given any solution which reproduces the original projection. These are not the only alternate solutions, but their existence proves that the problem is inherently ambiguous.

One approach to ambiguous reconstruction problems is *regularization*, which consists of adding penalty terms to encourage the solution to have some set of properties (depending on the regularization method). Regularization is used to guarantee that a unique solution can be found, but as a consequence of ambiguity this solution may not be identical to the original luminosity density distribution. Furthermore, some regularized versions of the problem may still not have a unique solution. Two common forms of regularization are finding the solution with the smallest L_2 norm, or finding the solution with the fewest non-zero elements. Neither the norm of the solution vector, nor the number of non-zero
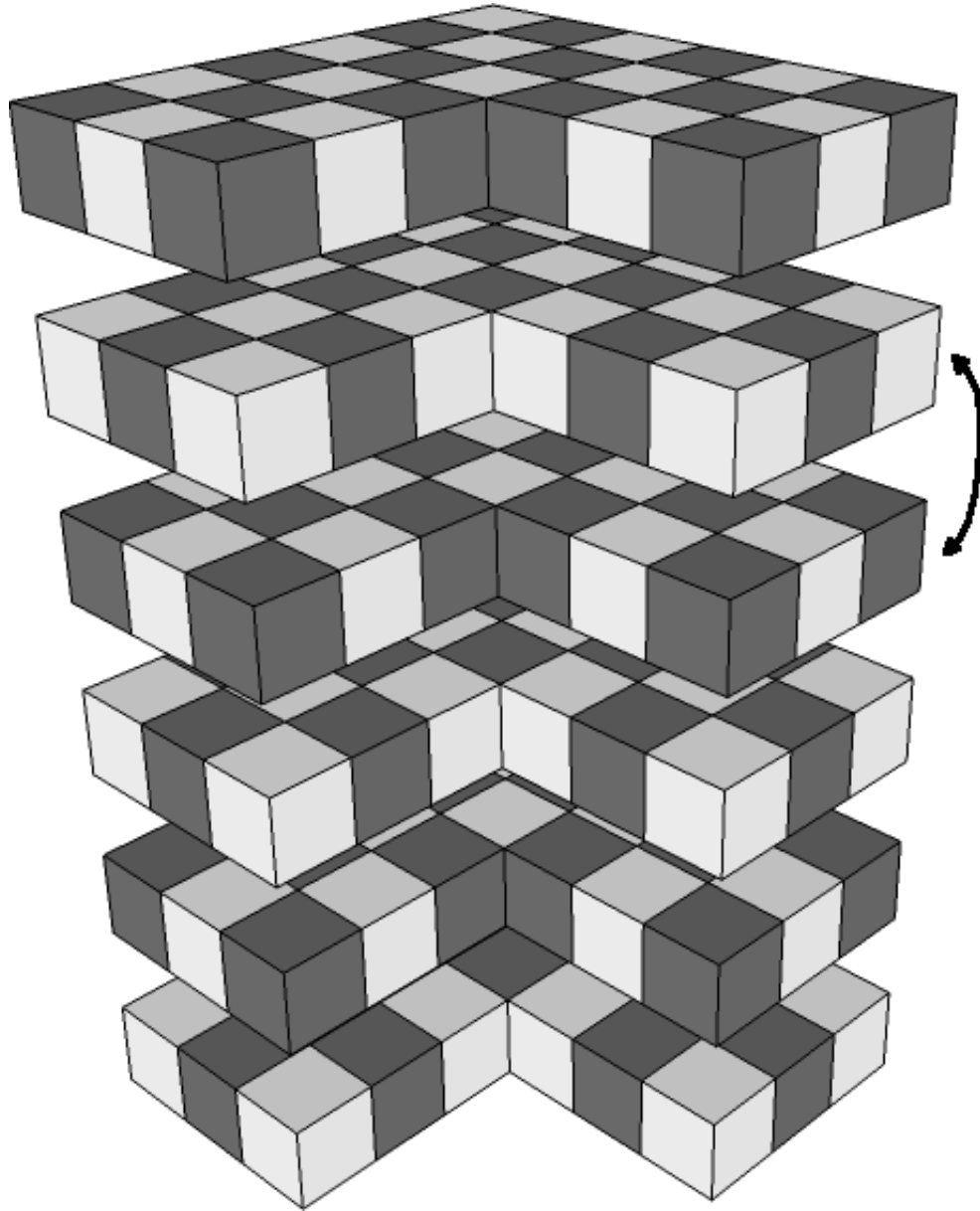
Figure 3.5: Voxels decomposed into z-layers; switching any two layers (as indicated by the double arrow) would not change the projection.

elements it contains, are altered by transposing z-layers; thus neither of these methods will produce a unique solution.

Although the ambiguity cannot be eliminated, it is worthwhile to examine how serious the ambiguity is. If the ambiguity is in some sense "small", it may not be objectionable. This raises an important question: How is the concept of "ambiguity" to be quantified in order to establish its severity and compare between approaches?

One possible means of quantifying ambiguity is enumerating the possible solutions. There are, for example, at most $n!$ distinct orderings of n z-layers. These are not the only reconstructions possible, however. Consider luminosity density distributions $\mathbf{l}_1$ and $\mathbf{l}_2$, which both project to the same image $\mathbf{i}$. Consider the following linear combination of the two solutions:

$$
\begin{aligned}
G_P(\alpha \mathbf{l}_1 + (1-\alpha)\mathbf{l}_2) &= G_P(\alpha \mathbf{l}_1) + G_P((1-\alpha)\mathbf{l}_2) \qquad (3.9) \\
&= \alpha G_P \mathbf{l}_1 + (1-\alpha) G_P \mathbf{l}_2 \\
&= \alpha \mathbf{i} + (1-\alpha)\mathbf{i} \\
&= \mathbf{i}
\end{aligned}
$$

This solution applies for any $\alpha \in \mathbb{R}$. Thus, there are in fact infinitely many distinct solutions if there are two or more. Clearly, then, counting solutions does not quantify ambiguity.

A better way to quantify ambiguity is to compare the number of equations with the number of unknowns. In the case in question, there are n^2 equations and n^3 unknowns; the projection matrix G_P is rank-deficient. The degree of rank deficiency is at least $n^3 - n^2$; since this is $O(n^3)$, the ambiguity is very serious.

Since the model of reconstruction using n^3 independent voxels results in severe ambiguity which cannot be eliminated, it is necessary to examine other restrictions that may be placed on the reconstruction process. The next section examines possible models for the luminosity density functions which, while still flexible, can reduce or eliminate ambiguity.

3.1.4 Symmetry

Spherical   Cylindrical   Reflective   Triaxial Ellipsoidal
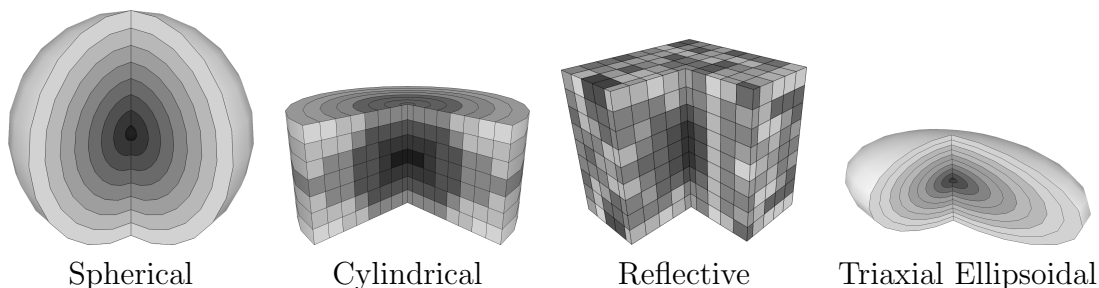
Figure 3.6: Four types of symmetry applicable to galaxies.

Ambiguity can be reduced by imposing symmetry constraints. Galaxies can be effectively approximated by symmetric functions. A lenticular galaxy, for example, is rotationally symmetric about its axis. Also, existing empirically-derived models of galaxies tend to imply both cylindrically symmetry about an axis and reflective symmetry across the equatorial plane. An elliptical galaxy can be modelled by a stretched spherically symmetric function. Spherical, cylindrical, reflective, and triaxial ellipsoidal symmetry are described here, as these are the types most relevant to the study of galaxy images. These types of symmetry are illustrated in Figure 3.6.

Although real objects are unlikely to be precisely symmetric, the best-fit symmetric function (for a given type of symmetry) is still a useful concept. For example, the projection of this best-fit luminosity density distribution represents the portions of the image that can be "explained" under a given symmetry assumption, and the residual after this reconstructed image is subtracted from the original image contains the parts of the image that cannot be explained this way. The residual can be used to find anomalies in the image which would otherwise be overwhelmed by strong symmetric components.

Perhaps the most important application of finding the best-fit symmetric function is that it allows even subtle asymmetries to be isolated and studied without interference from bright symmetric components. In the case of a disk galaxy, spiral arms could be isolated from the rest of the disk by subtracting the cylindrically symmetric component even if they were barely brighter than the disk itself. In the case of NGC 4452 (Table 1.1), the disk is not perfectly flat; this is much easier to see when the symmetric components are removed. Some two-dimensional techniques—notably GALFIT [18, 17]—are used to isolate fainter structures, but they assume specific classes of two-dimensional function, and so are less flexible than a best-fit symmetric function.
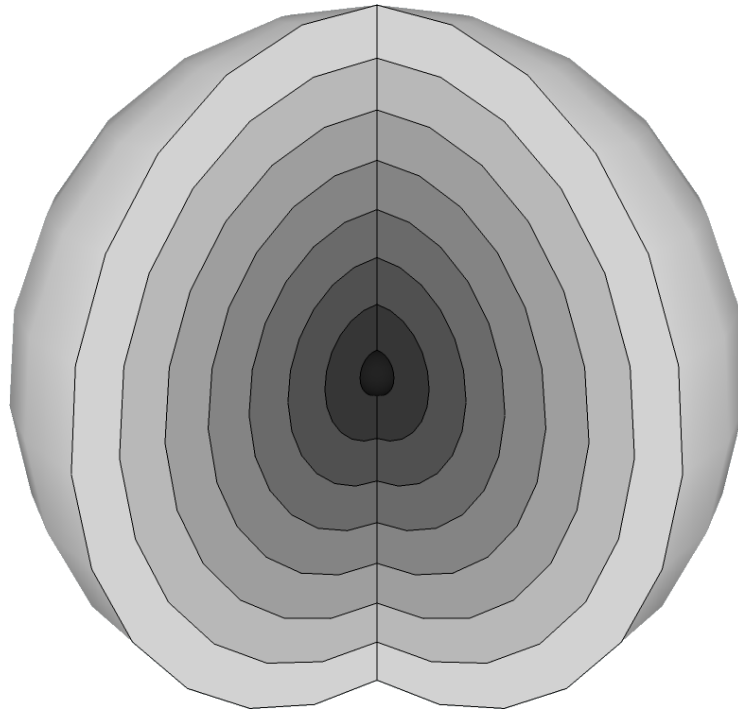
Spherical

Figure 3.7: Structure of a spherically symmetric volume.

A restrictive but often physically plausible type of symmetry is spherical symmetry. A spherically symmetric density function can be thought of as a series of spherical shells such that the value of the function is constant on each shell. A spherically symmetric distribution can be defined by its value at each radius. The number of discrete radii in a volume is $O(n)$; each pixel provides an equation, so there are n^2 equations and $O(n)$ unknowns. This is a dramatic improvement over the n^3 unknowns in a model that uses independent voxels. An algorithm for generating a projection matrix for a spherically symmetric function is described in detail in Section B.1.1, and an algorithm for reconstruction is shown in Section B.4.1. This is the model used in Sections 4.1.1 and 4.1.1. Spherical symmetry is quite restrictive, however, so it is important to examine the possibilities of generalization. Two generalizations are described: Section 3.1.4 describes cylindrical symmetry, and Section 3.1.4 describes triaxial ellipsoidal "symmetry".

Cylindrical
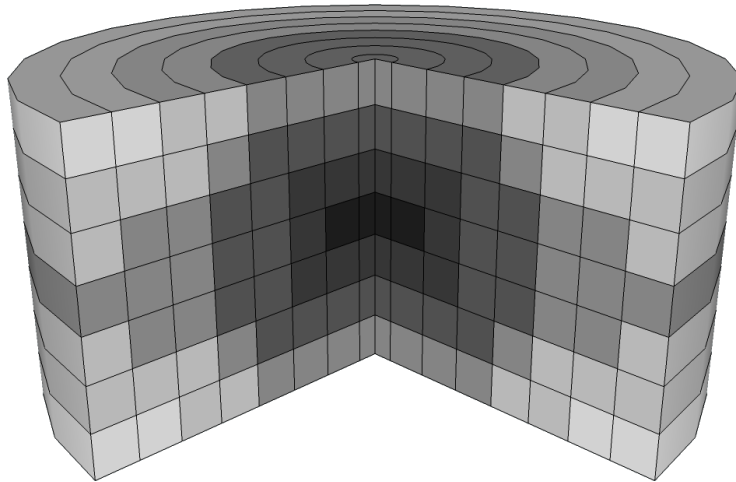
Figure 3.8: Structure of a cylindrically symmetric volume.

Cylindrical symmetry is a form of symmetry which is more general than spherical symmetry but nonetheless considerably reduces the space of solutions consistent with the projected image. A cylindrically symmetric density function can be considered as a set of rings with constant value, each of which has a radius and a position along the axis of the cylinder. Alternatively, it can be viewed as a two-dimensional function rotated about one edge. Under this symmetry assumption, the density function can be represented by a two-dimensional function (of radius and height), so the total number of unknowns is $O(n^2)$ and unique reconstructions are possible, as I will show in Section 3.2.2. This is asymptotically fewer than the $O(n^3)$ samples required for representing independent voxels, although the difference is not as dramatic as for the assumption of spherical symmetry. Cylindrical symmetry, combined with reflective symmetry about the equatorial plane, is the basis of the models used for reconstruction in Sections 4.1.1, 4.1.1, 4.1.2, and 4.3. An algorithm for generating the projection matrix for this case is described in detail in Section B.1.2; the full reconstruction algorithm is shown in Section B.4.2.

Cylindrical symmetry has the disadvantage of more parameters than spherical symmetry, but it is strictly more general. A cylindrically symmetric density distribution is symmetric with respect to rotation about one axis; a spherically symmetric distribution is symmetric with respect to rotation about any axis. If an arbitrary axis of symmetry

is chosen, the spherically symmetric function can be seen as a special case of a cylindrically symmetric function. Obviously there are many varieties of cylindrically symmetric functions which are not spherically symmetric, such as a long, narrow cylinder.

The viewing angle affects the projection of a cylindrically symmetric function. Analysis of the details of cylindrically symmetric reconstruction will require a discussion of rotation methods. This discussion will be found in Section 3.2.2.

Reflective
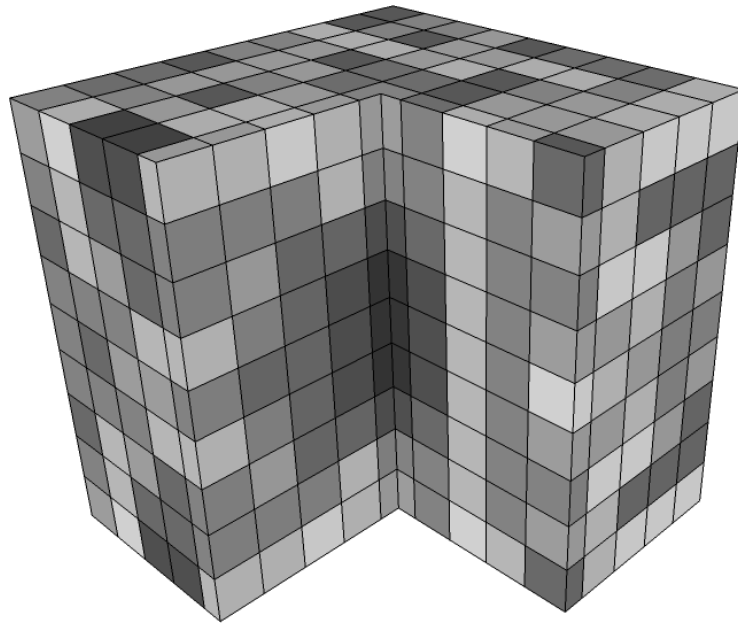
Figure 3.9: Structure of a reflectively symmetric volume; the plane of reflection is horizontal, at the vertical midpoint of the distribution shown

Reflective symmetry (across a plane) reduces the number of variables needed to represent the density function by approximately a factor of two. This does not change the rate at which the number of unknowns scales relative to the number of equations, but it may be useful to enforce this type of symmetry for other reasons. For example, a cylindrically symmetric density distribution may well also have reflective symmetry about its equatorial plane. It is not necessarily useful to apply this type of symmetry to all types of function;

spherical symmetry, for example, implies reflective symmetry about any plane through the center of the sphere, and for planes that do not pass through the origin spherical symmetry is only possible for an infinite, uniform density distribution.

This type of symmetry requires the specification of the plane across which the density distribution may be reflected; in general, this would require three components (two of which define a direction normal to the plane and one of which defines the distance of the plane from the origin). However, one would often use an assumption about the position of the plane of symmetry, or define the origin as the intersection of planes of symmetry, which would reduce the number of parameters.

Triaxial Ellipsoidal
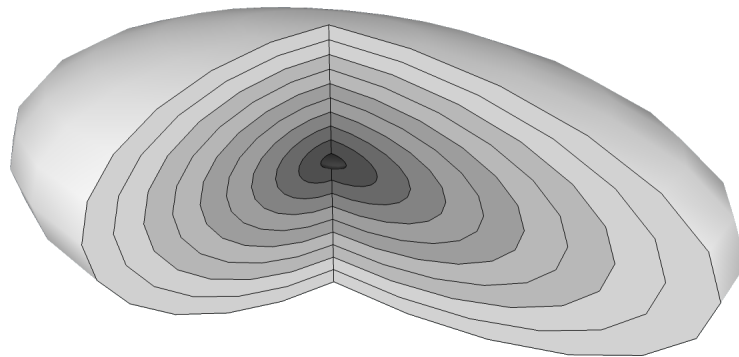
Figure 3.10: Structure of triaxial ellipsoidally symmetric volume.

Spherical symmetry is simple, but it does not necessarily reflect the common shapes of galaxies and parts of galaxies. The model commonly used by astronomers for elliptical galaxies and the bulges of disk galaxies is a triaxial ellipsoid [24, 15, 23]. This shape is a generalization of a sphere, specifically a scaling along three orthogonal axes where the scaling factors may not be equal. If all three scaling factors are equal, the result is spherical symmetry. If any two scaling factors are equal, the distribution can be represented by a cylindrically symmetric function (Section 3.1.4). If, however, no two scaling factors are equal, this sort of symmetry represents a function that cannot be represented by either model of symmetry previously described. If the scaling factors and rotations are assumed to have been estimated by other means, the number of unknowns is identical to the number of unknowns in spherical symmetry.

3.2 Characterizing Ambiguity

This section describes the characteristics of ambiguity in the reconstruction process, and ways of handling it, in detail. First, a proof is presented that shows that there is a unique solution for reconstructing a spherically symmetric distribution. Next, I show that for a cylindrically symmetric luminosity density distribution, the reconstruction may or may not be unique, depending on the viewing angle. The third section examines the relationship between ambiguity and the properties of the projection matrix. The fourth section discusses regularization methods. Finally, the fifth section describes how symmetry models can be defined.

3.2.1 Spherical Symmetry

Under the assumption of spherical symmetry, there are fewer unknowns than equations. This suggests that there is a unique solution to the reconstruction problem; however, it is possible, that a system of linear equations may have infinitely many solutions even if there are more equations than unknowns. This occurs if enough of the equations are linearly dependent upon each other. It is therefore necessary to prove that a unique solution exists for a particular symmetry assumption even if there are fewer unknowns than equations.
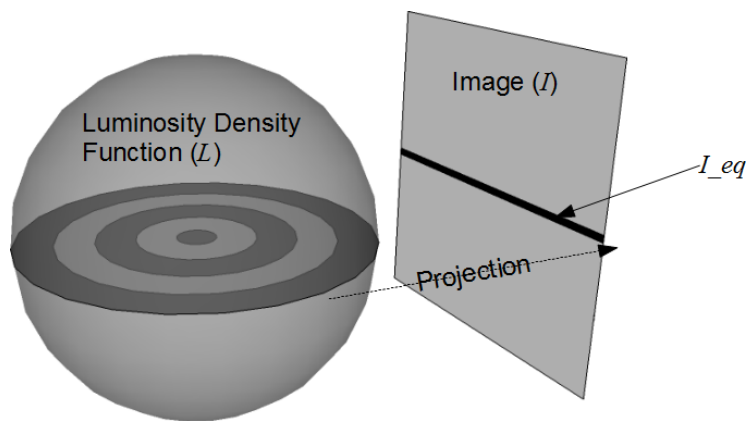
Figure 3.11: Diagram showing key ideas in the proof that the reconstruction of a spherically symmetric is unique

The Abel transform [7], which will be used in establishing uniqueness properties for spherically and cylindrically symmetric reconstruction, is related to the projection of a

two-dimensional, circularly symmetric density distribution to a one-dimensional "image". A circularly symmetric function is two-dimensional, but it can be described by a one-dimensional function of radius; that is, $g(x,y) = f(r)$ where $r = \sqrt{x^2 + y^2}$. The Abel transform of $f(r)$ is defined to be the projection of $g(x,y)$ onto a line, which is given by $f_A(x) = \int_{-\infty}^{\infty} f(\sqrt{x^2 + y^2})\mathrm{d}y$ [7].

Lemma 1. *The Abel transform is invertible. Given the Abel transform $f_A(x)$, $f(r)$ can be found using the following equation:*

$$f(r) = -\frac{1}{\pi} \int_r^{\infty} \frac{f_A'(x)}{\sqrt{x^2 - r^2}} \mathrm{d}x. \tag{3.10}$$

This is stated (but not formally proven) by Bracewell [7]. Because the Abel transform is invertible, there is a unique $f(r)$ for a given $f_A(x)$ (assuming, of course, that $f_A(x)$ is the Abel transform of some function).

Theorem 1. *Given a projection I of a spherically symmetric density function L, there is no other spherically symmetric density function that projects to I; that is, L is unique.*

Proof.   1. Let $L(x,y,z)$ be a spherically symmetric luminosity density distribution.

2. Let $I(x,y) = \int_{-\infty}^{\infty} L(x,y,z)\mathrm{d}z$ be the projection of L along the z-axis.

3. Since L is spherically symmetric, $L(x,y,z) = f(r)$, where $r = \sqrt{x^2 + y^2 + z^2}$, for some function f.

4. Substituting f into the definition of I gives

$$\begin{aligned} I(x,y) &= \int_{-\infty}^{\infty} f(r)\mathrm{d}z \\ &= \int_{-\infty}^{\infty} f(\sqrt{x^2 + y^2 + z^2})\mathrm{d}z. \end{aligned} \tag{3.11}$$

5. Let $I_{eq}(x) = I(x,0)$ represent the projection of the equatorial plane of L (see Figure 3.11)

6. $I_{eq} = \int_{-\infty}^{\infty} f(\sqrt{(x^2 + 0^2 + z^2)})\mathrm{d}z = \int_{-\infty}^{\infty} f(\sqrt{(x^2 + z^2)})\mathrm{d}z$

7. I_{eq} is therefore the Abel transform of f [7]

8. By Lemma 1 f is uniquely determined by I_{eq}.

28

9. Since $L(x, y, z) = f(\sqrt{x^2 + y^2 + z^2})$, L is uniquely determined by f.

10. Since $I_{eq}(x) = I(x, 0)$, I_{eq} is uniquely determined by I.

11. L is therefore uniquely determined by I when spherical symmetry is assumed.

$\square$

This result is of great significance to this project. It demonstrates conclusively that imposing symmetry constraints can not only reduce ambiguity but can eliminate it entirely. In Sections 4.1.1 and 4.1.1, it is easy to see how effective three-dimensional reconstruction can be when ambiguity is eliminated.

3.2.2 Cylindrical Symmetry

In the case of cylindrical symmetry, the number of unknowns is comparable to the number of equations (both are $O(n^2)$). It is worth examining whether or not the best-fit reconstruction is uniquely determined, as is the case for spherical symmetry, or if infinitely many solutions exist.

The models used thus far—independent voxels and spherical symmetry—have no preferred orientation. The luminosity density function may be defined in the viewer's coordinates just as easily as in any other coordinate system. Cylindrical, reflective, and ellipsoidal symmetries, on the other had, have a natural coordinate system which can be defined based upon the directions of symmetry. It is now necessary to distinguish between world coordinates (the natural coordinate system of the density distribution) and viewer coordinates (the coordinates defined by the projection process), both of which are important in defining the reconstruction problem clearly. Rotation can be accounted for using a rotation matrix R which converts a point in viewer coordinates to a point in world coordinates. The point does not move relative to structures in the luminosity density distribution; multiplying by R simply changes the representation of its fixed location. This rotation is shown in Figure 3.12 More formally, given a point $\mathbf{p} = \begin{bmatrix} x & y & z \end{bmatrix}^{\mathrm{T}}$ in viewer coordinates, the equivalent point in world coordinates is

$$\mathbf{p}' = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} = R\mathbf{p}. \tag{3.12}$$

This conversion between coordinate systems can be included in the projection process, so that L is defined in world coordinates and the projection process is along the z-axis of the
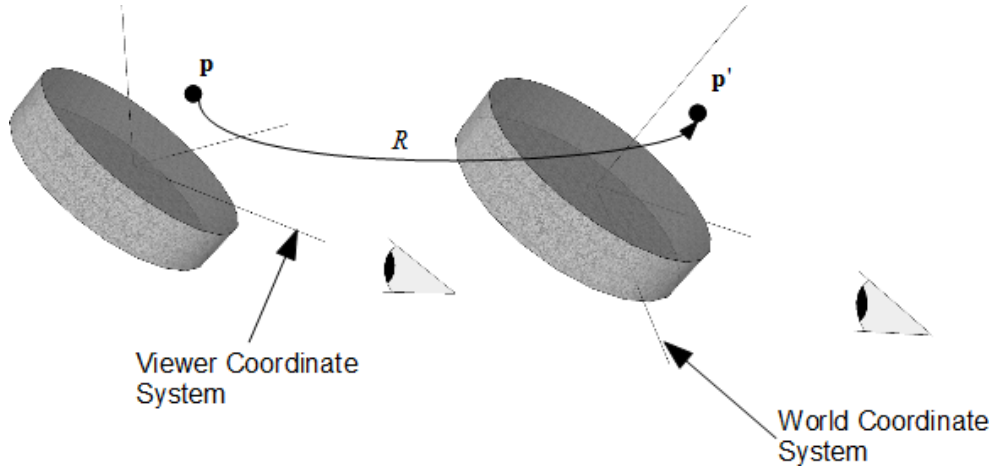
Figure 3.12: Conversion from viewer coordinates to world coordinates for a point **p**.

viewer's coordinates:

$$I(x, y) = \int_{z=-\infty}^{\infty} L\left(R\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) \mathrm{d}z. \tag{3.13}$$

This formulation of rotation allows for three axes of rotation. This is not always necessary. For example, rotation of a cylindrically symmetric function about its axis of symmetry changes nothing. Furthermore, the x- and y-axes of the viewer coordinates are essentially arbitrary. By default, they are determined by the x- and y-axes of the image, but they can be defined otherwise if it is more convenient to analyse. Conceptually, this could be seen as rotating the camera without changing the direction in which it is pointed. These factors can make a significant difference; for a cylindrically symmetric distribution, the x- and y-axes of the viewer coordinate system can be chosen such that the only rotation needed is about the x-axis. For the sake of simplicity, this single rotation will be referred to as the *viewing angle* for cylindrically symmetric distributions.

In addition to the direction of the axes, the location of the origin must be determined. The best location for the origin is the natural center of the luminosity density distribution. For a galaxy, the galaxy's center is the ideal position for the origin; in general, there is likely to be a natural origin for the coordinate system.

Much like projection without rotation, projection with rotation has a discrete form. In

this form,

$$I(x, y) = \sum_{z=z_{min}}^{z_{max}} L(R \begin{bmatrix} x \\ y \\ z \end{bmatrix}).$$ (3.14)

Assume, for the moment, that L is defined continuously but can be represented by a finite number of points, between which the values are linearly interpolated (for further details, see Section 3.3.1). Therefore, each pixel value in the image is a linear combination of the values of these sampling points. Like projection without rotation, projection with rotation can be represented as matrix multiplication. Let $G_P(R)$ represent the matrix corresponding to projection along the direction specified by the rotation matrix R. Then $\mathbf{i} = G_P(R)\mathbf{l}$. When the viewing angle is irrelevant, or R is the identity matrix, the projection matrix will continue to be denoted by the slightly briefer G_P.

By considering projection of a cylindrically symmetric luminosity density distribution perpendicular to and parallel to the axis of symmetry of the distribution, two important facts can be shown: first, that the solution *may be* uniquely determined; and second, that the solution is *not always* uniquely determined. As a result, some viewing angles are more likely to lead to reconstructions with artefacts than others, and particular care must be taken when handling ambiguity for cylindrically symmetric functions.

Viewer coordinates   World coordinates   Coordinates for f
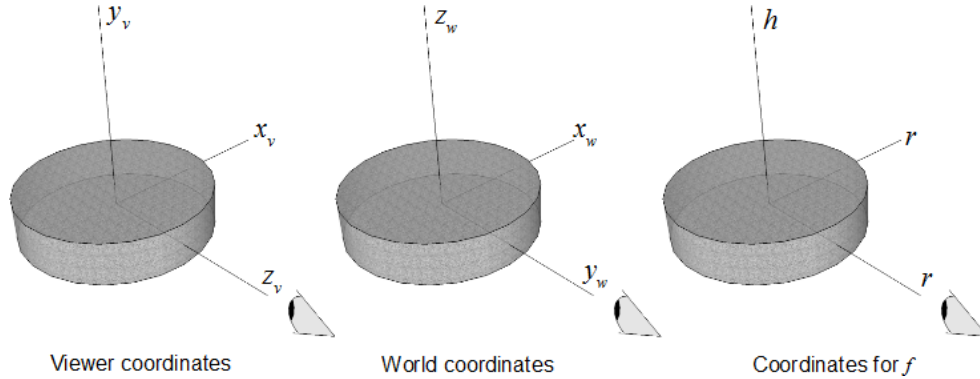
Figure 3.13: Coordinate systems used in the proof that the projection of a cylindrically symmetric luminosity density distribution in the edge-on direction uniquely determines its three-dimensional structure.

In the first case, the projection is formed along a direction perpendicular to the distribution's axis of symmetry. It is convenient to describe this case as an "edge-on view"; in the case of a disk galaxy this is an intuitive description, since it corresponds to viewing the thin edge of the galaxy.
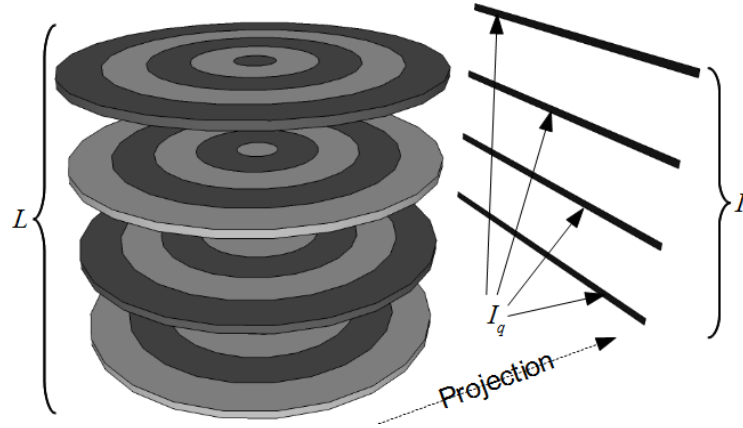
Figure 3.14: Diagram showing key ideas of the proof that a cylindrically symmetric distribution viewed edge-on has a unique reconstruction.

Theorem 2. *Given a projection I of a cylindrically symmetric density function L along a direction perpendicular to its axis of symmetry, there is no other cylindrically symmetric density function that projects to I; that is, L is unique.*

Proof. This proof involves relationships between several coordinate systems; in the interest of clarity, these coordinate systems are shown in Figure 3.13. The proof that there is a unique solution for this case "breaks down" the cylinder into circularly symmetric layers, demonstrates that each layer can be uniquely reconstructed, and "builds up" a unique reconstruction of the luminosity density distribution. This process is illustrated in Figure 3.14.

1. Let $L(x_w, y_w, z_w)$ be a cylindrically symmetric luminosity density distribution, where x_w, y_w, and z_w are the world coordinates.

2. Let $I(x_v, y_v) = \int_{-\infty}^{\infty} L(x_w, y_w, z_w) dz_v$ be the projection of L along the z-axis of the viewer coordinate system, where x_v, y_v, and z_v are the viewer coordinates.

3. In the edge-on view, $x_w = x_v$, $y_w = z_v$, $z_w = y_v$.

4. Since L is cylindrically symmetric, $L(x_w, y_w, z_w) = f(r, h)$, where $r = \sqrt{x_w^2 + y_w^2}$ and $h = z_w$.

5. Substitution gives $r = \sqrt{x_v^2 + z_v^2}$ and $h = y_v$.

6. Let $I_q(x_v) = I(x_v, q)$ for an arbitrary y-value q.

7. $I_q(x_v) = \int_{-\infty}^{\infty} L(x_v, z_v, q) \mathrm{d}z_v$.

8. $I_q(x_v) = \int_{-\infty}^{\infty} f(r, h) \mathrm{d}z_v$.

9. $I_q(x_v) = \int_{-\infty}^{\infty} f(\sqrt{x_v^2 + z_v^2}, q) \mathrm{d}z_v$.

10. Let $f_q(r) = f(r, q)$, which is circularly symmetric.

11. $I_q(x_v) = \int_{-\infty}^{\infty} f_q(\sqrt{x_v^2 + z_v^2}) \mathrm{d}z_v$.

12. I_q is the Abel transform of f_q. By Lemma 1, I_q uniquely determines f_q for each q.

13. For each q, I uniquely determines I_q.

14. By 12 and 13, I uniquely determines f_q for each q.

15. The set of all f_q suffices to define f.

16. By 14 and 15, I uniquely determines f.

17. f uniquely determines L.

18. By 16 and 17, I uniquely determines L.

$\square$

Given this result, it is important to note that it is applicable only to a cylindrically symmetric distribution viewed edge-on. For other viewing angles, L_y is not guaranteed to be circularly symmetric; thus the above proof is only valid for edge-on projections. Not only is the proof given invalid for other cases, the result does not hold for all viewing angles, as will be shown below.

The second case—that of a projection along the distribution's axis of symmetry—can be called, for convenience, the "face-on view". This, again, refers to a flat disk such as a disk galaxy. I will show that there exist infinitely many distributions that project to a function $\mathcal{P}$ along their axes of symmetry if and only if there is any such distribution; this will show that the reconstruction problem can be ambiguous when cylindrical symmetry is imposed.

Theorem 3. *Given a projection I of a cylindrically symmetric density function L along a direction parallel to its axis of symmetry, there are infinitely many $L' \neq L$ which also project to I.*

Proof.   1. Let L represent a cylindrically symmetric luminosity density distribution, and let I represent its projection along a direction parallel to its axis of symmetry.

2. The viewer and world coordinate systems are identical from this viewing angle (that is, R is the identity matrix).

3. Define a luminosity density distribution L' such that

$$L'(x, y, z) = \begin{cases} -1 & -k < z < 0 \\ 1 & 0 < z < k \\ 0 & \text{otherwise} \end{cases}$$

for some $k > 0$.

4. Let $L''(x, y, z) = L(x, y, z) + cL'(x, y, z)$ for an arbitrary $c \in \mathbb{R}$.

5. The projection of L'' is $I''(x, y) = \int_{z=-\infty}^{\infty} \left(L(x, y, z) + cL'(x, y, z) \right) \mathrm{d}z$.

6. By the properties of the integral, $I''(x, y) = \int_{z=-\infty}^{\infty} L(x, y, z) \mathrm{d}z + c \int_{z=-\infty}^{\infty} L''(x, y, z) \mathrm{d}z$.

7. From the definition of L', it is clear that $\int_{z=-\infty}^{\infty} L'(x, y, z) \mathrm{d}z = 0$.

8. Substituting and simplifying yields $I''(x, y) = \int_{z=-\infty}^{\infty} L(x, y, z) \mathrm{d}z = I(x, y)$.

9. Since $I''(x, y) = I(x, y)$ for each x and y, $I'' = I$; the projection of L'' is identical to the projection of L.

10. Since this holds for any $c \in \mathbb{R}$, there are infinitely many luminosity density distributions which have the projection I.

□

While this proof gives infinitely many alternative density distributions, similar proofs could be constructed that give equally valid transformations of other forms.

I have shown that there is a unique reconstruction for a cylindrically symmetric function viewed edge-on and infinitely many reconstructions for such a function viewed face-on.

These results bookend the range of possible viewing angles; the behaviour of the reconstructions at intermediate viewing angles remain to be explored from a theoretical perspective in our framework, although Zaroubi *et al.* provide an analysis from a Fourier perspective [25] showing that ambiguity is present at intermediate viewing angles and decreases as the angle approaches edge-on. From an empirical perspective, there appears to be some degree of ambiguity for all intermediate viewing angles in our framework, which also decreases smoothly as the viewing angle approaches the edge-on case. Detailed results are shown in Appendix A.1. The empirical results which assume cylindrical symmetry, especially those shown in Section 4.1.2, vividly demonstrate the effects of ambiguity and show why additional assumptions are needed.

3.2.3   The Projection Matrix

I now describe how the properties of the projection matrix G_P relate to ambiguity. In many cases, the reconstruction problem is underdetermined; Section 3.1.4 describes a variety of scenarios in which G_P may be rank-deficient. In these cases, the correct reconstruction cannot be determined without additional assumptions. In the absence of additional assumptions, it is necessary to express the range of possible solutions; this can be done readily using the null space of the projection matrix, which is the set of solution vectors which when multiplied by G_P result in an image vector which is uniformly zero. This is important for two reasons: first, it is important for understanding the origins of artefacts in the three-dimensional reconstructions shown in Chapter 4; second, it is very significant if the three-dimensional model is to be used in its own right. Accordingly, although Chapter 4 only includes tests in which the three-dimensional reconstruction influences the two-dimenional image, test results demonstrating the applications of the null space are included in Appendix A.

Theorem 4. *Given any reconstruction, a solution with an identical projection can be produced by adding any linear combination of vectors in the null space to the original solution.*

Proof. This statement follows from the linearity of the projection process. Assume that $\mathbf{l}'$ can be added to a solution $\mathbf{l}$ without changing its projection. Then the following proof applies:

$$G_P\mathbf{l} + G_P\mathbf{l}' = G_P\mathbf{l}$$
$$G_P\mathbf{l} - G_P\mathbf{l} + G_P\mathbf{l}' = G_P\mathbf{l} - G_P\mathbf{l}$$
$$G_P\mathbf{l}' = \mathbf{0} \tag{3.15}$$

It follows that $\mathbf{v}'$ must be an element of the null space of G_P. Similarly, it can be shown that, given any $\mathbf{v}'$ in the null space of G_P, the projections are identical:

$$
\begin{aligned}
P(\mathbf{l} + \mathbf{l}') &= P\mathbf{l} + P\mathbf{l}' \\
&= P\mathbf{l} + \mathbf{0} \\
&= P\mathbf{l}
\end{aligned}
\tag{3.16}
$$

From these two results, it can be concluded that a vector $\mathbf{l}'$ can be added to any reconstructed $\hat{\mathbf{l}}$ without changing its projection if and only if $\mathbf{l}'$ is in the null space of G_P. The null space of G_P therefore expresses the range of possible solutions which are equivalent (in terms of reconstruction error) to any given base solution $\hat{\mathbf{l}}$. $\square$

There is precedent for the use of the null space to express uncertainty in the theory of inverse problems. For example, Aster *et al.* describe how to use the null space of the forward operator in general terms [2].

In addition to expressing uncertainty, it is possible to use the null space to search for reconstructions with an optimal reconstruction error which also have other qualities. The dimensionality of this search space is often considerably smaller than the original space of possible luminosity density distributions even after symmetry constraints are imposed. This is a great advantage when the search criteria are complex and difficult to optimize. Rather than optimizing over the entire space of possible reconstructions using a difficult and inefficient method, it is instead possible to find one of the projections which is optimal in terms of reconstruction error using a simple method and then search equally good reconstructions for the best according to a complex set of criteria. For example, it may be desirable that the solution approximately correspond to a specific nonlinear parametric function for some values of the parameters; this could not be implemented using linear constraints, but could be incorporated into a more complex objective function. Since, as shown above, all of the solutions found by searching the null space result in an identical reconstructed image, this does not affect the use of three-dimensional reconstructions in modelling two-dimensional images.

3.2.4 Regularization

Regularization is one way of handling the ambiguity in the reconstruction problem. Using an appropriate regularization method is equally important; an inappropriate method guides the solution away from more plausible possibilities. I test two types of regularization in this thesis: L_2-norm regularization and a new regularization method I propose for disk galaxies

or other disk-like objects, which takes advantage of the three-dimensional structure of the solution to produce much better results.

Regularization which minimizes the L_2-norm of the solution vector is a commonly-used technique in many applications. With this regularization method, the objective function is changed from

$$E = \|G_P \mathbf{l} - \mathbf{i}\|_2^2$$

(the sum of squared errors) to

$$E = \|G_P \mathbf{l} - \mathbf{i}\|_2^2 + \lambda \|\mathbf{l}\|_2^2$$

(which trades off reconstruction error with the norm of the solution vector) [2]. This encourages solutions with a small L_2-norm, and discourages extreme values. The effects of L_2-norm regularization on reconstruction with noisy images are shown in Table 4.9, "L_2-Norm Regularization;" although the solutions do not display the extreme values that occur without regularization (shown in Table 4.8), the three-dimensional structures are still implausible. An algorithm for implementing L_2-norm regularization is shown in Section B.2.1.

Regularization which minimizes the L_2-norm of the solution does not take advantage of the three-dimensional nature of the reconstruction problem. A three-dimensional approach can address the problems observed with L_2-norm regularization. The choice of regularization method when using three-dimensional properties depends on beliefs about the three-dimensional structures that occur in nature. For a disk galaxy, a reasonable belief to use in creating a regularization method is that the galaxy is approximately flat. Equatorial concentration regularization encourages "flat" disks by penalizing high luminosity density at positions far from the equatorial plane. With the penalty term, the new objective function becomes

$$E = \int_x \int_y (I(x,y) - \hat{I}(x,y))^2 \mathrm{d}y \mathrm{d}x + \lambda \int_r \int_h 2\pi r h \hat{L}(r,h) \mathrm{d}h \mathrm{d}r. \qquad (3.17)$$

In this equation, r and h are cylindrical coordinates in the representation of $\hat{L}$. For nonzero luminosity density, the penalty is proportional to the square of the distance from the equatorial plane. This is in some sense analogous to a moment of inertia, although the distance h is the distance to a point from the equatorial plane rather than from the axis of symmetry. Multiplication by $2\pi r$ accounts for the fact that a ring of equal luminosity density at a radius of r contains a volume in world coordinates proportional to r. Equatorial concentration regularization improves reconstruction performance dramatically, as shown in Tables 4.10 and 4.11. An algorithm for implementing equatorial concentration regularization is shown in Section B.2.1
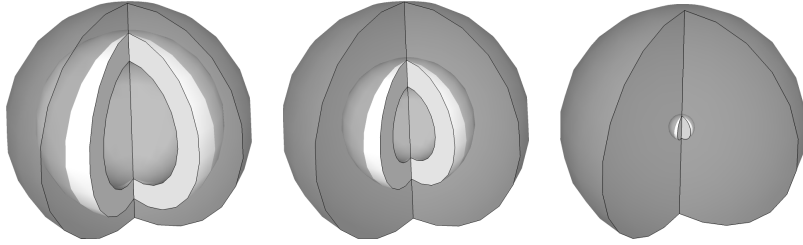
Figure 3.15: Examples of basis functions for a spherically symmetric luminosity density function.

3.2.5 Symmetry Models

An efficient and effective means of imposing symmetry is vital to the practicality of the proposed reconstruction system. In this context, symmetry can be viewed as a series of equality constraints on values of the luminosity density distribution.

An obvious possible method for implementing these constraints is to simply add rows describing each equality constraint to the projection matrix and using the inverse of this larger matrix to find the best-fit solution. Unfortunately, this would not strictly enforce the symmetry constraints. The symmetry constraints would be treated in the same way that the constraints on pixel values are treated, and it would be possible for a trade-off to occur between image reconstruction error and asymmetry, especially if the projection contains noise. Since the desired result is a strictly symmetric luminosity density distribution, I take a different approach.

My approach takes advantage of the equality relationships by defining a new set of basis vectors. Each of these new basis vectors represents a set of regions of space over which luminosity density must be constant. Each basis vector has the required symmetry, so any linear combination must also have it. Examples of suitable basis vectors are shown in Figure 3.15 and Figure 3.16.

Any luminosity density function that can be expressed using one of these models can also be expressed using individual voxels. Since the conversion is a change of basis, there exists a matrix β_{model} for each model such that

$$\mathbf{l}_{voxel} = \beta_{model}\mathbf{l}_{model}. \tag{3.18}$$

Since β_{model} can be used to convert from an arbitrary model basis to a vector representation, this conversion can be incorporated into the projection process:

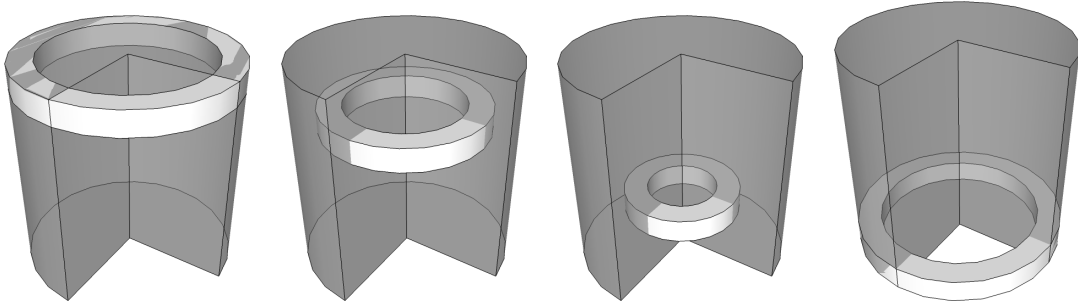$$\mathbf{i} = G_P(R)(\beta\mathbf{l}_{model}). \tag{3.19}$$

Figure 3.16: Examples of basis functions for a cylindrically symmetric luminosity density function.

By the associativity property of matrix multiplication, it is easy to see that

$$\mathbf{i} = (G_P(R)\beta)\mathbf{l}_{model}. \tag{3.20}$$

This method is shown algorithmically in Section B.1.3. In practice, it is generally easier to calculate $G_P\beta$ directly, rather than by matrix multiplication. This matrix, which will be denoted $G_P(R, \beta)$, represents each pixel in the projected image as a linear combination of the model basis vectors.

Because symmetries are defined by a set of symmetric basis functions, it is easy to define new types of symmetry. Even non-symmetric structures can be readily defined, although it is likely that symmetric functions will be the most commonly applicable ones.

One possible method for defining a new structure is to take a set of symmetric basis functions and apply a distortion to them. This distortion need not be linear or affine. Basis functions could be defined this way to represent, for example, spiral arms or bar structures. Boxy and disky elliptical galaxies could also be modelled this way. This would be analogous to the methods used in GALFIT version 3 for two-dimensional image modelling, except it does not assume any specific function profile [17].

3.3 Implementation Considerations

Section 3.2 establishes theoretical results about the projection process, but practical issues remain. For example, the discretization process (necessary to reduce the problem to a finite size) has been referred to but not yet made systematic. Also, it has been established that

the space of possible solutions can be defined using one of the solutions and the null space of the projection matrix, but there has been no discussion of methods for finding the first solution. This section discusses these and other practical matters for the implementation of this reconstruction system.

3.3.1 Discretization and Interpolation

The main difficulty in discretizing the luminosity density function arises from the need to define projection matrices for multiple viewing angles. Rotation will generally result in the need for values of the luminosity density distribution between the discrete grid point, no matter how fine the grid becomes. The luminosity density distribution must be defined continuously to handle these cases correctly. On the other hand, the number of values used to represent the luminosity density distribution must be finite.

Both requirements can be met if it is assumed that the function can be linearly interpolated between a finite number of sampling points arranged on a three-dimensional rectangular grid. Since the grid is three-dimensional, the process is more properly called trilinear interpolation (it is linear in each of three dimensions). Because of its linearity, trilinear interpolation is well-suited to the linear algebra-based approach taken throughout this thesis.

3.3.2 Finding Solutions Despite Ambiguity

The simplest method for finding a base solution of an underdetermined system of equations is the Moore-Penrose pseudoinverse, G_P^+ [21]. This is a generalization of the matrix inverse, applicable in cases where a matrix is not invertible, and even if it is not square. Depending on the number of rows and columns in the matrix, the pseudoinverse may be defined as either $G_P^+ = (G_P^T G_P)^{-1} G_P^T$ or $G_P^+ = G_P^T (G_P G_P^T)^{-1}$.

Left-multiplication by the pseudoinverse functions similarly to left-multiplication by the inverse in a case where the inverse exists. The main difference in practice is that the solution is not unique; the solution with the smallest L_2 norm is produced, but others exist. As described previously, these other solutions can be found using the null space of G_P.

3.3.3 Efficiency

Generating G_P is relatively time-consuming. For each of n^2 pixels, the contributions of n sampling points are added, one at a time, to the projection matrix. This process, as

implemented, is $O(n^3)$. The projection matrix is $n^2 \times m$, where m is the number of basis functions; m may be as large as n^3, so the projection matrix may have as many as $O(n^5)$ elements. Fortunately, however, G_P is very sparse; only $O(n^3)$ of the elements are non-zero, so a sparse representation of the matrix does not have such extreme space requirements.

The projection matrix is not dependent on the data, only on the viewing angle and model, so it is possible to calculate a set of projection matrices and their pseudoinverses in advance. If a very large number of reconstructions are to be performed, these pre-computed matrices may be reused extensively; each new problem instance can be solved with a single (admittedly large) matrix multiplication, which is a great advantage when applying this framework in any context where there are large numbers of instances to reconstruct.

3.3.4 Estimating Viewing Angle

Up to this point, it has been assumed that either the viewing direction is known, or it is unimportant. In practice, however, the viewing direction will not be known *a priori*. Instead, the most likely viewing angle must be inferred from the data. This implies that some measure of the likelihood of a particular viewing angle—or at least the relative likelihood of two different viewing angles—must be defined.

One plausible way to define relative likelihood is to assume that the most likely viewing angle is the one which allows the most accurate reconstruction of the data. This method would have the advantage of making no additional assumptions about the nature of the structure. Unfortunately, there is no guarantee that there will not be many different viewing angles which explain the data equally well. Experimental results indicate that this is the case for cylindrical symmetry, so this method will not be practical.

A method with a solid theoretical basis uses Freeman's generic viewpoint assumption [11]. Under the generic viewpoint assumption, the probability that an object is viewed from a given angle is independent of the structure of the object. This is a reasonable assumption for galaxies; a standard assumption in astronomy and cosmology is that the universe is homogeneous and isotropic (it has no "preferred direction" and "looks the same" from every position). The probability distribution over viewing angles need not be uniform; depending on the symmetries present in the distribution and the way the angle is specified, it may not be. It only needs to be independent of the object's structure. Using this assumption in a Bayesian framework, the probability that a given image $\mathbf{i}$ was generated by projection of a luminosity density distribution $\mathbf{l}$ viewed from an angle θ is:

$$P(\mathbf{l}, \theta | \mathbf{i}) = \frac{P(\mathbf{i} | \mathbf{l}, \theta) P(\mathbf{l}) P(\theta)}{P(\mathbf{i})}. \tag{3.21}$$

The prior probability $P(\mathbf{i})$ of observing a given image is only a scaling factor, not influenced by $\mathbf{l}$ or θ. The probability that a distribution $\mathbf{l}$ projected along a viewing angle θ will produce an image $\mathbf{i}$ after the addition of noise depends on how similar the projection and original image are:

$$P(\mathbf{i}|\mathbf{l},\theta) = \exp -\|\mathbf{i} - G_P(\theta)\mathbf{l}\|^2. \tag{3.22}$$

Integration over θ gives the probability $P(\mathbf{l}|\mathbf{i})$ of a given luminosity density distribution regardless of viewing angle. Also, integration over $\mathbf{l}$ gives the probability $P(\theta|\mathbf{i})$ of a given viewing angle regardless of the luminosity density distribution.

In practice, it is difficult to perform this type of optimization as such. I present a method for approximating a solution. This method relies on the following three assumptions:

- For a given luminosity density distribution, there is only a small range of viewing angles for which its projection approximately reproduces the original image. If the viewing angle at which a distribution best reproduces the image $\mathbf{i}$ is denoted θ^*, $\int_0^{\frac{\pi}{2}} P(\mathbf{i}|\mathbf{l},\theta)\mathrm{d}\theta \approx \int_{\theta^*-\delta}^{\theta^*+\delta} P(\mathbf{i}|\mathbf{l},\theta)\mathrm{d}\theta$ (where δ is some small constant).

- If the best-fit distribution for a given viewing angle θ^* is $\mathbf{l}^*$, there is no other viewing angle $theta$ such that $\|\mathbf{i} - G_P^\theta\| < \|\mathbf{i} - G_P^{\theta^*}\|$.

- The "best" $\mathbf{l}$ is the best-fit solution when some viewing angle θ is assumed; if the problem is ambiguous, it is at least among the set of solutions with the same projection.

Although no proof will be presented for these assumptions, they have been found to be a reliable guide in practice.

Based on the first assumption, the most likely distribution is the one for which the probability peak near the optimal viewing angle for that distribution encloses the largest possible area. It is reasonable to define the following measure of quality for a luminosity density distribution $\mathbf{l}^*$ which best reproduces the original image when projected along a viewing angle θ^*:

$$Q(\mathbf{i}, \mathbf{l}^*, \theta^*) = \exp\left(-\|\mathbf{i} - G_P(\theta^* - \epsilon)\mathbf{l}^*\|^2 - \|\mathbf{i} - G_P(\theta^* + \epsilon)\mathbf{l}^*\|^2\right) \tag{3.23}$$

The constant ϵ should be chosen such that the above equation can be used as a proxy for $\int_{\theta^*-\delta}^{\theta^*+\delta} P(\mathbf{i}|\mathbf{l},\theta)\mathrm{d}\theta$, which in turn is assumed to approximate θ^*, $\int_0^{\frac{pi}{2}} P(\mathbf{i}|\mathbf{l},\theta)\mathrm{d}\theta$.

Based on the second and third assumptions, the quality of a viewing angle estimate can be approximated by the quality Q of the best-fit luminosity density distribution of

that viewing angle. This is similar to Freeman's use of the generic viewpoint assumption in finding the illumination direction and shape of a surface simultaneously [11]. Finding the viewing angle with the highest quality estimate will allow the calculation of the most probable luminosity density distribution.

Initialization is also an important consideration for an implementation of viewing angle estimation; a good initial estimate can dramatically reduce the range of viewing angles that must be searched. This is more strongly dependent on assumptions about the nature of the luminosity density distribution beyond assumptions about symmetry. The case of a disk galaxy will be considered here.

A disk galaxy, as the name suggests, can be thought of as occupying a volume corresponding to a disk, with a thickness much smaller than its diameter. The galaxy will, therefore, project to a roughly elliptical area. A rough estimate of the orientation can be obtained using the ratio of the major and minor axes of the projected ellipse. If the thickness is assumed to be negligible, the aspect ratio of the ellipse determines the viewing angle precisely; since this is not the case, the estimate introduces error. This estimate can be used to initialize a search for a better estimate, but should not be used alone unless the true distribution is known to be either very thin or insensitive to errors in viewing angle estimation. The search space can be reduced by considering the nature of the error; increasing the thickness of the disk can only increase the aspect ratio of the ellipse, not decrease it, so the true value cannot be closer to a face-on view than the initial estimate. Thus, the search only needs to proceed in one direction.

The estimated viewing angle is found using an iteratively-refined grid search. Using a coarse grid, an initial estimate of the maximum can be found. A second grid search can be performed on a smaller interval around the estimated maximum using a finer grid size. This process can be repeated until the grid size becomes sufficiently small. Using this type of search rather than gradient search reduces the chances of missing the global optimum in favour of a local optimum, although it cannot eliminate it entirely. The complete algorithm for estimating the viewing angle is shown in detail in Section B.3.

Where the reconstruction problem is ambiguous, there is no guarantee that the base solution will have the highest quality measurement Q out of all of the possible solutions. In practice, this has not been found to be problematic; using the quality measurement for the base solution provides excellent results and avoids the need to search the null space to find the highest-quality solution.

Chapter 4

Empirical Results

We present empirical results for synthetic data and natural images. With synthetic images, it is possible to compare the reconstruction to ground truth data; this is not possible with a natural image. Synthetic images are therefore used for validation of the accuracy of the reconstruction given my model, and natural images are used to establish that qualitatively reasonable reconstructions are obtained under these assumptions.

In all images shown, black corresponds to high luminosity, a light grey to zero luminosity, and white to negative luminosity. The images are inverted from the natural interpretation for clarity and because there are large areas of zero luminosity which would otherwise be very dark. The "residual" image is the difference of the reconstructed image and the original image; it represents the portions of the image that cannot be explained by the model used.

4.1   Reconstruction with Synthetic Data

In generating synthetic data, it is possible to create images with varying levels of realism. The least realistic images are noise-free projections of simple geometric shapes with uniform luminosity; such images are very different from natural images of galaxies, but are effective illustrations of the operation of our system since the details are readily visible. More realistic are images of distributions based on those believed to correspond well to the distribution of luminosity in real galaxies; we use a mixture of a disk (which has an exponential distribution of luminosity with radius and with height) and a bulge (which is spherically symmetric and projects to a de Vaucouleur's Law surface brightness profile
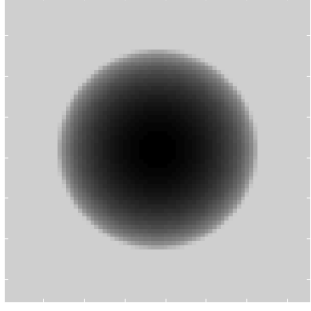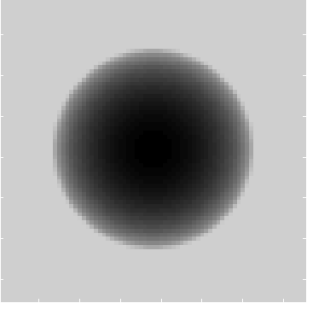
Input	Reconstructed	Residual

Table 4.1: Reconstruction of a spherically symmetric uniform distribution, under the assumption of spherical symmetry. The initial image is noise-free.

[24]). Realism can be improved by adding random noise with a realistic distribution [16] to these projections. Experiments are performed to evaluate the performance of the reconstruction system with each category, and we present the results in order of increasing realism.

4.1.1 Simple Shapes

The following tests are presented to provide examples of the performance of our system using distributions that are easy to visualize. Although not realistic as images of galaxies, these cases are useful conceptual tools for understanding and visualizing the more complex cases discussed later. The shapes used are solid and hollow spheres, thick disks (cylinders with a radius much larger than their height), and rings with a square cross-section.

Sphere

The simplest of all finite, uniform distributions is the spherical distribution. Since it is spherically symmetric, the viewing angle is irrelevant. Table 4.1 shows the performance of the reconstruction framework in the case of spherical symmetry. As expected, the image is reproduced exactly.

Table 4.2 shows the effects of adding Gaussian noise to the projection. Noise added to the projection degrades the accuracy of the reconstruction, although the reconstruction appears better, qualitatively, than the noisy projection. It is interesting that the largest

variations from the desired projection occur near the center of the projection of the reconstructed distribution. This occurs because there are fewer pixels for which the line of sight passes through very small radii than for which the line of sight passes through larger radii. This means that the central region of the distribution is constrained by a smaller number of noisy measurements; it is natural that the resulting estimate is less accurate.

Spherical Shell

A slightly more complex shape is a spherical shell (of non-negligible thickness) surrounding an empty volume. Table 4.3 shows the projection of such a distribution and its corresponding reconstruction. Again, the reconstruction is exact when no noise is present; this is a consequence of the properties of spherical symmetry described in Chapter 4.

Table 4.4 shows the results of the reconstruction process for a spherical shell with varying degrees of noise. The results are very similar to those obtained with a uniform spherical distribution, as shown in Table 4.2.

Thick Disk

The simplest test case shown for cylindrical symmetry is the uniform thick disk. This is an excellent test case for cylindrical symmetry, since a uniform cylinder like a thick disk is the simplest possible case of cylindrical symmetry which is not also spherically symmetric.

Since viewing angle is significant for cylindrically symmetric functions, it is worth revisiting the definition of viewing angle to be used for these cases. The viewing angle θ for a cylindrically symmetric function is defined to be the angle between the direction of projection and the axis of symmetry of the function. This is illustrated in Figure 4.1.

The test data and results are shown for 5 views, ranging from 0° to 90° from the axis of symmetry. Table 4.5 shows, in the column labelled "Input," the projections at each angle, including the edge-on view ($\theta = 90°$) and the face-on view ($\theta = 0°$). No noise is added to these images; these examples are intended to demonstrate performance in the simplest cases.

Table 4.5 also shows the minimum-norm solutions (that is, the solutions with optimal reconstruction error where the L_2-norm of the solution vector is smallest) in face-on and edge-on views. There is a clearly visible tendency for the accuracy of the reconstruction to decrease as the viewing angle approaches 0°. The original image is reproduced exactly when it is noise-free, but the luminosity density distribution found is not the same as the

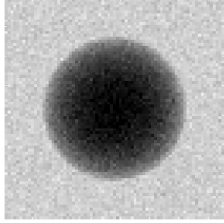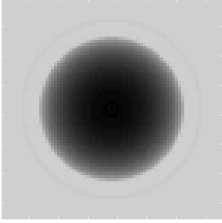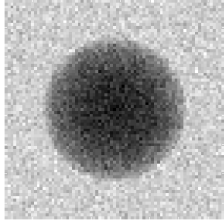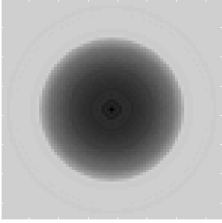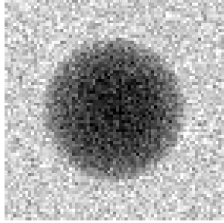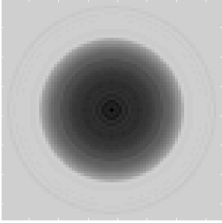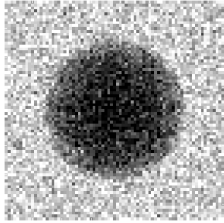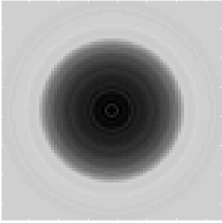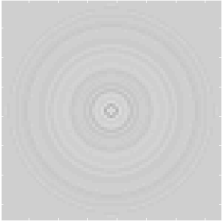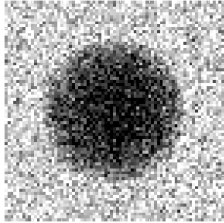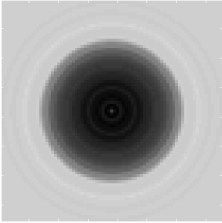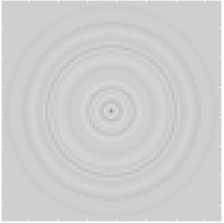| σ	Noisy Input	Reconstructed	Residual
0.05			
0.10			
0.15			
0.20			
0.25			

Table 4.2: Results of the reconstruction process applied to noisy data. The residual is relative to the noise-free projection. The parameter σ is the standard deviation for the Gaussian noise, as a proportion of the value of the brightest pixel in the projection.

47

Input	Reconstructed	Residual

Table 4.3: Reconstruction of a thick spherical shell, under the assumption of spherical symmetry. The initial image is noise-free.
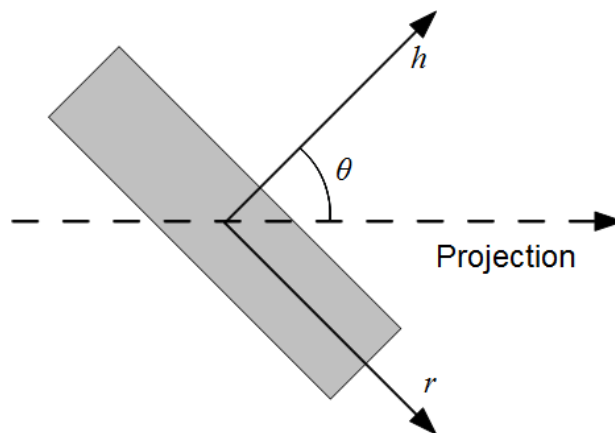
Figure 4.1: Diagram showing the definition of the viewing angle θ for a cylindrically symmetric function.
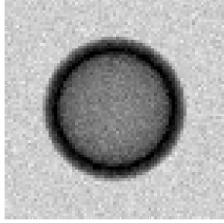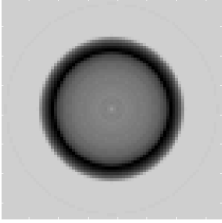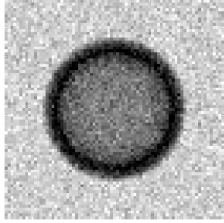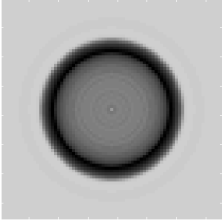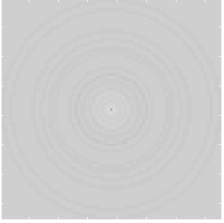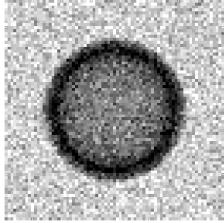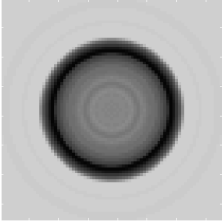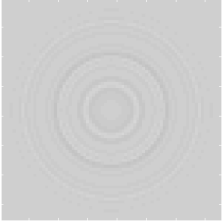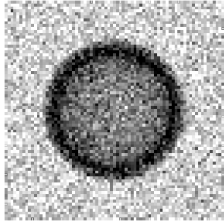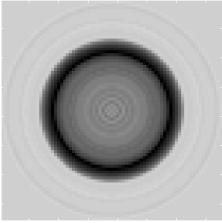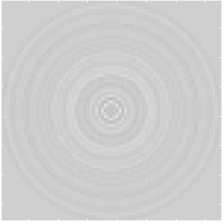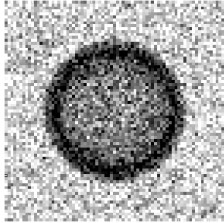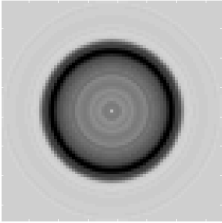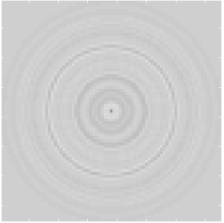
σ	Noisy Input	Reconstructed	Residual
0.05			
0.10			
0.15			
0.20			
0.25			

Table 4.4: Results of the reconstruction process applied to noisy data. The residual is relative to the noise-free projection. The parameter σ is the standard deviation for the Gaussian noise, as a proportion of the value of the brightest pixel in the projection.
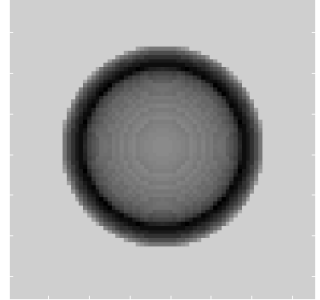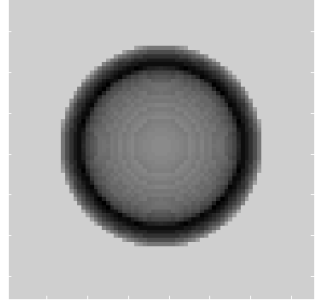
original. This is because the projection matrix represents an underdetermined system of equations; there are infinitely many exact solutions, but only the solution with the smallest L_2-norm is shown. The minimum-norm solution is not necessarily the luminosity density distribution which was used to generate the image. This uncertainty can be expressed using the null space of the projection matrix; these experimental results are described in Appendix A. This problem occurs for all cylindrically symmetric reconstructions, including realistic galaxies. Sections 4.1.2 and 4.1.2 describe unsuccessful and successful attempts to use regularization to address the ambiguity.

Ambiguity is most obvious in the case of a face-on view; there is no evidence about the distribution of luminosity density along the axis of symmetry (because it is parallel to the viewing direction). The minimum-norm solution spreads the luminosity across the entire volume being reconstructed, which has a lower norm than a solution in which the luminosity is concentrated toward the equatorial plane. Since the luminosity is much more concentrated in the luminosity density function originally used to generate the image, the error in the reconstructed luminosity density function is especially large in this case.

Ring

A somewhat more complex example of cylindrical symmetry is a uniform ring with a rectangular cross-section. Unlike many common distributions, such a ring is not convex. Although it is, again, not a realistic astrophysical distribution, such an unusual distribution shows the flexibility of a nonparametric algorithm in terms of the structures that can be reconstructed.

4.6 shows the ring-shaped distribution from a range of viewing angles in the "input" column. The structure is most clearly visible in the face-on view, but the nature of the distribution can also be discerned from the edge-on view.

The reconstructed distributions—found by selecting the minimum-norm solution—for each viewing angle are also shown in Table 4.6. It is interesting to note that the overall quality of the reconstruction appears to be higher than that for the thick disk. This is likely due to the absence of luminosity near the axis of symmetry; in the case of the thick disk, the errors are largely concentrated along the axis of symmetry.

4.1.2   Synthetic Galaxies

The following tests demonstrate the efficacy of our method on more realistic images, which are generated using luminosity density distributions that are based on existing models of
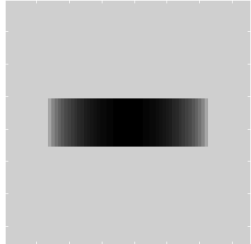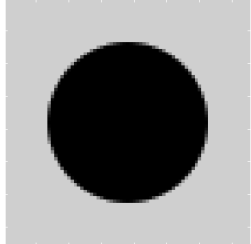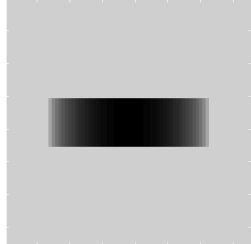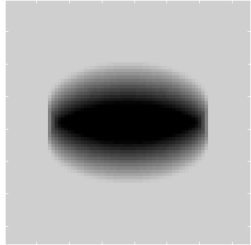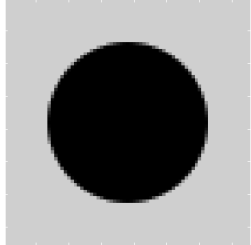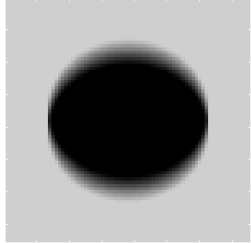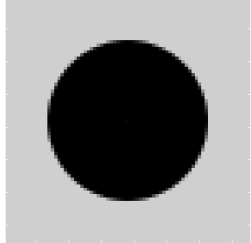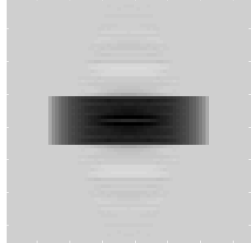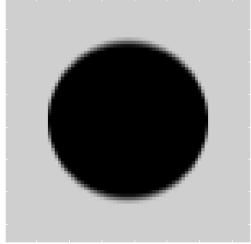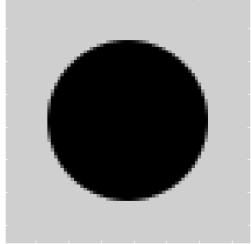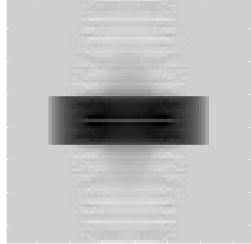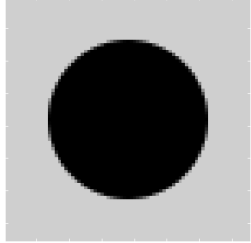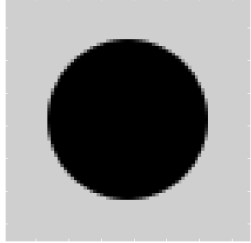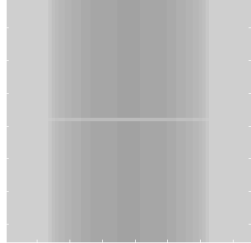
θ	Input	Face-On	Edge-On
90°			
60°			
30°			
10°			
0°			

Table 4.5: Disks reconstructed assuming cylindrical symmetry with reflective symmetry about the equatorial plane. Projections are shown for the face-on viewing angle and the edge-on viewing angle.
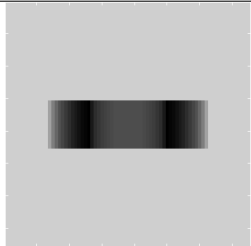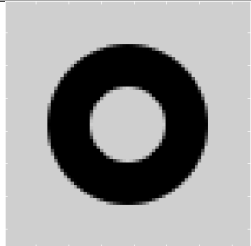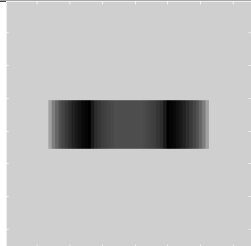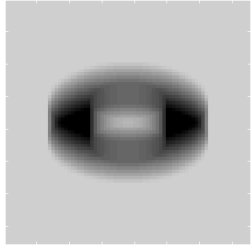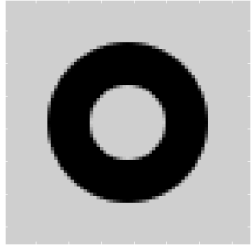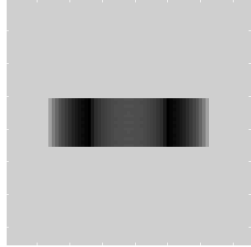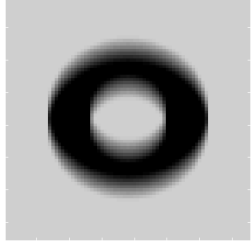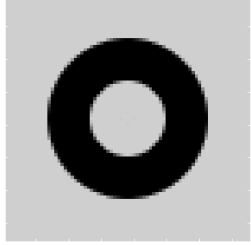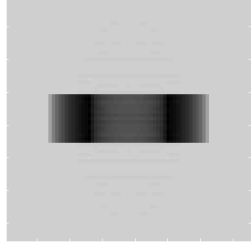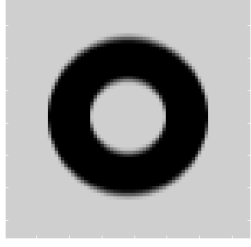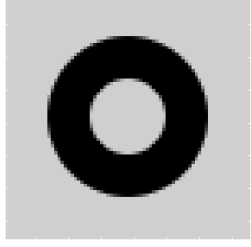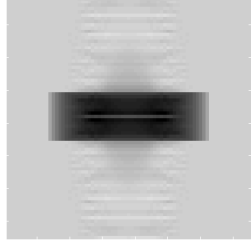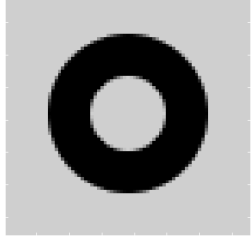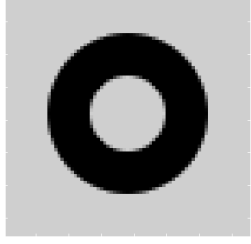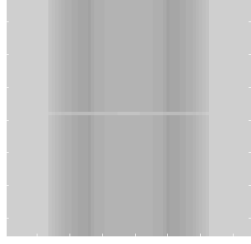
θ	Input	Face-On	Edge-On
90°			
60°			
30°			
10°			
0°			

Table 4.6: Rings reconstructed assuming cylindrical symmetry with reflective symmetry about the equatorial plane. Projections are shown for face-on and edge-on viewing angles.

galaxies. Specifically, the luminosity functions used are the sum of a bulge component and a disk component. The disk component is modelled as decreasing exponentially with radius and with distance from the equatorial plane: $L_B(r,z) = L_{0B}\exp\left(-\frac{r}{r_B}\right)\exp\left(-\frac{|z|}{z_B}\right)$, where L_{0B} is a characteristic luminosity, and r_B and h_B are scale lengths for the radius and height of the disk, respectively. The bulge component is modelled by a distribution that projects to a de Vaucouleur's Law surface brightness profile; this distribution must be found numerically. This model corresponds to an idealized lenticular galaxy (Hubble type S0).

Table 4.7 shows projections of this distribution from a range of viewing angles and the results of its reconstruction. The images are reproduced exactly (so the identical reproduced images are omitted), but the error in the three-dimensional function is significant in some cases. This, again, occurs because multiple solutions exist which reproduce the original image exactly but only the one with the smallest L_2-norm is returned. The errors in this case are greater than for the thick disk, which appears to be due to the high degree of central concentration of luminosity. There is a significant volume in which small quantities of luminosity can be incorrectly distributed, reducing the norm of the solution vector.

Adding Noise

The following tests use the same models as the noiseless images, but after projection a background sky level and realistically-distributed noise are added to the image. These images are the closest of our synthetic data tests to natural images.

In astronomical images, the noise takes on a Poisson distribution [16]. This is because the pixel values are photon counts—generally not very many photons, given the faintness of distant galaxies—and noise in this type of count data has a Poisson distribution. Poisson noise is often approximated using Gaussian noise with a standard deviation proportional to the square root of the pixel value [16]; this is the approach taken here.

Table 4.8 shows the results of reconstruction of noisy images of galaxies. The reconstructed *image* is very close to the original, but includes much of the noise. When viewed from the side, it is clear that the reproduction of noise involves a very complex structure with alternating patterns of very large positive and negative luminosity density. Although the two-dimensional view initially seems reasonable, the three-dimensional view clearly shows severe problems. This is an extreme case of over-fitting; regularization is obviously essential to proceed.

θ	Input	Face-On	Edge-On
90°			
60°			
30°			
10°			
0°			

Table 4.7: Simulated lenticular galaxies reconstructed assuming cylindrical symmetry with reflective symmetry about the equatorial plane. Face-on and edge-on projections are shown.

θ	Input	Reconstructed	Residual	Reconstructed (Edge-On)
90°				
60°				
30°				
10°				
0°				

Table 4.8: Reconstructions of a simulated lenticular galaxy with realistically-distributed noise. The least-norm solution consistent with the minimization of the sum of squared errors is shown. The desired edge-on projection is identical to that shown for the noise-free case in Table 4.7; the edge-on projection reconstructed for the edge-on viewing angle is approximately correct.

L_2-Norm Regularization

Regularization improves the results. Table 4.9 shows the results using L_2-norm regularization (see Section 3.2.4). The tendency toward large positive and negative densities is dramatically reduced. Unfortunately, the reconstructed volume is still very unrealistic. Luminosity density forms a sort of cone—with a sides parallel to the viewing direction—which allows the reproduction of much of the noise in the original image.

Equatorial Concentration: The Importance of Three Dimensions

The use of L_2-norm regularization still does not yield satisfactory results. Although it is performed on variables representing three-dimensional structures, it does not exploit three-dimensional relationships between them. As described in Section 3.2.4, it is possible to define a regularization method that encourages concentration of luminosity density toward the equatorial plane of the distribution by changing the objective function to

$$E = \int_x \int_y (I(x,y) - \hat{I}(x,y))^2 \mathrm{d}y\mathrm{d}x + \lambda \int_r \int_h 2\pi r h \hat{L}(r,h) \mathrm{d}h\mathrm{d}r. \tag{4.1}$$

Varying λ determines the strength of the regularization; depending on the value of λ, luminosity a considerable distance from the equatorial plane may be included in the model if the input image provides sufficiently strong evidence. The reverse is also true: if the evidence for a thick disk is weak, the reconstructed disk will be very thin. These effects both appear in Table 4.10, which shows the results when this type of regularization is performed with a moderate λ.

A significantly higher λ—corresponding to a stronger penalty for luminosity far from the equatorial plane—is used to generate the results shown in Table 4.11. The results are fairly similar, but with a stronger tendency to flatten the distribution, as would be expected. The choice of λ should be based on how flat the luminosity density distribution is believed to be.

For both values of λ, the results are a dramatic improvement over both L_2-norm regularization and the unregularized case in terms of eliminating noise in the reconstructed two-dimensional image and in terms of reproducing the three-dimensional properties of the distribution. This is because the regularization method corresponds to the three-dimensional structure of the image. Despite the simplicity of the penalty term used in regularization—it is convex and adds only a single parameter, λ—it dramatically improves qualitative performance.

θ	Input	Reconstructed	Residual	Reconstructed (Edge-On)
90°				
60°				
30°				
10°				
0°				

Table 4.9: Reconstructions of a simulated lenticular galaxy with realistically-distributed noise. L_2-norm regularization is used. Again, the original luminosity density distribution used is the same as for the noise-free case, so the correct edge-on projection can be found in Table 4.7.

θ	Input	Reconstructed	Residual	Reconstructed (Edge-On)
90°				
60°				
30°				
10°				
0°				

Table 4.10: Reconstruction of simulated lenticular galaxy with realistically-distributed noise. A penalty term is used to encourage concentration to the equatorial plane ($\lambda = \frac{1}{4\pi}$). Again, the original luminosity density distribution used is the same as for the noise-free case, so the correct edge-on projection can be found in Table 4.7.

θ	Input	Reconstructed	Residual	Reconstructed (Edge-On)
90°				
60°				
30°				
10°				
0°				

Table 4.11: Reconstruction of simulated lenticular galaxy with realistically-distributed noise. A penalty term is used to encourage concentration to the equatorial plane ($\lambda = \frac{1}{\pi}$).
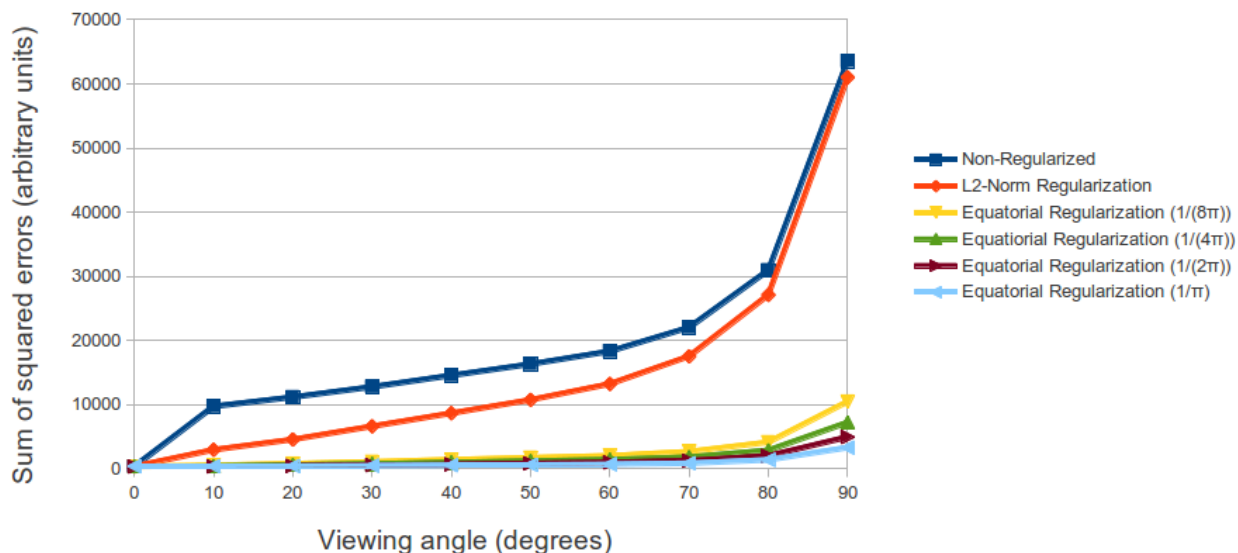
Figure 4.2: Quantitative comparison of the accuracy of the reconstructed image to the noise-free image for simulated lenticular galaxies with realistically-distributed noise. The 95% confidence intervals were calculated, but are much smaller than the symbol size at this scale.

Quantitative performance is also improved. Figure 4.2 shows the relationship between the sum of squared errors between the true noise-free image and the reconstructed noise-free image. Regularization for equatorial concentration dramatically outperforms L_2-norm regularization and unregularized reconstruction on this measure, except for the degenerate face-on case. These differences are statistically significant, with $p \ll 0.001$. The same applies to the sum of squared errors in the three-dimensional representation of the function; these are shown in Figure 4.3.

Assumptions about the three-dimensional structure of the data can produce dramatic improvements in both the three-dimensional reconstruction and the two-dimensional reconstructed image, but only if these assumptions are reasonably accurate. If this is not the case, the results can be misleading, as shown in Table 4.12. These examples are generated using the assumption that the distribution has moderate or strong equatorial concentration, when in fact the distribution is extended far from the equatorial plane. It is unsurprising that the images are reproduced poorly when the assumptions made are so
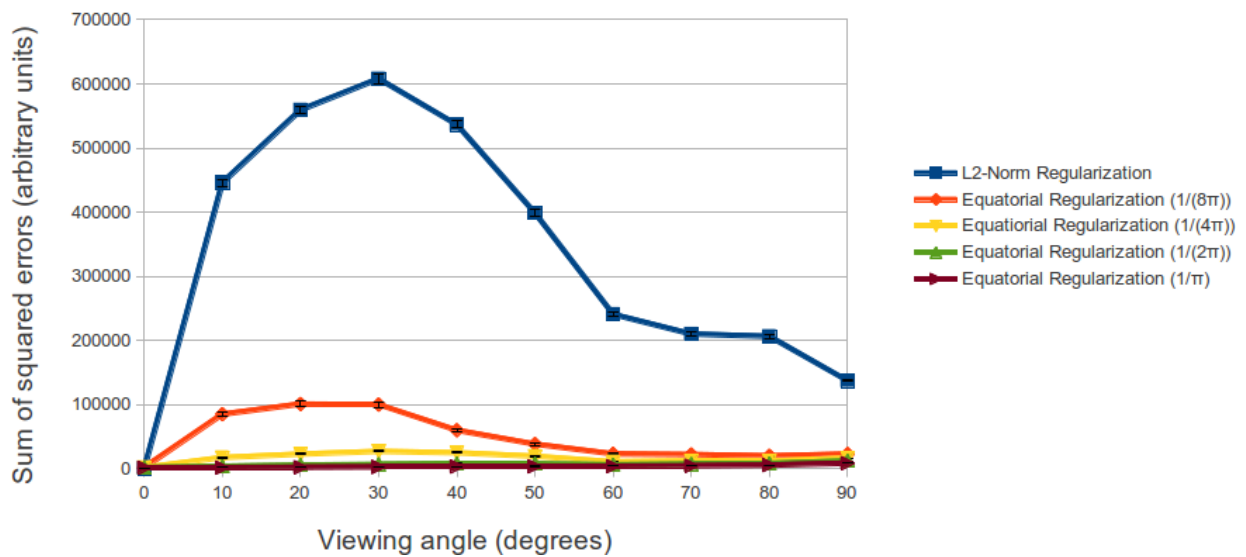
Figure 4.3: Quantitative comparison of the accuracy of the reconstructed volume to the original volume for simulated lenticular galaxies with realistically-distributed noise. The unregularized case is omitted because the values are far higher than for any regularized case (on the order of 10^{22}) and would dominate even a log-scale plot. The 95% confidence intervals are shown, but are often smaller than the symbol size at this scale.

far from the truth.

By finding a best-fit function with a specified type of symmetry and which follows structural assumptions such as equatorial concentration (defined based on a user's domain-specific expertise), it is possible to show which parts of the image are explained under this framework. In this case, the residual is significant in that it shows the portions of the image which cannot be adequately explained given the assumptions made. In the examples shown, these components of the image have been negligible, but in the case of real data, subtle asymmetries can be of great interest. Faint spiral arms, for example, are a sort of asymmetry which would be much more readily investigated when the portions of the image which can be explained by a cylindrically symmetric function are subtracted.

4.2   Viewing Angle Estimation

Estimation of the viewing angle using the generic viewpoint assumption (described in Section 3.3.4) is very effective. Table 4.13 shows the true and estimated viewing angles for a simulated galaxy in the noise-free case. Each estimated viewing angle is selected because it results in a reconstructed distribution that has the smallest change in the image with small changes in the angle at which the distribution is viewed.

True θ	$0.00°$	$30.00°$	$60.00°$	$90.00°$
Estimated θ (S0)	$-1.62°$	$30.00°$	$60.01°$	$90.01°$
Estimated θ (Disk)	$2.72°$	$29.97°$	$60.03°$	$84.48°$
Estimated θ (Ring)	$2.72°$	$30.00°$	$60.02°$	$84.97°$

Table 4.13: True and estimated viewing angles for noiseless images of a variety of 3-D distributions. A maximum range of 0.3 radians is used and the "nearby" viewing angles are 0.025 radians from the estimated angle.

4.3   Reconstruction with Real Data

It is impossible to obtain ground truth data for real galaxies, but the results of the reconstruction algorithm can be evaluated qualitatively. As an illustrative example, the lenticular galaxy NGC4452 4.4 is reconstructed under the assumption of cylindrical symmetry and reflective symmetry about the equatorial plane. The images of this galaxy used here are taken from the Sloan Digitial Sky Survey III (SDSS-III), Data Release 9 [1].
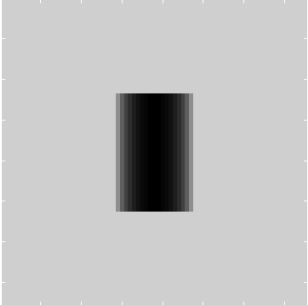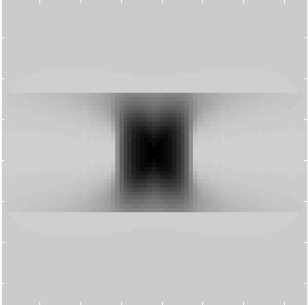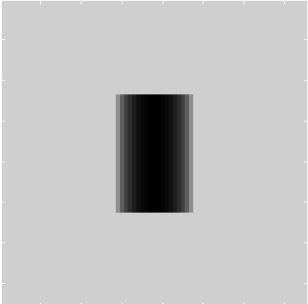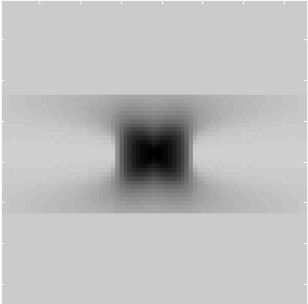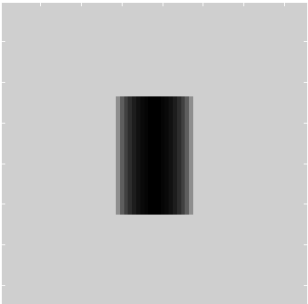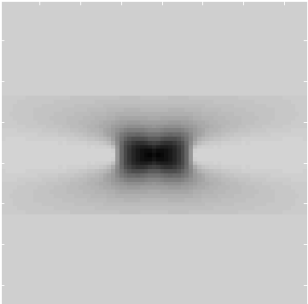
λ	Input	Reconstructed
$\frac{1}{8\pi}$		
$\frac{1}{4\pi}$		
$\frac{1}{2\pi}$		

Table 4.12: Examples of varying degrees of equatorial concentration applied to a "tall" cylinder.
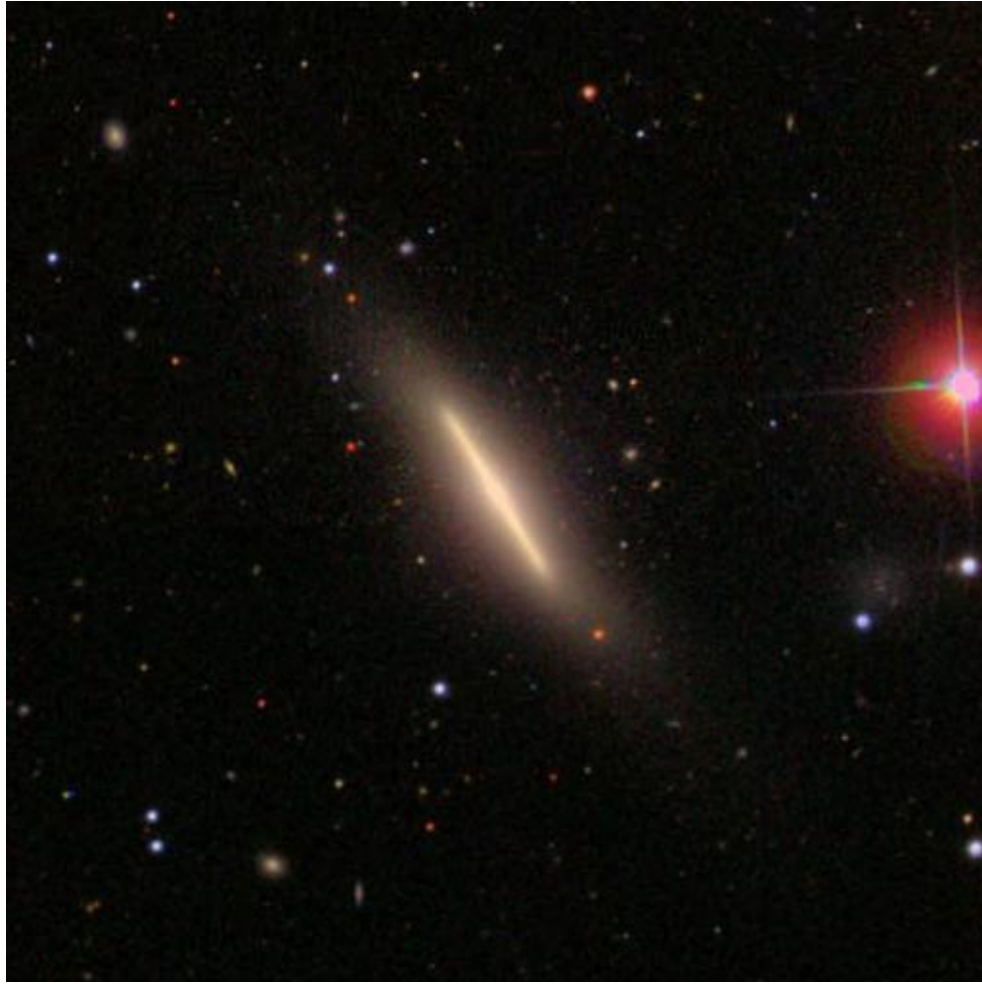
Figure 4.4: Lenticular galaxy NGC4452 (image from SDSS-III [1]).

Data from SDSS-III consists of images in five spectral bands. Since the reconstruction procedure used is monochromatic, the intensity values for the five bands are combined to create a grayscale image. The image is downsampled to 75×75 pixels to reduce the effects of the point-spread function on the image and to improve efficiency; since the original image is blurred by the point-spread function, it is not necessary to apply a Gaussian blur to the image before downsampling. Selection of the center of the object of interest and normalization of the orientation of the major axis of the projected image of the galaxy was performed manually. Viewing angle estimation is performed using the method described

in Section 3.3.4, without applying regularization during the reconstruction process. The estimated viewing angle is 1.5716 radians, or 90.05°.

Table 4.14 shows the results of reconstruction without regularization and with varying degrees of regularization. When no penalty term is imposed, the fitting procedure attempts to include objects surrounding the galaxy; even with relatively minor regularization, these extraneous objects are not included in the model. Strong regularization can cause the model to include less of the weaker luminosity far from the equatorial plane, but the bright regions of the galaxy are reproduced accurately across the entire range of λ tested. One point of interest is that the outer regions of the galaxy appear "twisted" in the projection relative to the inner regions. This is only faintly visible in the original image, but distinct in the residual luminosity; this demonstrates that the image produced is the closest image *that fits the model*.

Table 4.15 shows several different views of each reconstructed galaxy. Without regularization, objects near the galaxy in the image appear as rings in the model. Regularization eliminates these rings and results in a model which resembles the approximately-flat shape that is expected.

The example of NGC 4452 shows that the performance achieved on synthetic data can also be achieved on real data. Although no ground truth data is available for a distant galaxy, the reconstructions found are reasonable when a moderate degree of regularization is applied.

4.4   Comparison with 2-D Image Modelling

Since the three-dimensional reconstruction method models two-dimensional images much differently from existing methods, it is difficult to produce a quantitative comparison between their results. There are qualitative differences between the two systems. Because of these qualitative differences, however, it is worth contrasting the results of two-dimensional image modelling with image modelling through 3-D reconstruction.

The two-dimensional image modelling system used was developed specifically for this comparison, but incorporates methods and functions also used in systems used by astronomers. Developing the system for this specific experiment allows the system to use identical assumptions about the image formation process and image alignment to those used in the three-dimensional case. The two functions used were selected for the case of NGC 4452: an edge-on disk profile (from GALFIT Version 3 [17]) and a Sérsic profile (commonly used, including in GALFIT Version 3 [17]). Optimization is performed

λ	Input	Reconstructed	Residual
0			
$\frac{1}{8\pi}$			
$\frac{1}{4\pi}$			
$\frac{1}{2\pi}$			

Table 4.14: Original and reconstructed views of NGC4452 with equatorial concentration assumed. Each column shows a different view of the reconstructed luminosity density distribution. The strength of the assumption is determined by λ. All reconstructions shown are generated using an estimated viewing angle of 90.05°, obtained automatically.
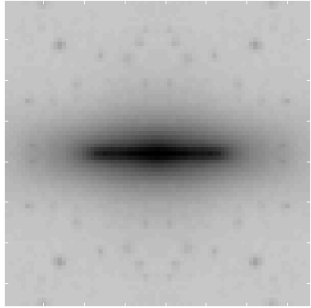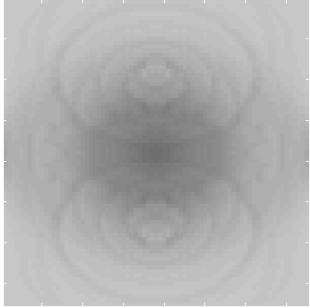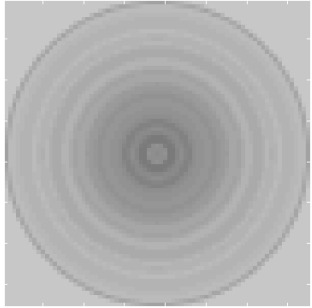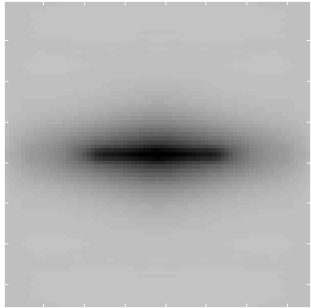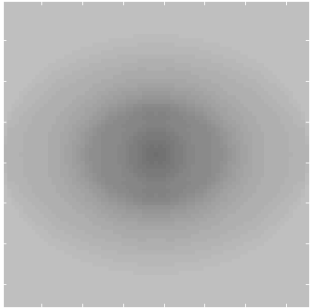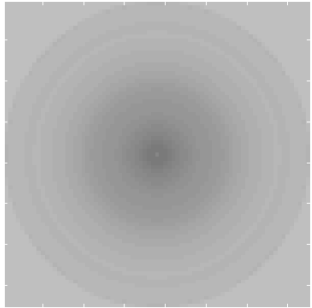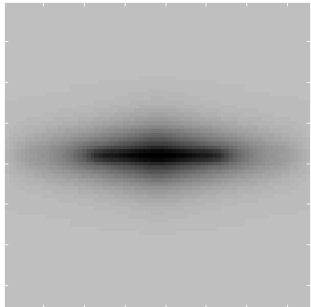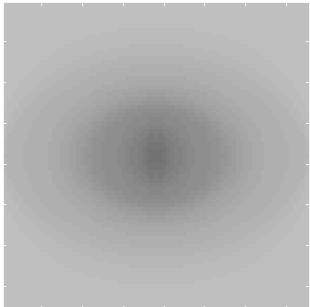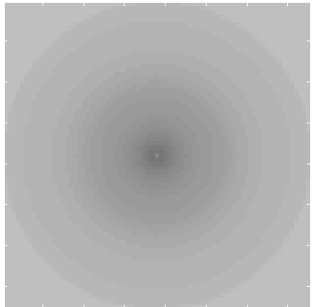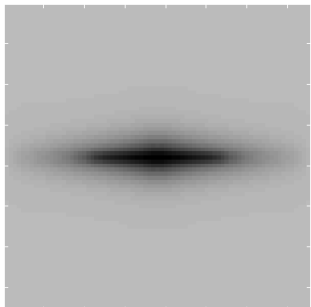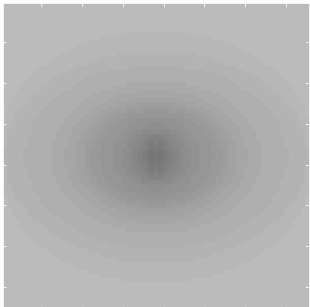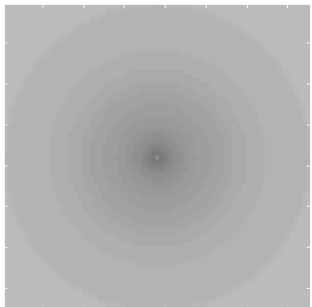
λ	90°-view	45°-view	0°-view
0			
$\frac{1}{8\pi}$			
$\frac{1}{4\pi}$			
$\frac{1}{2\pi}$			

Table 4.15: Reconstructions of NGC4452, assuming equatorial concentration. The strength of the assumption is varied by changing λ. All reconstructions shown use an estimated viewing angle of 90.05°, obtained automatically.

67

using a simple gradient-descent method; initialization is performed manually, and many tests (25, for the results shown here) are performed using random perturbations of these manually-selected conditions.

Table 4.16 shows the examples of results of the two-dimensional image model. The top row shows the results of the three-dimensional method, the middle row shows the generated example with the lowest reconstruction error, and the bottom row shows a different structure that may also result from the optimization process. In both two-dimensional cases there is a significant quantity of luminosity in the image which cannot be accounted for using the given functions, but can be accounted for using the three-dimensional method under the assumption of cylindrical symmetry. This shows one of the important differences between the two methods: the three-dimensional reconstruction method can be used to explain all of the portions of the image consistent with a given type of symmetry, whereas the two-dimensional image modelling methods explain the portions of the image consistent with a specific set of functions. Which method is appropriate to use depends upon the intended use of the reconstructed and residual images and the problem that the user is attempting to solve.

Using additional functions would allow the two-dimensional model to explain more of the image, but would also dramatically increase its complexity because of the number of parameters that would be introduced. Even with only two components, the model requires non-convex optimization and is sensitive to initial parameters. Automatic fitting becomes infeasible for large numbers of components due to the increasing number of parameters; GALFIT [18, 17], which includes a large library of functions, depends to a large extent on human guidance and interpretation to produce results. The three-dimensional method, on the other hand, uses a convex optimization process and does not require initialization, because it is a linear model. With appropriate methods for determining the center and orientation of the galaxy, it can be run entirely automatically.
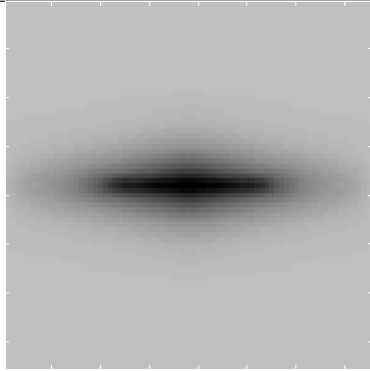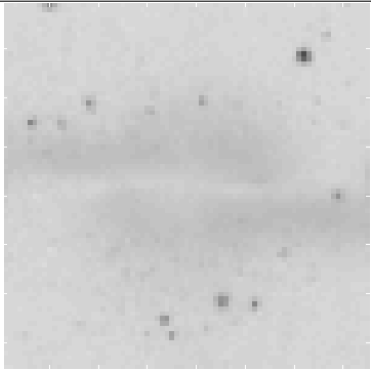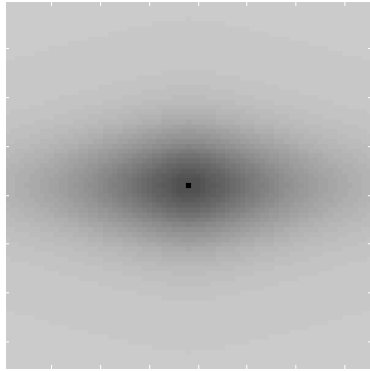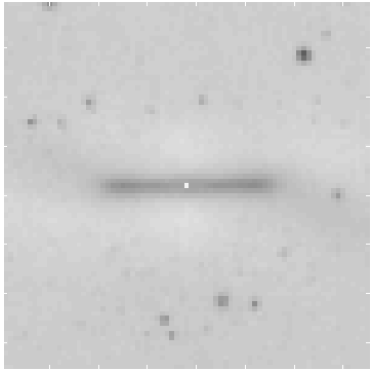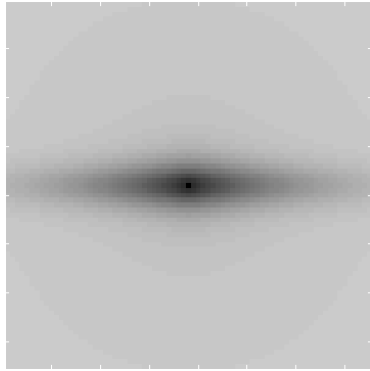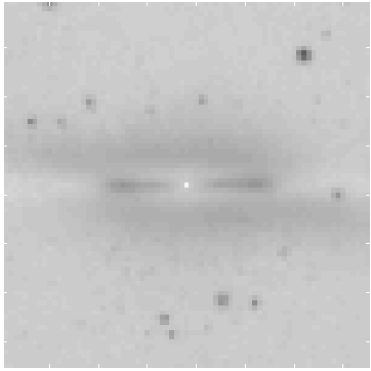
	Reconstructed	Residual
3-D ($\lambda = \frac{1}{4\pi}$)		
2-D (Best)		
2-D (Alternate)		

Table 4.16: Image models of NGC 4452 produced by fitting either a 3-D reconstruction or 2-D functions to the original image, shown with residuals. The top row shows the results obtained using 3-D reconstruction, with an estimated viewing angle of 90.05°. The two examples of 2-D reconstruction used different initial conditions but otherwise identical parameters; the middle row shows the lowest reconstruction error (out of 25 randomly-initialized tests) and the bottom row shows a contrasting structure which can be produced with a different initialization.

Chapter 5

Related Work in Computer Science

Volumetric reconstruction is a problem in computer vision and related fields that has been studied in a variety of contexts. Significant advances have been made due to progress in this area. One notable application in medicine is the computed tomography (CT) scan; this technique is made possible by sophisticated mathematical analysis of the reconstruction problem.

5.1   Inverse Problems

Inverse problems are a very broad class of problems which include reconstruction problems. In an inverse problem, the objective is to determine a set of model parameters $\mathbf{m}$ given a set of observations $\mathbf{d}$ and an operator G such that $G(\mathbf{m}) = \mathbf{d}$ [2]. The opposite problem—determining $\mathbf{d}$ given G and $\mathbf{m}$—is known as the forward problem.

   In general, inverse problems are difficult to solve. Even the operator G may not have an explicit analytical form, let alone G^{-1} [2]. A specific class of inverse problem is the linear inverse problem. In linear inverse problems, $G(a\mathbf{m}) = aG(\mathbf{m})$ and $G(\mathbf{m}_1 + \mathbf{m}_2) = G(\mathbf{m}_1) + G(\mathbf{m}_2)$. This class of inverse problem includes the problem of reconstruction from a projection, as discussed in Chapter 3.

5.2 The Projection-Slice Theorem and Computer Tomography

One of the best-known inverse problems involving projection is *tomography*. Tomography uses a series of projections from different viewpoints to determine the three-dimensional density distribution (usually the opacity of material, rather than luminosity, in most fields where tomography is used). It is best known for its medical applications in CT scanners, but the general technique is used in a wide range of other contexts including archaeology [14] and geophysics [2]. Algorithms used in tomography include Fourier methods related to the projection-slice theorem (which is reliable but slow) and back-projection (a faster but approximate method which is not discussed here).

Computed tomography, and the projection-slice theorem which makes it possible, is a commonly-used technique for estimating a three-dimensional density distribution from a series of projections, much like the projections of galaxies which are discussed here. It requires many projections, but makes no assumptions about the structure. Computed tomography techniques have a wide range of applications—notably in medical imaging—but are not applicable in cases where only one projection is available, such as for astronomical imaging. The reconstruction system described in this thesis requires assumptions about the structure of the density distribution, but can be readily applied when only one projection is available. Although the two methods are applicable in very different cases and operate in very different ways, it is important to describe computed tomography methods in order to establish why the properties of the two approaches differ.

The projection-slice theorem is an elegant and powerful theorem describing the relationship between the Fourier transform of a projection and the Fourier transform of the original function. It states that the Fourier transform of the projection at an angle θ is equal to the slice through the origin at angle θ of the Fourier transform of the original function [7].

Since projections from different angles correspond to different slices of the two-dimensional Fourier transform of the original function, it is possible to use a series of such projections to reconstruct the Fourier transform, and using the inverse Fourier transform reconstruct the original function. In doing this, however, it is necessary to use a substantial number of projections to obtain sufficient resolution in the frequency domain to obtain an accurate reconstruction [7].

In a medical CT scan, the projections are x-ray images of the patient. The images are taken at a range of angles and the three-dimensional distribution of x-ray absorption in the patient's body is reconstructed in accordance with the projection-slice theorem. Since

x-ray absorption depends on the characteristics of tissue, the reconstructed distribution gives physicians a three-dimensional model of the internal structures in the patient's body through a non-invasive procedure. This is obviously a very useful tool. Unfortunately, the impossibility of observing a galaxy from a range of angles makes it impossible to use a directly analogous method to "diagnose" the properties of a galaxy. Tomography is applicable only in situations where many views are available; my framework only requires a single view, and is in important ways complementary to tomography.

5.3   Active Sensors

Computed tomography uses projections at a range of angles to reconstruct a frequency-domain representation of the density distribution and from this obtains a spatial-domain (voxel) representation. There are other methods, however, that use an active sensor to reconstruct the spatial-domain representation of the density distribution directly, without moving the sensor or using multiple sensors. Multiple images are used, which are generated by changing the probing techniques of the active sensor, but structural assumptions can reduce the number of images required [12]. Because an active sensor must be used, however, these methods are impossible to apply to galaxies. Active sensor methods are still worth discussing because of the interesting similarities to and differences from the work described in this thesis; in particular, systems described by Hawkins *et al.* and Gu *et al.* are discussed.

Active sensor methods are intended for mapping the three-dimensional density distribution of a participating medium. The term "participating medium" refers to a medium with a three-dimensional distribution which interacts with light. In the cases examined by Hawkins *et al.* [13] and Gu *et al.* [12], the medium scatters light to a degree proportional to its density.

In the system described by Hawkins *et al*, a laser sheet projected perpendicular to the optical axis of the camera is used to scan the volume in question at successive depth levels. Since, at each stage, only one depth level is illuminated, all of the scattered light reaching the camera must originate at that depth level. The authors explain that in their test cases, it was not necessary to account for the attenuation of light due to the effects of the medium before and after scattering due to the low overall density of their test cases. With insignificant attenuation, the amount of light detected at a pixel (x, y), the position z of the laser sheet, and the scattering properties of the medium are sufficient to determine the density of the medium at the voxel (x, y, z). The density at each point in the medium is calculated in this way to reconstruct the entire density distribution.

This method is not without disadvantages. The scattering properties of the medium must be determined through calibration experiments [13]. Since the fraction of the original light that reaches the camera is relatively small, the laser used in the scanning process must be quite powerful [13, 12]. The use of sequential scanning (with one scan for each depth layer) also substantially slows the image acquisition process; one frame is needed for each depth layer [13, 12]. Also, the need for an active sensor can be considered a disadvantage, since it precludes the use of the technique in many situations.

Some of these disadvantages can be addressed using a technique called compressive structured light, described by Gu *et al.* [12]. This method makes a weak structural assumption about the density distribution (specifically, that it has a sparse representation in the basis used) to reduce the number of images required for the reconstruction.

Compressive structured light uses probe light with randomly-generated patterns which cover the entire volume to be reconstructed to generate each image, rather than scanning individual z-layers. This creates an underdetermined system of linear equations; scattered light from each voxel along a line of sight adds up to the corresponding pixel value, and different voxels are illuminated to different degrees. Since the solution is believed to be sparse, the solution with the smallest L_0-norm (the number of non-zero values in the solution vector) is returned as the reconstruction. The authors indicate that it is reasonable to assume sparsity when the volume in question is largely transparent, with a low density of the medium that interacts with the probe light. Wisps of smoke, for example, occupy a relatively small proportion of a cube that encloses them. Compressive structured light is able to produce accurate reconstructions using as few a one-quarter as many images as the method described by Hawkins *et al.* [12].

Active sensor methods are interesting and have great practical potential, but they are applicable in different situations from the method described in this thesis. Many situations—including astronomy—preclude the use of a laser scanner to acquire images. Active sensors also do not eliminate the need for multiple images; they eliminate the need for multiple viewing angles, but there must be multiple illumination conditions. Interestingly, as in the methods presented in this thesis, structural assumptions can significantly reduce the number of images required.

5.4 Non-Projection-Based Single-Image Reconstruction

A variety of techniques exist for single-image reconstruction in settings where the image is not formed by projection. The image formation models used in these methods are very different from projection, so the methods are not as closely related to my framework as multiple-image projection models are; it is, however, useful to briefly describe some of the techniques for single-image reconstruction that have been used in other contexts.

Zhang and Yuan propose [26] a model called a Nonlinear Deformable Superquadric (NDS), a flexible parametric model for a surface which can be deformed along its length; the model can use any superquadric function [4] as its basis before deformation, and is deformed by a Bezier spline. The NDS can be fitted to the silhouette of an object, but its three-dimensional structure depends on the initial assumptions made about the superquadric on which it is based. Saito and Tsunashima describe [20] a shape-from-shading method that fits a superquadric to the image; a genetic algorithm is used to optimize the parameters of the superquadric. Prasad, Zisserman, and Fitzgibbon describe [19] a system for reconstructing very general surfaces using object boundaries and discontinuities combined with the assumption of smoothness. All of these methods are similar to this framework in that they use a single image to reconstruct three-dimensional structure given some model of the expected structure, but the techniques used to reconstruct an opaque surface are very different from those used to reconstruct a luminosity density distribution.

5.5 A Fourier Method for Cylindrical Symmetry

Zaroubi *et al.* describe [25] a method for reconstructing cylindrically symmetric luminosity density distributions using the projection-slice theorem, emphasizing astronomical applications. Their method is the most similar to the one described in this thesis—it relies on symmetry assumptions and uses only a single projection—but is less versatile.

This method, like computed tomography, is based on the projection-slice theorem. Unlike the computed tomography, however, structural assumptions (specifically, cylindrical symmetry) are used to reduce the number of projections required. A coordinate transformation is used to account for varying viewing angles relative to the axis of symmetry, so it is not limited to the case of an edge-on view as it would be if it simply used the inverse Abel transform on each image row [25].

Ambiguity is noted to be an issue in this case. The authors refer to the ambiguity as the "cone of ignorance". As described in this thesis, the cone of ignorance shrinks to nothing for a precisely edge-on view, and is largest for a precisely face-on view. The extent of the cone of ignorance varies between these extremes; this confirms the experimental results for the dimensionality of the null space of the projection matrix for these viewing angles [25].

The Fourier-domain method used by Zaroubi *et al* is a very different approach from the voxel-based approach presented in this thesis. It is closely connected (through the projection-slice theorem) to conventional computed tomography, which is very interesting. The voxel-based approach does appear to have certain advantages, however. The voxel-based approach operates exclusively in the spatial domain, and allows complex regularization and penalty terms based upon the shape of the reconstructed function to be implemented quite readily. This is what makes regularization to encourage equatorial concentration possible, and as shown in Chapter 4 this type of regularization is essential to obtaining valid results. Also, the equations relating the projection to the reconstructed volume in the Fourier method were derived specifically for cylindrically symmetric density distributions. To adapt the method to other structural assumptions would require extensive modification, whereas the voxel-based method presented in this thesis can use a wide variety of model bases with minimal alteration.

Chapter 6

Related Work in Astronomy

6.1 Astronomy and the Projection-Slice Theorem

One example of a technique from image processing finding important applications in astronomy is the use of the projection-slice theorem to dramatically improve the performance of radio telescopes. In 1956, Bracewell described a sophisticated use of the projection-slice theorem in two dimensions which allowed small radio telescopes to produce results with a resolution equivalent to the that achievable with a much larger telescope using multiple images [5].

The angular resolution of a circular radio telescope (in radians) is inversely proportional to its diameter in wavelengths. To achieve sufficiently high resolution at long wavelengths, however, the telescope may need to be hundreds or thousands of meters in diameter. However, a long, narrow telescope has excellent resolution along its long axis, but poor resolution along its short axis, and a series of images acquired by such a telescope is equivalent to the projection of the full-resolution 2-D image onto a line, as shown in Figure 6.1 [6]. Given a set of these one-dimensional projections obtained at different angles, the projection-slice theorem or back-projection can be used to reconstruct the full-resolution image; this allows a much more practical telescope to obtain equivalent results, albeit more slowly [5, 6].

Although this is a case of reconstructing a two-dimensional function using one-dimensional projections, it demonstrates the use of the general technique of reconstruction from projection where the original data cannot be obtained directly. It also demonstrates the historical connection between computation and astronomy, even as early as 1956.
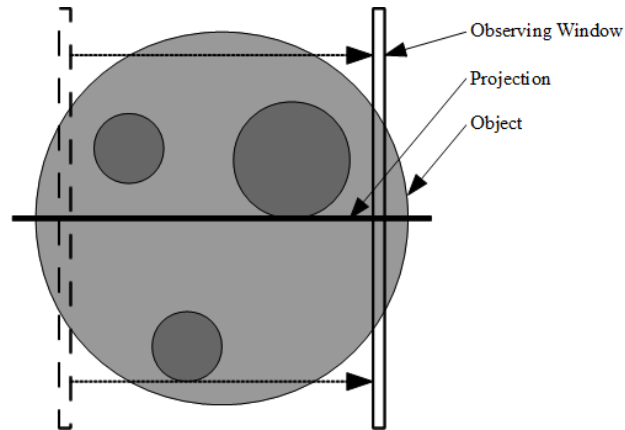
Figure 6.1: A diagram showing the production of one projection from a long but narrow "window" of observation. Each window gives one point in the projection; sliding the window across the object gives the entire projection.

6.2 Surface Brightness Profile Estimation and Bulge-Disk Decomposition

A commonly-used method for quantitatively describing the structure of a galaxy is its surface brightness profile. Generally, a model of a surface brightness profile is fit to the image of the galaxy. It is important to note that these methods are two-dimensional; they describe the projected image, rather than directly describing the three-dimensional structure of the galaxy. This section describes various models and the methods used for fitting them to image data.

6.2.1 One-Dimensional

In the ideal case, it would be possible to extract a one-dimensional profile from the image. This profile represents surface brightness as a function of radius. If a clear one-dimensional profile can be extracted, the model fitting process can proceed in one dimension.

Baggett, Baggett, & Anderson describe [3] a fitting method for the one-dimensional case using a bulge-and-disk model. This fitting method uses a ring-shaped (inner-truncated) model for the disk, which allows a clearer distinction in the model between bulge and disk. This was found to be a useful feature for many spiral galaxies in preliminary studies. In

the tests reported, one-dimensional brightness profile was found using the profile along the major axis of the image of the galaxy, to minimize distortion caused by viewing the galaxy obliquely.

6.2.2   Two-Dimensional

Unfortunately, it is often not possible to extract a one-dimensional profile to which to fit a model. If only the profile along the major axis is considered and the rest of the image data is discarded, noise in the image can overwhelm subtle distinctions between model components, such as a slice through a spiral arm being interpreted as a sequence of pixels made slightly brighter by noise. For galaxies which are not precisely symmetric, it is also difficult to define the "real" profile; one-dimensional profiles differ to a degree, and selecting a single profile to which the models will be fitted introduces bias into the process. For these reasons, surface brightness models often use the entire image, even though the need to fit parameters such as aspect ratios results in a higher-dimensional optimization problem.

Byun and Freeman describe [9] a 2D method for fitting a bulge-disk model to a galaxy image. This method uses the following definitions for the bulge surface brightness μ_B and disk surface brightness μ_D:

$$\mu_B(x,y) = \mu_{eff} + 8.325 \left(\left(\frac{\sqrt{(x^2 + y^2)(b/a)_B}}{R_{eff}} \right)^{0.25} - 1 \right) \tag{6.1}$$

$$\mu_D(x,y) = \mu_0 + 1.0857 \left(\frac{\sqrt{(x^2 + y^2)(b/a)_D}}{R_h} \right) \tag{6.2}$$

In these equations, R_{eff} represents the effective radius of the bulge, μ_{eff} represents the surface brightness at the effective radius, R_h represents the scale length of the disk, μ_0 is the central brightness of the disk, x and y represent the image coordinates of a point along the major and minor axes respectively, and $(b/a)_B$ and $(b/a)_D$ represent the ratios between the major and minor axes of the bulge and disk respectively [9].

De Souza, Gadotti, and dos Anjos described [10] a system for performing 2D bulge-disk decomposition. Only the bulge and disk components are modelled, not spiral arms, bars, or other smaller structures; the system is intended to isolate these substructures, not to model them. The disk component is modelled with an exponential function:

$$I_d(r) = I_{0d} e^{\frac{-r}{h}} \tag{6.3}$$

where I_{0d} represents the central surface brightness of the disk and h represents the scale length of the disk [10]. The bulge component is modelled with a Sérsic profile:

$$I_b(r) = I_{0b}10^{-b_n\left(\frac{r}{r_{eb}}\right)^{\frac{1}{n}}} \tag{6.4}$$

where I_{0b} represents the central intensity, r_{eb} represents the effective radius, n is a parameter determining the rate at which intensity drops, and $b_n \approx 0.868n + 0.142$ ensures that half of the total intensity of the bulge is inside the effective radius.

In 2002, Peng *et al* introduced the GALFIT algorithm [18]. Later, in 2010, the same group introduced a more sophisticated and more general version of the algorithm, GALFIT Version 3 [17]. GALFIT Version 3 is a very sophisticated and versatile suite of functions, which is capable of fitting a wide variety of galaxy profiles very accurately; this is because it contains both a variety of profiles and a variety of distortion modes which can be applied to the profiles.

The profiles used in GALFIT Version 3 include the common Sérsic profile, narrow Gaussians, and many other, less common, profiles such as the modified Nuker profile. All of these profiles fulfil roles similar to the profiles already described; there is no need to provide complete definitions for all of them here.

In the original GALFIT, each component is transformed from circular to what the authors describe as a generalized ellipse. The radial coordinate of a point (x, y) is defined by:

$$r = \left(|x|^{c+2} + \left|\frac{y}{q}\right|^{c+2}\right)^{\frac{1}{c+2}} \tag{6.5}$$

The parameter q is the aspect ratio of the ellipse. The parameter c determines the shape of the generalized ellipse; for $c = 0$, the radius corresponds to a true ellipse, for $c < 0$ the generalized ellipse is "disky", and for $c > 0$ the generalized ellipse is "boxy" [18].

In addition to the use of components in the shape of a generalized ellipse, GALFIT 3 allows the shape of a component to be modified in several ways to account for complex components. Fourier modes, bending modes, coordinate rotation using a power-law-hyperbolic tangent, and coordinate rotation using a logarithmic-hyperbolic tangent are all available distortions of components. Fourier modes produce "bulge-like" or "petal-like" shapes in a component. This modification is defined by the equation

$$r(x, y) = r_0(x, y)\left(\sum_{m=1}^{N} a_m \cos(m(\theta + \phi_m))\right) \tag{6.6}$$

79

Bending modes distort the component shape in a variety of ways; shear is a special case where $m = 1$. This transformation operates on the y-axis; it may be necessary to rotate the image, apply the transformation, and rotate it back if the desired axis does not align with the original y-axis. The equation for this modification is:

$$y' = y + \sum_{m=1}^{N} a_m \left(\frac{x}{r_{scale}} \right)^m \tag{6.7}$$

Bending modes and Fourier modes can be combined to produce complex shapes. Another deformation, coordinate rotation, allows the algorithm to model spiral structures, such as spiral arms. A logarithmic spiral is useful locally in many spiral galaxies, but a single logarithmic spiral is often a poor fit for the full length of the spiral arms, so the logarithmic spiral must be modified. This is done by using a variation on a hyperbolic tangent function to transition smoothly from an inner constant function to an outer power-law or logarthimic function. A power-law spiral is also available, where the power is a free parameter; if the power is chosen carefully, this can approximate a logarithmic spiral when it is appropriate [17].

Optimization of a GALFIT model with respect to the χ^2 metric is difficult, since the number of parameters easily becomes large (sometimes over 40) and the problem is not convex. GALFIT uses the Levenberg-Marquardt method to find the parameters that minimize the χ^2 metric. GALFIT requires significant human intervention to determine the number of components and guide the fitting process [18]. This is an important disadvantage, as it slows the application of the algorithm to a large dataset, consisting of thousands or hundreds of thousands of galaxy images, significantly.

6.3 Reconstruction in Three Dimensions

In a May 2012 e-print [23], Tempel *et al.* describe a 3-D fitting method. This method uses single images, and fits two components (3-D analogues of the Sérsic function commonly used in 2-D fitting) to the image by modifying the parameters of the functions. As this system uses an additive bulge-and-disk model, it has many similarities to the one-dimensional and two-dimensional surface brightness profile estimation methods discussed above [23].

The luminosity density function used to model each of the two components (one for the bulge and one for the disk) is the Einasto's law distribution:

$$l(a) = l(0) \exp\left(-\left(\frac{a}{ka_0} \right)^{1/N} \right) \tag{6.8}$$

In this equation,

$$l(0) = \frac{hL}{4\pi q a_0^3} \qquad (6.9)$$

where L represents the overall luminosity of a component, q represents the ratio of axes of the ellipsoid, and $a = \sqrt{r^2 + \frac{z^2}{q^2}}$. Cylindrical symmetry about the z-axis is assumed for these components [23].

Fitting is performed using a simulated annealing algorithm to minimize the χ^2 error measure,

$$\chi^2 = \frac{1}{N_{dof}} \sum_v \sum_{(x,y) \in \text{mask}} \frac{(f_{obs}^v(x,y) - f_{model}^v(x,y))^2}{\sigma^v(x,y)^2} \qquad (6.10)$$

In this equation, v represents spectral band (one of u, g, r, i, and z). Furthmore, $(f_{obs}^v(x,y)$ and $f_{model}^v(x,y)$ represent observed and predicted surface brightness at the pixel (x,y) and spectral band v, respectively. Finally, $\sigma^v(x,y)$ represents the Poisson error surface brightness measurements at the pixel (x,y) in spectral band v [23].

Tests are presented using synthetic images generated based on SDSS Data Release 8 data. Images are generated through projection of luminosity density to produce a surface brightness image. The paper mentions the possibility of modelling ring structures and dust, but notes that the dataset used is insufficiently detailed for this to be practical. The SDSS point spread function and noise levels are applied to the projections to simulate the properties of SDSS images [23].

The reconstructed bulge/disk ratios are not very accurate, although there is no significant systematic bias. The estimated inclination angle is usually within five degrees of the true angle, provided that the disk is reasonably prominent. Structural parameters of the bulge and disk (the characteristic radius a_0 and the index N) are reconstructed with reasonable accuracy [23].

The Fourier-based system described by Zaroubi *et al.* is more general in that it assumes cylindrical symmetry but makes no other assumptions about the luminosity density distribution. This method was designed for applications in astronomy, but does not make any assumptions that are specific to galaxies, so it was discussed in Section 5.5.

The method described by Tempel *et al.* is strictly less general than the methods proposed in this thesis because the components used have an inherent symmetry (specifically, cylindrical symmetry). The method described by Zaroubi *et al.* assumes only cylindrical symmetry, but cannot be as readily adapted to other forms of symmetry as the method described here can be because the reconstruction method was derived specifically for cylindrically symmetric functions.

Chapter 7

Conclusion and Future Work

I have addressed the problem of three-dimensional reconstruction from a single projection to explain features of the original projection given assumptions about the three-dimensional structure of the original density function. Specifically, the problem was considered in the context of telescopic images of galaxies. The reconstruction system presented in here solves this problem using a spatial-domain representation of the luminosity density distribution (as opposed to a frequency-domain representation) and techniques from linear algebra. Assumptions about symmetries reduce the number of unknowns to a manageable degree, and remaining uncertainty can be explicitly expressed using the null space of the matrix describing the projection process. Penalty terms allow general classes of structure to be encouraged in the reconstruction process, without which the reconstruction is likely to be poor despite the reduced degree of ambiguity.

Briefly, the main contributions of this thesis are as follows:

- Providing a far more adaptable framework for single-projection reconstruction than was previously available

- Proving the properties of the framework for a variety of practical cases

- Expanding the range of possible techniques for investigating the structure of galaxies by allowing astronomers to find best-fit images compatible with a range of highly expressive models

The framework I have described is a significant contribution to computer vision in that it allows significantly more complex inference about three-dimensional structures when

only a single image is available. This case has been largely ignored in computer science, and although work has been done with the technique in astronomy my framework is much more versatile and expressive because it can use a wide range of symmetry constraints and regularization methods. Since any image formed by a process that can be modelled as a purely light-emitting or purely light-absorbing process can be interpreted as a projection, the framework is applicable in many areas beyond astronomy. For example, X-ray images are formed by a light-absorption process; denser materials absorb more of the incident X-rays than less dense materials. There are many types of X-ray images, including medical images and images taken of inanimate objects for inspection purposes; a framework that allows inference from a single image is clearly useful. Because many properties of the framework are proven, the limits of inference within the framework are clearly defined. It is possible for users of the framework to determine if multiple interpretations of the image (or a part of the image) are compatible with their assumptions about the data. Perhaps more importantly, users can be confident that, if the framework indicates a unique solution, there are no other solutions which satisfy their assumptions.

From the perspective of astronomy, this work expands the range of possible analyses of the structure of galaxies. Because of its versatility, the framework described here enables a user to determine which aspects of an image of a galaxy can be explained by a luminosity density distribution with specified symmetry and other structural properties (as in Table 4.14). Whereas most existing techniques for modelling images of galaxies use specific two-dimensional functions and two-dimensional properties, the new reconstruction-based method uses three-dimensional properties that galaxies are believed to exhibit. Thus, the new reconstruction framework expands options for analysis.

There are a wide range of directions for future work. There are fundamental aspects of the framework that can be developed further; for example, there are a wide range of forms of symmetry and regularization which have not been explored in this thesis. A systematic study of the full range of models and the corresponding uniqueness properties would be an interesting and valuable research topic in computer science. Detailed investigation of applications to areas other than astronomy, such as X-ray images, would be another valuable area for further research.

Considering applications to astronomy in particular, there are many other possible areas for research. The framework might, for example, be adapted to account for optical effects such as the point spread function, or to combine information from multiple spectral bands. Further evaluation of the method might be possible if a problem is found where both existing parametric methods and the reconstruction-based framework can be applied to the same problem and the quality of results compared. Naturally, a very important direction for future research is the application of the framework to problems in astronomy.

Appendix A

Applications of the Null Space

The null space of the projection matrix is an important factor in the reconstruction process. Since the best-fit solution for a spherically symmetric luminosity density is always unique, the null space for such a projection matrix is always empty. Since an empty null space is significant in that it indicates that the best-fit solution is unique but is not otherwise particularly interesting, this section focuses on the projection matrices for cylindrically symmetric functions with symmetry about the equatorial plane. The thrust of the arguments about the use of the null space would apply to other structures which the null space of the projection matrix is nonempty, with suitable changes to the descriptions of specific structures.

A.1   Singular Value Spectrum

The singular value spectrum of the projection matrix is of great significance in the case of cylindrical symmetry. Figure A.1 shows the spectrum for a variety of viewing angles. The largest singular value has rank 1; the smallest singular value has the largest rank. There is a clear point for each viewing angle where the singular values drop dramatically. Although the singular values after this point are not zero, this appears to simply be a numerical matter. It is reasonable to consider these extremely small singular values to be zero; these are therefore the singular values corresponding to the null space.

The proportion of nonzero singular values can be seen to increase as the viewing angle approaches 90°, where all of the singular values are nonzero. Since the dimensionality of the null space is equal to the number of singular values which are equal to zero, this indicates
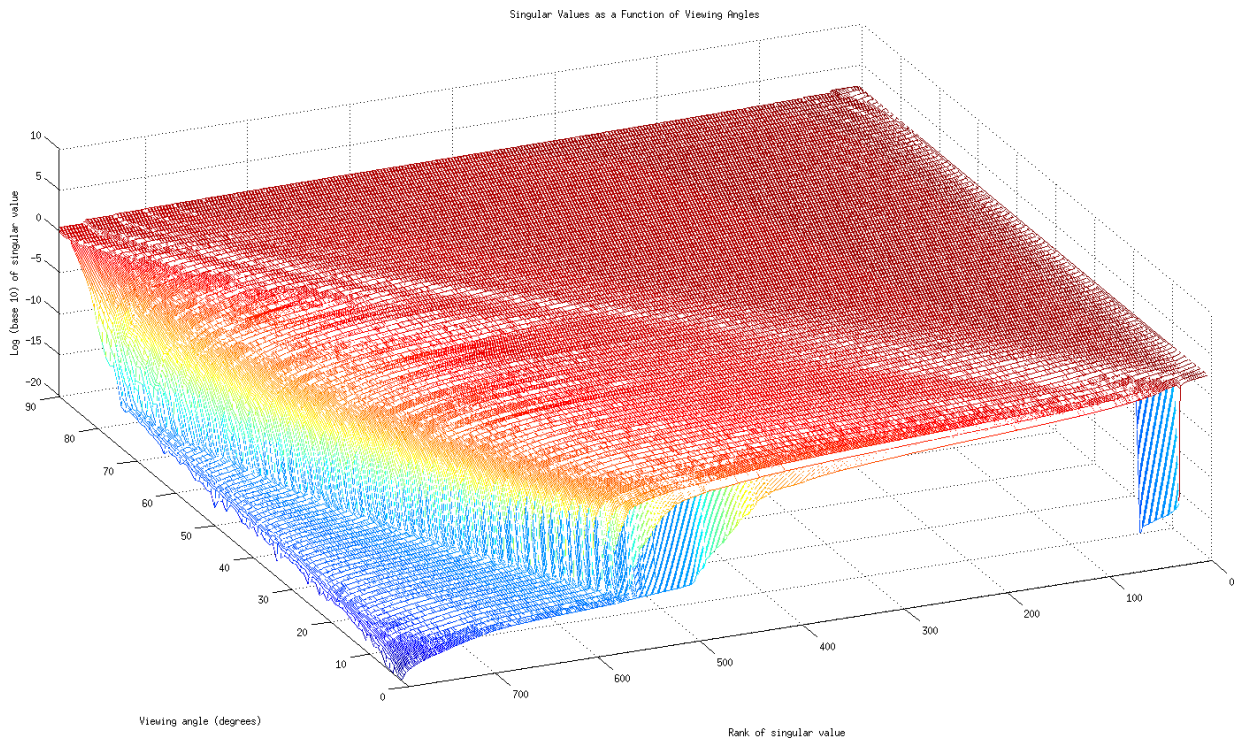
Figure A.1: Spectrum of singular values for a projection matrix corresponding to a cylindrically and reflectively symmetric luminosity density distribution. The singular values are shown on a log scale.

decreasing ambiguity. This is as expected, given the experimental and theoretical results already shown. At the other extreme, most of the singular values for a viewing angle of 0° are below the lower limit of the graph.

A.2   Null Space Vectors

In showing the null space vectors, it is helpful to reduce the resolution. If the resolution is very high, the number of null space vectors becomes unwieldy; a low resolution is convenient for demonstration purposes. The null space vectors are shown as projections of the corresponding volumes in the face-on and edge-on directions. This allows a better comparison between the null space vectors and reconstructions, although only general comparisons can be made due to the difference in resolution.
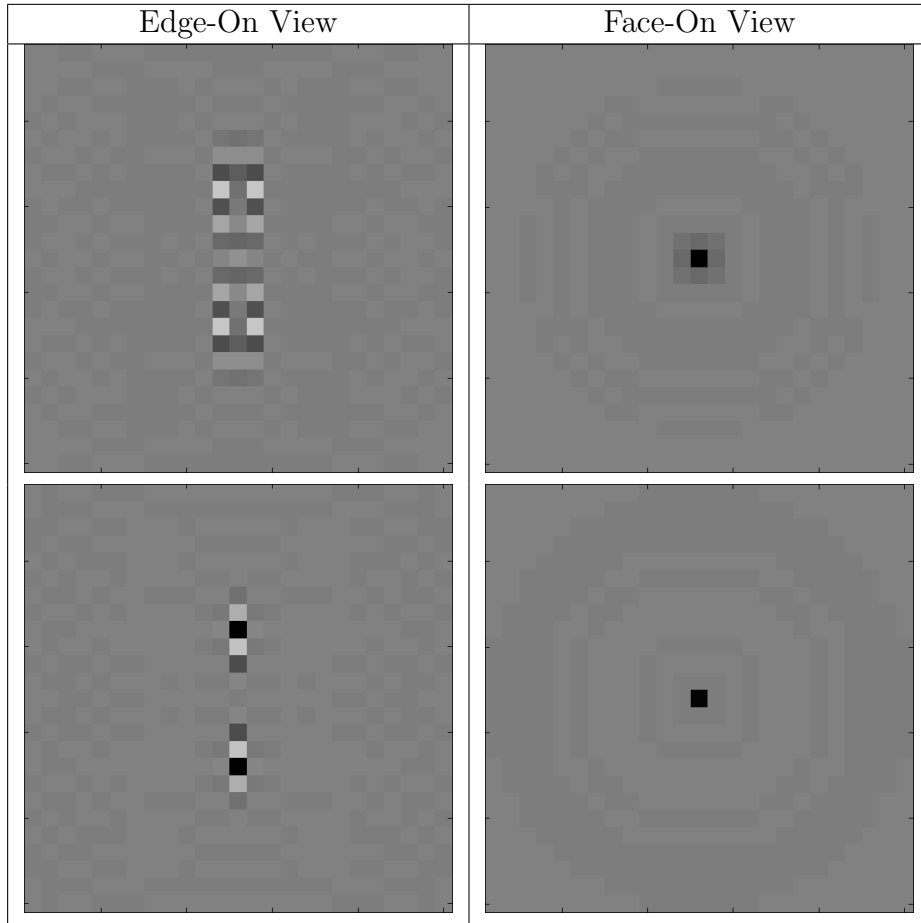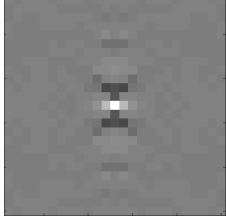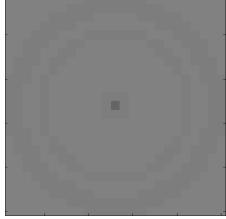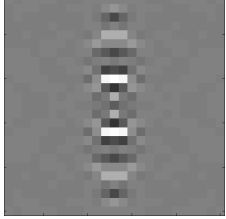
Edge-On View	Face-On View
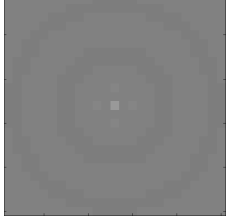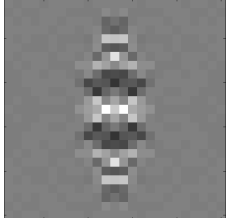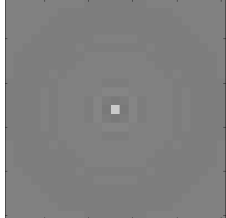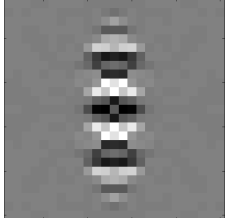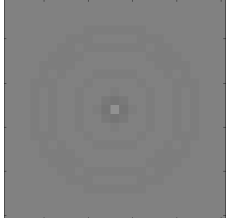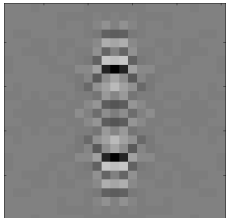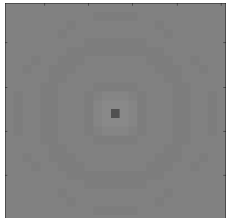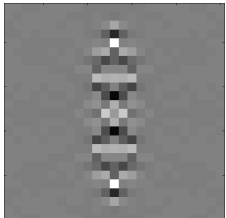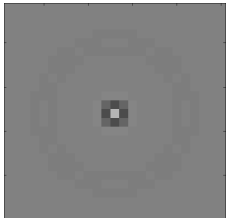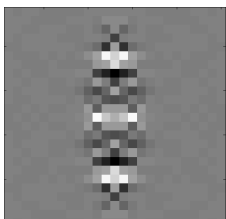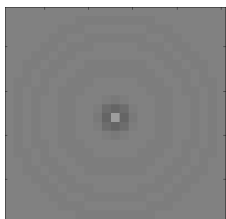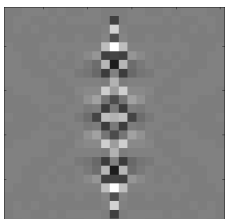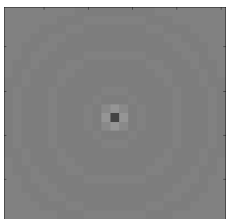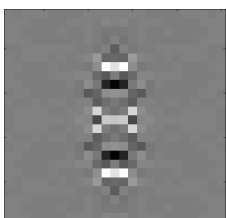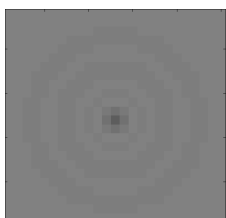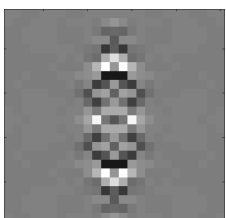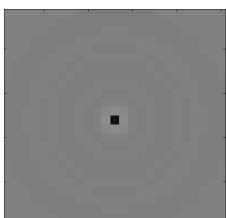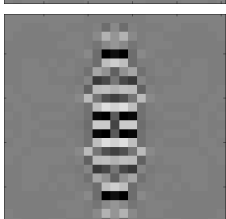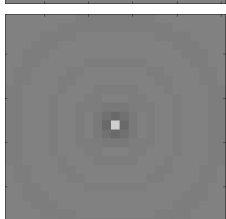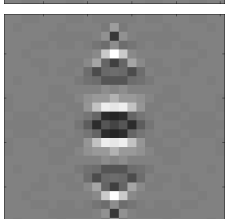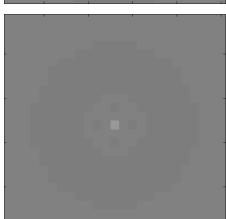
Table A.1: Null space vectors for a 60° viewing angle projection matrix, shown as projections of the volume form. The projection matrix assumes cylindrical symmetry and reflective symmetry about the equatorial plane.

For a viewing angle of 60° and an image side length of 25, the null space has only two dimensions (with basis vectors shown in Table A.1). Both basis vectors have a similar form: a core of alternating positive and negative values near the axis of symmetry, with only slight variation further from the axis. These features are also reflected in the reconstruction residuals, suggesting that the residuals can indeed be explained by a linear combination of null space vectors.

Edge-On View	Face-On View	Edge-On View	Face-On View

Edge-On View	Face-On View	Edge-On View	Face-On View

Table A.2: Null space vectors for a 30° viewing angle projection matrix, shown as projections of the volume form. The projection matrix assumes cylindrical symmetry and reflective symmetry about the equatorial plane.

For a viewing angle of 30°, the null space has a much higher dimension (17). Basis vectors for the null space of this projection matrix are shown in Table A.2. The same patterns observed in the smaller null space of the 60° are present here.

A.3 Exploring the Null Space

In the examples shown, the quality of the reconstruction depends on the viewing angle. It is important to bear in mind, however, that the original images are reconstructed accurately in all cases. The errors in the three-dimensional structures shown in Chapter

[4](#) are introduced by the implicit prior probability distribution inherent in the use of the minimum-norm solution. As shown in Section [3.2.3](#), the error must be expressible as a linear combination of null space vectors. It makes sense, therefore, to search the null space to satisfy a given prior probability distribution rather than searching the space of all possible distributions. This can dramatically reduce the dimensionality of the search space, and it guarantees the accuracy of the reconstruction of the image.

Demonstrating the efficacy of the method of optimizing a prior probability distribution over the null space is a difficult problem, in that it requires a specific prior probability distribution. If the prior is formulated poorly, the results will be poor regardless of the validity of the method. The results here use an ideal prior probability distribution, corresponding to minimizing the sum of squared differences with respect to the ground truth function. This is, obviously, not the form that a prior probability distribution would take in practice, but it allows the testing of the method without worrying that the prior was badly defined. The results presented are intended to show the performance of the method given a high-quality prior probability distribution; obtaining such a prior is a domain-specific problem, better suited to purely astronomical research than to this work.

Table [A.3](#) shows reconstructions after the ideal linear combination of null space vectors has been added. The difference between the reconstructions after the optimizing the ideal prior and the simple least-norm solutions shown in Table [4.7](#) is dramatic, especially for the face-on viewing angle. Only a few examples are shown, since the reconstructions are entirely accurate in every case; there is nothing to be gained by showing ten identical results.
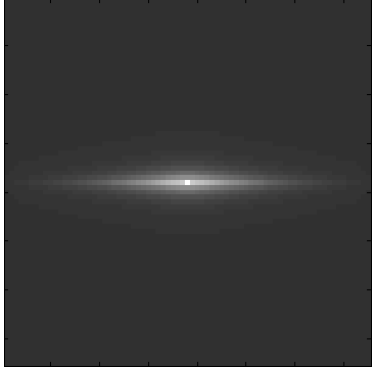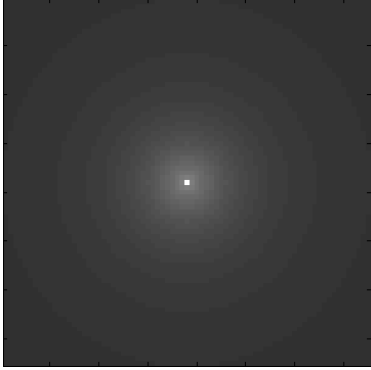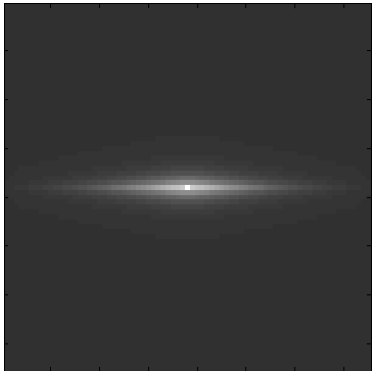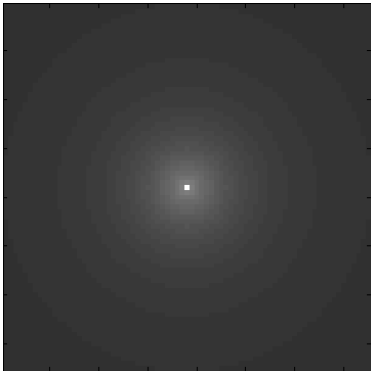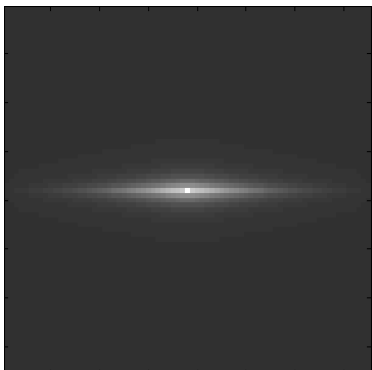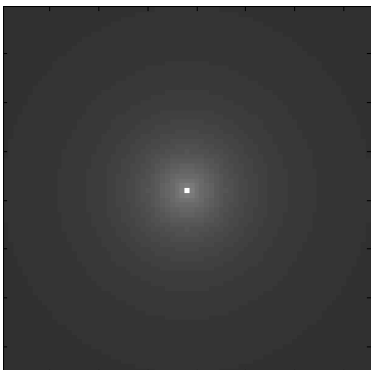
θ	Edge-On View	Face-On View
0°		
30°		
60°		

Table A.3: Reconstructions with residual accounted for by a linear combination of null space vectors

Appendix B

Algorithms

The mathematical framework defined in this thesis leads to a variety of algorithms. Since the primary concerns of the thesis are the framework and its properties, detailed specifications of the algorithms were not included in the main body of the thesis. Nonetheless, these algorithms are of interest and are important for replicating the results shown. This appendix provides formal definitions for the algorithms used to generate the results shown in Chapter 4.

In the interest of brevity, I define a number of simple utility functions. The details of the implementations of these functions are not important to the algorithms in which they are used, and they are all simple to implement. These functions are:

- $\mathbf{i} = \text{PackImg}(I)$: "Packs" a two-dimensional image as a vector.

- $I = \text{UnpackImg}(\mathbf{i})$: "Unpacks" a two-dimensional image from its vector representation.

- $\mathbf{l} = \text{PackVol}(L)$: Packs a three-dimensional luminosity density distribution as a vector.

- $L = \text{UnpackVol}(\mathbf{l})$: Unpacks a three-dimensional luminosity density distribution from its vector representation.

- $k = \text{IndImg}(x, y)$: Converts a pixel location in the image to the position of the corresponding pixel's value in the vector produced by PackImg.

- $k = \text{IndSph}(r)$: Converts a spherical radius to a column index in the projection matrix.

- $k = \text{IndCyl}(r, h)$: Converts a cylindrical radius and height to a column index in the projection matrix.

- $k = \text{IndVox}(r, h)$: Converts the coordinates (x, y, z) of a voxel to its position in the vector representation

These functions define the methods for converting between vector, image, and three-dimensional representations. So long as some such function is defined, the algorithms are equally valid regardless of the details of the conversion process.

B.1 Generating Projection Matrices

Generating the projection matrices is a vital step in the reconstruction process. The properties of these matrices were defined in Chapter 3, but the implementation details for specific cases were not shown. Section B.1.1 shows how a projection matrix for a spherically symmetric function is generated; Section B.1.2 shows the same process for a cylindrically symmetric function which is also symmetric about the equatorial plane. Section B.1.3 shows a less efficient method for generating a projection matrix for an arbitrary basis. For simplicity, both algorithms assume a square $n \times n$ input image, and assume that n is odd. It is simple to implement less restricted versions of these algorithms, but these would be less concise.

B.1.1 Spherical Symmetry

Input	Odd image size n
Output	Projection matrix G_P
As a function	$G_P = \text{ProjSph}(n)$

$s \leftarrow \lfloor \frac{n}{2} \rfloor$

$x \leftarrow -s$

while $x \leq s$ **do**

    $y \leftarrow -s$

    while $y \leq s$ **do**

        $z \leftarrow -s$

        while $z \leq -s$ **do**

            $r \leftarrow \sqrt{x^2 + y^2 + z^2}$

            $r_{int} \leftarrow \lfloor r \rfloor$

$$r_{frac} \leftarrow r - r_{int}$$
$$G_P(\text{IndImg}(x, y), \text{IndSph}(r_{int})) \leftarrow 1 - r_{frac}$$
$$G_P(\text{IndImg}(x, y), \text{IndSph}(r_{int} + 1)) \leftarrow r_{frac}$$
$$z \leftarrow z + 1$$
end while
$$y \leftarrow y + 1$$
end while
$$x \leftarrow x + 1$$
end while
return G_P

B.1.2   Cylindrical and Reflective Symmetry

Input   Odd image size n, viewing angle θ
Output   Projection matrix G_P
As a function $G_P = \text{ProjCylRefl}(n, \theta)$
$$s \leftarrow \lfloor \tfrac{n}{2} \rfloor$$

$$R \leftarrow \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos -\theta & -\sin -\theta \\ 0 & \sin -\theta & \cos -\theta \end{bmatrix}$$

$$x \leftarrow -s$$
while $x \leq s$ **do**
$$\quad y \leftarrow -s$$
$\quad$**while** $y \leq s$ **do**
$$\quad\quad z \leftarrow -s$$
$\quad\quad$**while** $z \leq -s$ **do**

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \leftarrow R \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$r \leftarrow \sqrt{x'^2 + y'^2}$$
$$h \leftarrow |z'|$$
$$r_{int} \leftarrow \lfloor r \rfloor$$
$$r_{frac} \leftarrow r - r_{int}$$
$$h_{int} \leftarrow \lfloor h \rfloor$$
$$h_{frac} \leftarrow h - h_{int}$$

$$G_P(\text{IndImg}(x,y), \text{IndCyl}(r_{int}, h_{int})) \leftarrow (1 - r_{frac})(1 - h_{frac})$$
$$G_P(\text{IndImg}(x,y), \text{IndCyl}(r_{int} + 1, h_{int})) \leftarrow r_{frac}(1 - h_{frac})$$
$$G_P(\text{IndImg}(x,y), \text{IndCyl}(r_{int}, h_{int} + 1)) \leftarrow (1 - r_{frac})h_{frac}$$
$$G_P(\text{IndImg}(x,y), \text{IndCyl}(r_{int} + 1, h_{int} + 1)) \leftarrow r_{frac}h_{frac}$$

$z \leftarrow z + 1$
    end while
    $y \leftarrow y + 1$
  end while
  $x \leftarrow x + 1$
end while
return G_P

B.1.3   Generic Basis

Input   Odd image size n, rotation matrix R, matrix β_{model} which defines each basis vector as a linear combination of voxel values (see Section 3.2.5).
Output   Projection matrix G_P
As a function $G_P = \text{ProjGeneric}(n, R, \beta_{model})$

$s \leftarrow \lfloor \frac{n}{2} \rfloor$
$x = -s$
while $x \leq s$ **do**
    $y = -s$
    while $y \leq s$ **do**
        $z = -s$
        while $z \leq s$ **do**

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \leftarrow R \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$x_{int} \leftarrow \lfloor x \rfloor$
$x_{frac} \leftarrow x - x_{int}$
$y_{int} \leftarrow \lfloor y \rfloor$
$y_{frac} \leftarrow y - y_{int}$
$z_{int} \leftarrow \lfloor z \rfloor$
$z_{frac} \leftarrow z - z_{int}$
$G_P(\text{IndImg}(x,y), \text{IndVox}(x_{int}, y_{int}, z_{int})) \leftarrow (1 - x_{frac})(1 - y_{frac})(1 - z_{frac})$

$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int}, y_{int}, z_{int} + 1)) \leftarrow (1 - x_{frac})(1 - y_{frac})z_{frac}$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int}, y_{int} + 1, z_{int})) \leftarrow (1 - x_{frac})y_{frac}(1 - z_{frac})$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int}, y_{int} + 1, z_{int} + 1)) \leftarrow (1 - x_{frac})y_{frac}z_{frac}$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int} + 1, y_{int}, z_{int})) \leftarrow x_{frac}(1 - y_{frac})(1 - z_{frac})$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int} + 1, y_{int}, z_{int} + 1)) \leftarrow x_{frac}(1 - y_{frac})z_{frac}$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int} + 1, y_{int} + 1, z_{int})) \leftarrow x_{frac}y_{frac}(1 - z_{frac})$$
$$G_P(\text{IndImg}(x, y), \text{IndVox}(x_{int} + 1, y_{int} + 1, z_{int} + 1)) \leftarrow x_{frac}y_{frac}z_{frac}$$

        end while

      end while

    end while

   $G_P \leftarrow G_P \beta_{model}$

   return G_P

B.1.4   Background Luminosity

In some cases, a background value contributes to each pixel in the image. In astronomy, this is called the *sky level*. If this is not accounted for, it will interfere with the reconstruction process. If the sky level is known, it can simply be subtracted from the input image during preprocessing; it will not affect the reconstruction. If it is unknown, it can be incorporated into the reconstruction problem as another element in the luminosity density distribution which contributes equally to every pixel. This is the approach I have taken for reconstruction with realistic images of galaxies.

B.2   Regularization

Regularization methods were described in Section 3.2.4 in terms of changes to the objective function. Here I describe methods for altering the projection matrix such that multiplication by the pseudoinverse of the new matrix results in the solution to the regularized version of the problem.

B.2.1 L_2-Norm

Input	Projection matrix G_P, regularization strength λ, number of columns c
Output	Regularized matrix G_P
As a function	$G_P = \text{RegL}_2(G_P, \lambda, c)$

$j \leftarrow 1$
while $j \leq c$ **do**
    $G_P(k, j) \leftarrow \lambda$
    $\mathbf{i}(k) \leftarrow 0$
    $k \leftarrow k + 1$
end while
return G_P

B.2.2   Equatorial Concentration

Input	Projection matrix G_P, regularization strength λ, odd image size n
Output	Regularized matrix G_P
As a function	$G_P = \text{RegEq}(G_P, \lambda, n)$

$s \leftarrow \lfloor \frac{n}{2} \rfloor$
$k = \text{rows}(G_P) + 1$
$r \leftarrow -s$
while $r \leq s$ **do**
    $h \leftarrow 0$
    while $h \leq s$ **do**
        $G_P(k, \text{IndCyl}(r, h) \leftarrow 2\lambda\pi rh$
        $\mathbf{i}(k) \leftarrow 0$
        $k \leftarrow k + 1$
    end while
end while
return G_P

B.3   Estimating Viewing Angle

Estimation of the viewing angle is a somewhat complex process. Here I present an algorithm using the heuristic defined in Section 3.3.4. For practical purposes I define two additional utility functions: MajorAxis(I) and MinorAxis(I). These functions return the lengths of the major and minor axes of the projected ellipse of the galaxy shown in image I; there are a variety of ways to implement these functions (especially in terms of how any extraneous objects are removed), so they are not explicitly defined in terms of implementation.

Input	Square input image I, (odd) image size n
Output	Estimated viewing angle θ
As a function	$\theta = \text{EstViewingAngle}(I, n)$

$\theta_{init} = \frac{\pi}{2} - \sin^{-1}\left(\frac{\text{MajorAxis}(I)}{\text{MinorAxis}(I)}\right)$

$\theta = \theta_{init} + 2.5 g_{inc}$

while $g_{inc} > g_{final}$ **do**

    $\theta_{min} \leftarrow \theta - 5g_{inc}$

    $\theta_{max} \leftarrow \theta + 5g_{inc}$

    $k \leftarrow -5$

    $E_b est \leftarrow \infty$

    while $k \leq 5$ **do**

        $\theta_k \leftarrow \theta + k g_{inc}$

        $G_P \leftarrow \text{ProjCyl}(n, \theta_k)$

        $\mathbf{i} \leftarrow \text{Pack}(I)$

        $\hat{\mathbf{l}} \leftarrow G_P^+ \mathbf{l}$

        $\hat{\mathbf{i}} \leftarrow G_P \mathbf{l}$

        $E \leftarrow \|\mathbf{i} - \hat{\mathbf{i}}\|^2$

        if then $E < E_{best}$

            $E_{best} = E$

            $k_{best} = k$

        end if

        $k \leftarrow k + 1$

    end while

    $\theta \leftarrow \theta + k_{best} * g_{inc}$

    $g_{inc} \leftarrow \frac{g_{inc}}{5}$

end while

return θ

B.4   Reconstruction

The following algorithms describe the reconstruction problem in general, from the output image and any other parameters (such as regularization strength or viewing angle) to the reconstructed luminosity density distribution and image. Many variations on these algorithms are possible, especially in the choice of regularization.

B.4.1   Spherical Symmetry

Input   Square input image I, (odd) image size n
Output   Reconstructed luminosity density distribution $\hat{L}$, reconstructed
                        image $\hat{I}$
As a function $[\hat{L}, \hat{I}] = \mathrm{ReconSph}(I, n)$

  $G_P \leftarrow \mathrm{ProjSph}(n)$
  $\mathbf{i} \leftarrow \mathrm{PackImg}(I)$
  $\hat{\mathbf{l}} \leftarrow G_P^+ \mathbf{i}$
  $\hat{\mathbf{i}} \leftarrow G_P \mathbf{l}$
  $\hat{I} \leftarrow \mathrm{UnpackImg}(\hat{\mathbf{i}})$
  $\hat{L} \leftarrow \mathrm{UnpackVol}(\hat{\mathbf{l}})$
  return $\hat{L}, \hat{I}$

B.4.2   Cylindrical and Reflective Symmetry

Input   Square input image I, (odd) image size n, viewing angle θ,
                        regularization strength λ
Output   Reconstructed luminosity density distribution $\hat{L}$, reconstructed
                        image $\hat{I}$
As a function $[\hat{L}, \hat{I}] = \mathrm{ReconSph}(I, n, \theta, \lambda)$

  $G_P \leftarrow \mathrm{ProjSph}(n, \theta)$
  $G_P \leftarrow \mathrm{RegEq}(G_P, \lambda)$
  $\mathbf{i} \leftarrow \mathrm{PackImg}(I)$
  $\hat{\mathbf{l}} \leftarrow G_P^+ \mathbf{i}$
  $\hat{\mathbf{i}} \leftarrow G_P \mathbf{l}$
  $\hat{I} \leftarrow \mathrm{UnpackImg}(\hat{\mathbf{i}})$
  $\hat{L} \leftarrow \mathrm{UnpackVol}(\hat{\mathbf{l}})$
  return $\hat{L}, \hat{I}$

References

[1] C. P. Ahn, R. Alexandroff, C. Allende Prieto, S. F. Anderson, T. Anderton, B. H. Andrews, É. Aubourg, S. Bailey, E. Balbinot, R. Barnes, and et al. The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey. *The Astrophysical Journal Supplement*, 203:21, December 2012.

[2] R. C. Aster, C. H. Thurber, and B. Borchers. *Parameter Estimation and Inverse Problems*, volume 90 of *International Geophysics Series*. Elsevier Academic Press, Amsterdam, 2005.

[3] W. E. Baggett, S. M. Baggett, and K. S. J. Anderson. Bulge-disk decomposition of 659 spiral and lenticular galaxy brightness profiles. *The Astronomical Journal*, 116:1626–1642, 1998.

[4] A. H. Barr. Superquadrics and angle-preserving transformations. *IEEE Computer Graphics and Applications*, 1(1):11–23, 1981.

[5] R. N. Bracewell. Strip integration in radio astronomy. *Australian Journal of Physics*, 9(2):198–217, 1956.

[6] R. N. Bracewell. Image reconstruction in radio astronomy. *Topics in Applied Physics*, 32:81–104, 1979.

[7] R. N. Bracewell. *Fourier Analysis and Imaging*. Kluwer Academic/Plenum Publishers, 1st edition, 2003.

[8] N. Brandt. http://www2.astro.psu.edu/users/niel/astro1/slideshows/class21/slides-21.html.

[9] Y. I. Byun and K. C. Freeman. Two-dimensional decomposition of bulge and disk. *Astrophysical Journal*, 488:563–574, 1995.

[10] R. E. de Souza, D. A. Gadotti, and S. dos Anjos. BUDDA: A new two-dimensional bulge/disk decomposition code for detailed structural analysis of galaxies. *The Astrophysical Journal Supplement Series*, 153:411–427, 2004.

[11] W. T. Freeman. The generic viewpoint assumption in a framework for visual perception. *Nature*, 368:542–545, 1994.

[12] J. Gu, S. K. Nayar, E. Grinspun, P. N. Belhumeur, and R. Ramamoorthi. Compressive structured light for recovering inhomogeneous participating media. In *European Conference on Computer Vision*, 2008.

[13] Tim Hawkins, Per Einarsson, and Paul Debevec. Acquisition of time-varying participating media. *ACM Trans. Graph.*, 24(3):812–815, July 2005.

[14] H. Hoffman, W. E. Torres, and R. D. Ernst. Paleoradiology: advanced CT in the evaluation of nine Egyptian mummies. *RadioGraphics*, 22(2):377–385, 2002.

[15] H. Mo, F. van den Bosch, and S. White. *Galaxy Formation and Evolution*. Cambridge University Press, 1st edition, 2010.

[16] R. Molina, J. Nunez, F. J. Cortijo, and J. Mateos. Image restoration in astronomy: a Bayesian perspective. *Signal Processing Magazine, IEEE*, 18(2):11–29, 2001.

[17] C. Y. Peng, L. C. Ho, C. D. Impey, and R. Hans-Walter. Detailed decomposition of galaxy images. II. beyond axisymmetric models. *The Astronomical Journal*, 124:266–293, 2002.

[18] C. Y. Peng, L. C. Ho, C. D. Impey, and R. Hans-Walter. Detailed decomposition of galaxy images. *The Astronomical Journal*, 139(6):2097–2129, 2010.

[19] M. Prasad, A. Zisserman, and A. Fitzgibbon. Single view reconstruction of curved surfaces. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2, pages 1345–1354, 2006.

[20] H. Saito and N. Tsunashima. Superquadrics parameter estimation from shading image using genetic algorithm. In *20th International Conference on Industrial Electronics, Control, and Instrumentation*, volume 2, pages 978–983, 1994.

[21] G. Strang. *Linear Algebra and its Applications*. Thompson Brooks/Cole, 4th edition, 2006.

[22] R. Szeliski. *Computer Vision: Algorithms and Applications*. Springer, 2010.

[23] E. Tempel, A. Tamm, R. Kipper, and P. Tenjes. Uncertainties in SDSS galaxy parameter determination: 3D photometrical modelling of test galaxies and restoration of their structural parameters, 2012.

[24] S. van den Bergh. *Galaxy Morphology and Classification.* Cambridge University Press, 1st edition, 1998.

[25] S. Zaroubi, G. Squires, Y. Hoffman, and J. Silk. Deprojection of rich cluster images. *The Astrophysical Journal Letters*, 500(2), 1998.

[26] C. Zhang and B. Yuan. Representation and reconstruction of three-dimensional objects using nonlinear deformable superquadric models. In *IAPR Workshop on Machine Vision Applications*, pages 171–174, 1992.