OBSERVATIONS AND MODELING OF THE GAS DYNAMICS OF THE BARRED SPIRAL GALAXY NGC 3359

BY

JOHN ROGER BALL

A DISSERTATION PRESENTED TO THE GRADUATE COUNCIL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1984

for Mary

ACKNOWLEDGMENTS

I must first offer my sincere thanks to my advisors, Drs. James Hunter and Steve Gottesman. They have been deeply involved with the progress of this dissertation, from its conception to the dotting of the last i and crossing of the last t. They have been equally involved in my development as an astronomer, and from each of them I have learned much more than I could detail here. Their guidance and their support are very warmly appreciated.

I would also like to express my gratitude to Dr. Jim Huntley of Bell Laboratories, who has been a member of my supervisory committee in all but name. Besides making available his hydrodynamical code, Jim has provided many insights into its results, and into the dynamics of barred spiral galaxies. Above all, he has lent a sympathetic ear to many a long-distance call of bewilderment, for which I thank him profusely.

I thank the remaining members of my committee, Drs. Haywood Smith and Gary Ihas, for their interest and comments on this work. Additionally, I thank Dr. T. D. Carr for his reading of Chapter II, and Dr. R. B. Piercey for reviewing the manuscript and attending my defense on short notice.

I acknowledge with pleasure the financial support, during my writing of this dissertation, of the National Science Foundation, under grant AST 81 16312, and of the Graduate School of the University of Florida.

I am greatly indebted to Dr. Debra Elmegreen, of IBM Watson Research Center, for the use of the near infrared data which forms the basis of Chapter IV.

The original observations described in this dissertation were obtained with the Very Large Array of the National Radio Astronomy Observatory. The National Radio Astronomy Observatory is operated by Associated Universities, Inc., under contract with the National Science Foundation. I can hardly exaggerate the helpfulness of the staff, as well as many of the visitors, of the NRAO, who patiently bore the thousand questions of the neophyte observer struggling to achieve competence. Although there are many who deserve mention, I will single out only the inimitable Drs. Jacqueline van Gorkom and Rick Perley, to whom go my most heartfelt thanks.

For expert assistance with the preparation of various figures, I am grateful to Paul Gombola and Hans Schrader.

The physical quality of the volume which the reader has in his hands is largely the product of the labors of Irma Smith, who has truly gone the extra mile in helping to prepare this dissertation. Without her masterful touch, putting together a satisfactory copy would have been next to impossible. Therefore, I thank Irma for helping to preserve my sanity, and hope that she has not done so at the cost of her own.

There are also several people whose support demands a word of thanks, although their contributions to this dissertation are less direct. It is not only a pleasure, but an honor, to mention two gentlemen whose tremendous abilities and dedication to the art of teaching have a great deal to do with my decision to attempt a career in the physical sciences: Mr. Howard Slaten, of Durant, Oklahoma, High School, and Dr. Frank Hart, of the University of the South.

No one understands the travails of graduate school like another graduate student, and what limited capacity I still have for normal

behavior owes a good deal to the understanding and friendship of the astronomy graduate students who have been here at the University of Florida during my stay, from 1977 to 1984. I also, needless to say, learned a lot of astronomy from, and with, them. I thank all of them, for innumerable reasons, and, especially, I thank Chris St. Cyr, Joe Pollock, Greg Fitzgibbons, and Glenn Schneider.

The encouragement of my parents, Lee and Kate Ball, through the past twenty-eight years is a gift which I value more than I can hope to express.

My deepest thanks of all go to my wife, Sandy, who has been with me constantly, however many miles have been between us. Although the hardships incurred by my pursuit of this degree have fallen on her as well as on me, she has given me only encouragement, comfort, inspiration, and hope. My gratitude to her for sharing the burdens of these years goes far beyond what can be expressed here; let me simply say that I thank her for all that she has done, and for all that she is: she is, above all, the point.

TABLE OF CONTENTS

Page

ACKNOWLEDGMENTS iii

LIST OF TABLES. viii

LIST OF FIGURES ix

ABSTRACT. xi

Chapter

I. INTRODUCTION. 1

An Overview of Disk Galaxy Dynamics. 1
Dynamics of Barred Spirals 7
The Barred Spiral Galaxy NGC 3359. 11

II. RADIO OBSERVATIONS. 17

Neutral Hydrogen as a Kinematical Tracer 17
Elementary Aperture Synthesis Theory 21
VLA Observations of NGC 3359 39

III. DISTRIBUTION AND KINEMATICS OF NEUTRAL HYDROGEN 92

Techniques of Spectrum Integration 92
Neutral Hydrogen Distribution in NGC 3359. 100
Kinematics of the Neutral Hydrogen 120
The Companion Galaxy to NGC 3359 136

IV. THE STELLAR BAR OF NGC 3359 148

Surface Photometry: Introduction. 148
Projection Effects in Triaxial Figures 155
Brightness Distributions 167
The Attraction of a Triaxial Homoeoid. 182

V. HYDRODYNAMICAL MODELS 195

The Beam Scheme. 195
Models Using Triaxial Bars 207
Models Using Oval Distortions. 222

VI. SUMMARY OF CONCLUSIONS. 258

Observational Results. 258
Hydrodynamical Models. 261

Page

APPENDIX A. 265

APPENDIX B. 269

REFERENCES. 275

BIOGRAPHICAL SKETCH 283

LIST OF TABLES

Table		Page
1-1	GLOBAL PROPERTIES OF NGC 3359.	14
2-1	CALIBRATION PROPERTIES	46
2-2	CONTINUUM SOURCES IN THE NGC 3359 FIELD.	58
2-3	SINGLE CHANNEL SIGNAL AND NOISE CHARACTERISTICS. . . .	90
3-1	LOGARITHMIC SPIRAL FIT TO FIGURE 3-3	109
3-2	SUMMARY OF NEUTRAL HYDROGEN OBSERVATIONS	138
4-1	BAR PROJECTION PARAMETERS.	168
4-2	EVALUATION OF INTEGRALS FROM GRADSHTEYN AND RYZHIK (1980).	188
5-1	BEAM PROPERTIES, ONE-DIMENSIONAL CASE.	198

LIST OF FIGURES

Figure		Page
1-1	The barred spiral galaxy NGC 3359, from a print of the National Geographic Society--Palomar Sky Survey blue plate.	13
2-1	Geometry of the two-element interferometer	22
2-2	The (u,v) coordinate system.	30
2-3	The tracks swept out in the (u,v) plane by a single interferometer pair	35
2-4	Single-channel clean maps of the neutral hydrogen in NGC 3359.	62
2-5	Single-channel dirty map of the H I in NGC 3359, at the center velocity of 1016 km/s.	83
2-6	Single-channel clean maps of neutral hydrogen in NGC 3359, showing a larger area than in Figure 2-4. . . .	86
3-1	Contours of neutral hydrogen surface density in NGC 3359.	102
3-2	The same integrated H I map as in Figure 3-1, shown here in grey tone format.	104
3-3	Distributions of H I and H II in NGC 3359.	107
3-4	Mean deprojected H I surface density in NGC 3359, as a function of radius.	114
3-5	Continuum emission in NGC 3359	118
3-6	Contours of heliocentric, line-of-sight velocity in NGC 3359.	122
3-7	Angle-averaged rotation curve of NGC 3359.	127
3-8	Rotation curve determined independently from velocities along the two halves of the major axis.	128

Figure		Page
3-9	Rotational velocities of the first-order mass models used in this chapter to estimate the mass of NGC 3359.	134
3-10	The observed global spectrum of neutral hydrogen in NGC 3359.	137
3-11	Contours of integrated H I column density in the area surrounding the isolated companion to NGC 3359. . . .	140
3-12	Relative locations of NGC 3359 and its satellite	142
3-13	The observed global spectrum of H I in the companion galaxy to NGC 3359.	144
4-1	The I passband image of NGC 3359	153
4-2	Geometry of a triaxial bar in a disk galaxy.	157
4-3	Contour plot representation of Elmegreen's near infrared plate of NGC 3359.	162
4-4	Near infrared brightness profiles, perpendicular to the major axis of the disk of NGC 3359	171
4-5	Brightness profiles along the bar of NGC 3359.	178
5-1	Gas response in the first of the purely triaxial models of NGC 3359.	212
5-2	Gas response in the second of the purely triaxial models.	216
5-3	Gas response in the pure oval distortion model	228
5-4	Gas response in the first of the hybrid oval distortion and triaxial models	235
5-5	Gas response in the second of the hybrid oval distortion and triaxial models	239
5-6	Gas response in the third of the hybrid oval distortion and triaxial models	243
5-7	Gas response in the final hybrid oval distortion and triaxial model.	247
5-8	The rotation curve derived from the model illustrated in Figure 5-7	254

Abstract of Dissertation Presented to the Graduate Council
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

OBSERVATIONS AND MODELING OF THE GAS DYNAMICS OF THE BARRED SPIRAL GALAXY NGC 3359

By

John Roger Ball

August 1984

Chairman: James H. Hunter, Jr.
Major Department: Astronomy

We have conducted a detailed study of the dynamics of the neutral hydrogen gas in the bright northern barred spiral galaxy NGC 3359. Observations of the 21 cm line at the Very Large Array have been reduced to give single-channel maps with a spatial resolution of 18", and a velocity resolution of 25 km/s. The acquisition, calibration, and reduction of the data are discussed in some detail.

Maps of the integrated column density and mean velocity of the atomic hydrogen, derived from the channel maps, provide the principal data for an investigation of the dynamics associated with the spiral structure of the galaxy. On scales comparable to the resolution of this survey, approximately 1 kpc at the distance of NGC 3359, the gas is broken up into a somewhat chaotic distribution of local maxima and minima. However, on larger scales the column density shows a smooth, grand design spiral pattern with two principal spiral arms. The extent and density of these two arms are roughly equal in the 21 cm map, unlike the

optical image. These neutral hydrogen arms are very well correlated with the positions of H II regions.

The observed velocity field of NGC 3359 shows large kinematical deviations from simple rotation. The locations of the disturbances are closely correlated with the optical and 21 cm spiral structure, and indicate that significant spiral streaming motions are present in the arms.

We have performed numerical modeling of the galaxy, using the well-known "beam scheme" hydrodynamical code, to attempt to reproduce the observed kinematics. Two types of model have been employed. In the first, the spiral structure is driven by an inhomogeneous, triaxial bar, whose density distribution is determined from near infrared surface photometry. The alternative is an oval distortion of the underlying disk surface density. In models which depend solely on the triaxial bar, the kinematical perturbations in the galaxy's outer regions are far too weak, relative to those near the bar. The oval distortion models fit the observed velocities much more accurately. We conclude that the spiral structure in NGC 3359 cannot be driven entirely by the nonaxisymmetric force produced by the observed optical bar.

CHAPTER I
INTRODUCTION

An Overview of Disk Galaxy Dynamics

This dissertation is concerned with the application of certain hydrodynamical models to the spiral structure in the barred spiral galaxy NGC 3359. In this introductory chapter, we hope to explain the importance of our chosen subject, and to place it in its astrophysical context. Accordingly, we first will review, briefly, the topic of spiral structure in disk galaxies, pointing out the chief obstacles that arise in an analytical approach to the general problem. We will see that restricting our attention to the barred spiral galaxies eliminates one of the central, physical difficulties of the analytical theory, and, at the same time, dictates the use of numerical techniques. We then introduce the reader to the particular barred spiral, NGC 3359, which will be the subject of our investigation into the details of barred spiral dynamics. We close the chapter with an outline of the remainder of the present work.

The problem of spiral structure in galaxies is one of the major challenges currently facing dynamical astronomy. It seems probable that spiral galaxies make up quite a large proportion of the luminous mass of the universe, especially outside of rich clusters of galaxies, though perhaps not as large a numerical fraction of galaxies. For example, in our Local Group of galaxies, van den Bergh (1981) finds only three true spirals out of 29 known members, but these are estimated to contribute some 90% of the total luminosity contained in all 29 galaxies. By comparison

with the other principal classes of galaxies, namely, the elliptical, lenticular, and irregular galaxies (Sandage, 1961), spirals pose intriguing problems for dynamical study. Their structures are, to all appearances, much more complex than those of elliptical and lenticular galaxies, and yet remain ordered and fairly symmetrical, unlike the irregulars. The dynamics governing the spirals should, therefore, be both interesting and ultimately tractable, if difficult. Furthermore, there are several techniques available for measuring radial velocities in these galaxies, so their internal motions can be explored observationally.

However, the dynamics of spiral galaxies are not yet well understood. While the light distribution of a typical spiral is dominated by a fairly flat, thin, rotating disk, with the addition, in some types, of an inner "spheroidal" or "bulge" component, the mass distribution is often inferred to be markedly different. It is common to postulate a "dark halo," more or less spherical in shape, of some kind of matter which has a high mass-to-luminosity ratio. The visible disk is pictured as lying in the equatorial plane of this halo, though often the halo is believed to extend to considerably greater radius than the disk. Halos have been discussed extensively in the recent literature, and work on them is reviewed by van den Bergh (1978) and by Haud and Einasto (1983). We will have relatively little to say about them in this dissertation.

In light of the uncertainties in our knowledge of the masses and other large-scale properties of spiral galaxies, perhaps it is not surprising that there is not yet general agreement on a specific mechanism for the most striking dynamical features of these objects,

the spiral patterns themselves. All the promising theories of spiral structure advanced to date depend on the existence of shear in the galaxies' disks, due to differential rotation. It is well established that spirals do rotate in that way. However, the problem of deriving a self-consistent treatment of the disk dynamics, which explains the spiral patterns, remains an open one. Perhaps the best-known approaches are the Lin-Shu density wave theory, and the scenario of stochastic, self-propagating star formation (SSPSF). These two hypotheses are radically different.

The density wave theory is an analytical treatment, based on a linearization of the relevant equations (see below). Some elements of this treatment are present in the work of Lindblad (e.g., Lindblad, 1963), but the modern form of the theory derives primarily from Lin and Shu (1964), Toomre (1964), and Kalnajs (1965). Subsequent papers of special importance include, among others, those by Lin and Shu (1966), Lin, Yuan, and Shu (1969), Toomre (1969), Shu (1970), Lynden-Bell and Kalnajs (1972), Mark (1974, 1976), and Lau and Bertin (1978). Lin and Lau (1979) give a thorough mathematical presentation and summary of density wave theory. As Lin and Shu (1967) emphasize, this theory is an attempt to explain the grand design of spiral features over the whole face of a disk galaxy. It consists of a demonstration that a small, quasi-stationary perturbation in an otherwise symmetric disk can, at least under some conditions, be self-sustaining. The spiral arms are seen as manifestations of a "density wave," which is, in essence, a tightly wound, rotating, spiral potential minimum, through which the material of the galaxy flows. That is, the matter which one sees in the spiral arms at any given moment is not permanently bound to the arms. In

its orginal form, the theory made no effort to explain the origin of the spiral perturbation, and to do so remains one of its great unsolved problems. It has also grown increasingly clear that even the maintenance of the waves over many rotations is not adequately dealt with by this theory as it now stands. Toomre (1977) gives an excellent review of the theory's progress and its difficulties. Despite the persistence of the latter, the density wave theory is probably the most significant advance, to date, in our understanding of the basic dynamics involved. Also, it seems to succeed quite well, for some galaxies, in matching the observed morphology (Roberts, Roberts, and Shu, 1975).

The stochastic, self-propagating star formation theory takes a very different tack. It explains the global pattern as an effect of local interactions in the gas. Also, it is a numerical scheme by its very nature. The idea first appeared in a paper of Muller and Arnett (1976), but the theory in its present form has been developed, primarily, by Gerola and Seiden (1978), Seiden, Schulman, and Gerola (1979), Seiden and Gerola (1979), and Seiden (1983). The thermodynamical interpretation which underlies the seemingly heuristic approach of SSPSF is brought out in a very interesting pair of papers by Shulman and Seiden (1978, 1982). As Gerola and Seiden (1979) point out, it is not a dynamical theory, but a statistical model, which works in the following way. One begins with a small number of high-mass stars, randomly distributed about a galactic disk, which experience supernova events. For each of these, there is a probability P_{st} that the shock wave engendered by the supernova triggers the formation of a similarly massive star in a nearby gas cloud. The physical justification for this step has been studied extensively, by, among others, Elmegreen and Lada (1977), Elmegreen and Elmegreen (1978),

Lada, Blitz, and Elmegreen (1978), and Welter (1982). As this process slowly percolates through the available gas of the galaxy, clumps of OB stars are produced. These groups are then pulled out into trailing, spiral features by differential rotation. Not surprisingly, the appearance of the model galaxy depends sensitively on its rotation curve (through its shear), but the general tendency is toward rather fragmented spiral structure with many disconnected, or imperfectly connected, branches. The Sc galaxy M 101 is an example of one whose overall morphology is strikingly similar to models of this kind.

Although the density wave and self-propagating star formation theories differ in method, results, and, most importantly, in their assessment of what is fundamental to the physics of the problem, they need not be mutually exclusive. As discussed by Gerola and Seiden (1979), there is no conceptual barrier to the simultaneous operation of both processes in a single galaxy (Jensen, Talbot, and Dufour, 1981). Perhaps a more common view is that these two analyses may correspond to the subjective, but useful, division of spiral galaxies into grand design and "flocculent" types (e.g., Elmegreen, 1981).

The techniques to be used in the present study differ in spirit from both of those just discussed. As in the density wave theory, we will seek an explanation of spiral structure as a global mode in the galaxy under investigation. However, our approach is strictly through numerical experimentation. This methodology is forced upon us by the desire to explore dynamics which are inaccessible to the density wave, and related, theories. The fundamental difficulty arises from the nonlinearity of the problem, which is contained fully in the following set of equations. We employ hydrodynamical equations here to represent

the behavior of the constituent matter of a galaxy, as we will be concerned primarily with the gaseous component in this work. It can be shown that somewhat similar expressions hold for the case of stellar dynamics in the gravitational field of a galaxy (Freeman, 1975; Lin and Lau, 1979). Let us denote the gas density by ρ, the velocity vector by $\underline{v}$, the pressure by P, and the kinetic energy density by E. Neglecting the effects of magnetic fields, we have the equation of mass continuity for a compressible fluid:

$$\frac{\partial}{\partial t}(\rho)+\nabla\cdot(\rho\underline{v}) = 0; \tag{1-1}$$

the equation of motion:

$$\frac{\partial}{\partial t}(\rho\underline{v})+\nabla(\rho\underline{v}^2+P)-\underline{F} = 0; \tag{1-2}$$

and the energy equation:

$$\frac{\partial}{\partial t}(E)+\nabla\cdot[\underline{v}(E+P)]-Q = 0. \tag{1-3}$$

Here Q represents explicitly dissipative terms, such as viscosity and heat conduction. Finally, the force, $\underline{F}$, consists only of gravitational and viscous terms. We will not usually write the latter explicitly in this study (though they may not be negligible), because the numerical calculation with which we will model these equations contains the

effects of viscosity implicitly. If we omit the viscous force, then $\underline{F}$ is solely gravitational, and so is specified by Poisson's equation:

$$\underline{F} = -\nabla\Phi, \tag{1-4}$$

$$\nabla^2\Phi = \pi G\rho. \tag{1-5}$$

Since equations (1-1), (1-2), and (1-3) are nonlinear, the problem of galaxian dynamics, contained in equations (1-1) through (1-5), is not susceptible to linear analysis, although one can sometimes develop a set of linear equations which approximates this system for disturbances of small amplitude. The density wave theory is an example of that approach, in which the approximation is valid as the perturbation density and force, which are involved in the spiral pattern, go to zero. Unfortunately, there is no reason to believe that these perturbations should be small, in general. Even if the difficulties currently confronting the density wave theory are overcome, this theory can only be regarded as the precursor to a less restrictive theory.

Dynamics of Barred Spirals

A large subclass of spirals is that of the barred spiral galaxies, so named because of elongated concentrations of light at their centers (e.g., Sandage, 1961). These bars are generally interpreted as nonaxisymmetric mass concentrations, with the form of the bar rotating as a solid body, though the individual stars which compose it may well have more complicated orbits. Such nonaxisymmetric configurations have been found to be long-lived and robust in a great many numerical simulations of gravitationally interacting particles (Miller, 1971, 1976, 1978;

Hohl, 1971, 1972, 1976, 1978; Ostriker and Peebles, 1973; Sellwood, 1980; and, especially, Miller and Smith, 1979). The attack by Kalnajs (1972) upon the very difficult analytical problem of the stability of rotating disk galaxies points toward similar conclusions. Initially, these bars were viewed as aberrant instabilities in rotating disks. However, they seem to form without fail when the disks are dynamically "cold," i.e., have small velocity dispersions. It is more to the point, therefore, to say that these experiments demonstrate the stability of the equilibrium figure in a flattened system whose internal motions are dominated by rotation: namely, a bar. Indeed, one of the reasons for invoking dark halos, in nonbarred or "normal" spirals, is that a halo a few times more massive than its embedded disk helps to suppress bar development, as was first surmised by Ostriker and Peebles (1973). The ubiquity of bars in the N-body results immediately suggests a ready source of a strong, periodic, nonaxisymmetric driving perturbation on the background disk potential in a spiral galaxy. If the bar is taken to be the cause of the nonaxisymmetric perturbation in the rest of the galaxy, one major aspect of the problem of the origin and maintenance of spiral structure is resolved at once. One the other hand, the nonlinearity of the problem can no longer be evaded, as the "perturbation" force is far from negligible.

The theoretical study of barred spirals, then, is best carried out by numerical techniques. In recent years, many numerical experiments of this kind have been published. The first successful models were those of Sanders and Huntley (1976) and of Sorensen, Matsuda, and Fujimoto (1976). The former used the "beam scheme," a hydrodynamical code originally developed by Sanders and Prendergast (1974), to follow the gas response.

(The beam scheme is discussed in some detail in Chapter V of this dissertation.) Shortly thereafter, Huntley (1977), Huntley, Sanders, and Roberts (1978), and Sanders and Tubbs (1980) explored fairly thoroughly the barred spirals which can be generated by varying the dynamical parameters in the beam scheme. Huntley (1980) investigated the inclusion of the self-gravity of the gas. The beam scheme also served as the basis for the barred spiral models of Schempp (1982). In the meantime, other hydrodynamical codes were used for this purpose by Sanders (1977), by Berman, Pollard, and Hockney (1979), by Roberts, Huntley, and van Albada (1979), by Matsuda and Isaka (1980), and by van Albada and Roberts (1981). These codes give slightly different results from one another, especially in calculating the strengths and locations of shocks. However, these are differences of detail, arising largely from differences in the effective spatial resolution from one scheme to another. Both the essential physics, and the observable consequences, of the models are quite similar, for similar input potentials. Therefore, while it should be borne in mind that small-scale structure in the models depends somewhat on the particular code employed, this choice is less important in examining the global response of the gas. It is unlikely that the differences between the various computer programs will be very significant, for most galaxies, at the relatively coarse resolutions which typify radio observations, such as we report in this dissertation.

It is now clear that any of several hydrodynamical computer codes is capable of producing stable, trailing spiral density patterns in response to a rotating, barred potential. The goal of the present work is to examine the validity of these codes, by comparing in detail the predicted gas distribution and kinematics with observations of a real

galaxy. The code we use is a version of the beam scheme developed by Dr. J. M. Huntley, who has kindly made it available for this study.

The limitations of this work are quite fundamental, and we wish to alert the reader to them at once. Neither the present study, nor any other yet undertaken, provides a fully self-consistent treatment of the dynamical behavior of a complete galaxy. The discussion above should make it clear why that is a goal beyond our abilities at this stage of our understanding of galaxies. We will simply assume that an underlying, stellar disk has been perturbed by the formation of a bar into some stable configuration, and that the potential of this system remains constant, except for its rotation, during the time followed by our numerical computation. This is further assumed to be the only potential present in the model galaxy. That is, we ignore the self-attraction of the gas which is modeled. The justification for this assumption is that the density of gas, over most of a typical galaxian disk, is small compared to that in the disk-bar system. We also assume a gaseous component which is in a single, continuous, isothermal phase. Further, it is characterized by a viscous length scale equal to our computational cell size, and has no sources and only one sink, namely, outflow beyond our numerical grid. The intricate interactions of the interstellar medium with the galactic ecology are thus omitted entirely for simplicity's sake. We will elaborate on these questions in Chapter V. At a deeper level, the developing spiral pattern is not allowed to influence the imposed potentials of the disk and bar, so that there is no dynamical feedback in the models. (One consequence of this approach is that, if a model accurately reproduces the desired rotation curve, no further information is necessary about the axisymmetric component of the mass distribution

responsible for it. Thus, the presence or absence of a halo need not be considered explicitly.) In these ways does our approach fail to be strictly self-consistent, and to include all of the nonlinear behavior of the problem. Obviously, our investigation, even if successful in its stated aims, is simply one additional step toward a full understanding of spiral galaxy dynamics.

The Barred Spiral Galaxy NGC 3359

The galaxy whose dynamics we will attempt to model is NGC 3359, a system classified as S(B)cII by van den Bergh (1960), as SBc(rs) by Sandage (1961), as SBT5 by de Vaucouleurs, de Vaucouleurs, and Corwin (1976), and as SBc(s)I.8pec by Sandage and Tammann (1981). Figure 1-1 is an enlargement from the blue plate of the Palomar Observatory Sky Survey, showing the galaxy. The reader may also find the following published photographs of NGC 3359 of interest: a blue exposure with the Palomar 5-meter in Sandage (1961); yellow and hydrogen-alpha images from the Palomar Schmidt telescope, in Hodge (1969); and the near-infrared plate of Elmegreen (1981), also using the Palomar Schmidt, which will be used in Chapter IV of this study. Table 1-1 lists properties of NGC 3359, compiled from a variety of sources, which we will not attempt to verify independently in the present work.

The galaxy is dominated by a strong central bar, of length approximately 1.7 arc minutes, from which a fairly prominent, two-armed spiral pattern emerges. The spiral pattern is, however, rather asymmetric, with the arm that begins at the northern end of the bar being significantly more prominent than its counterpart. The asymmetry is equally clear in the map of NGC 3359 in H α by Hodge (1969), where a string of H II regions delineates the stronger arm much more sharply than the other.

Figure 1-1. The barred spiral galaxy NGC 3359, from a print of the National Geographic Society--Palomar Sky Survey blue plate. North is at the top, and east to the left. Several foreground stars are marked with crosses; the positions of these stars are used as fiducial points in figures throughout the present work.© 1960, National Geographic Society--Palomar Sky Survey. Reproduced by permission of the California Institute of Technology.

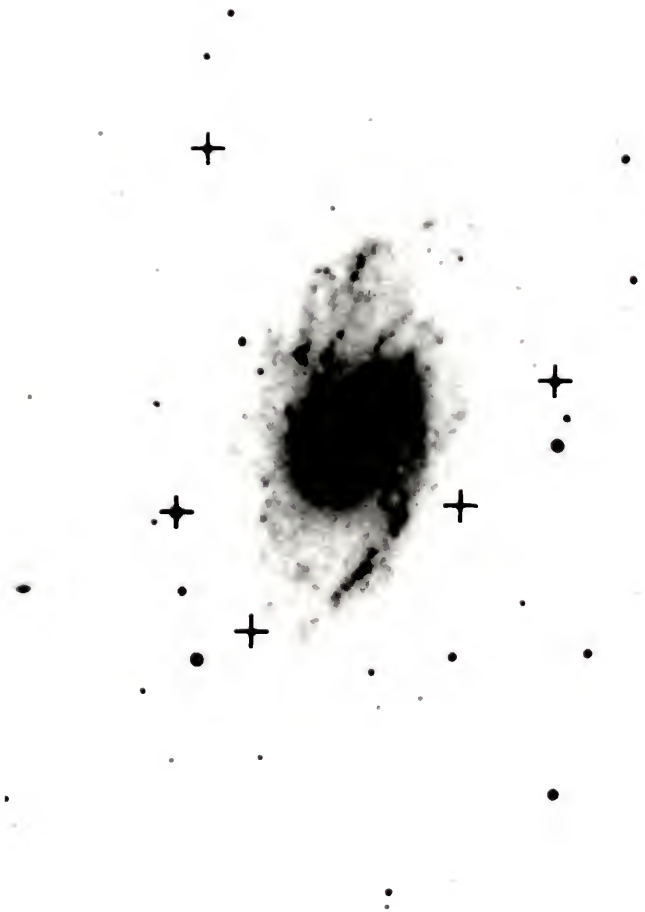


TABLE 1-1
GLOBAL PROPERTIES OF NGC 3359

Right Ascension[a]	$10^h43^m20^s.7$
Declination[a]	+63°29'12"
Morphological Type[b]	SBT5
Distance[c]	11 Mpc
Photometric Diameter[c] at 25 mag/(arc sec)^2	6!3
	20.2 kpc
Corrected Blue Magnitude[c]	10.60
Corrected Blue Luminosity[d]	$1.08 \times 10^{10} L_\odot$

[a]Gallouët, Heidmann, and Dampierre (1973).
[b]de Vaucouleurs, de Vaucouleurs, and Corwin (1976).
[c]de Vaucouleurs (1979)
[d]Calculated using above values of distance and magnitude, and with $M_B (\odot) = + 5.48$ (Allen, 1973).

Nevertheless, the open, two-armed spiral structure is quite similar to typical results of the hydrodynamical barred spiral models, so that one may reasonably hope to be able to model at least the first-order behavior of the system. This structure was one factor in the selection of NGC 3359 as our program object. The galaxy is also at a high, northerly declination, and is known to be bright in the 21 cm line of neutral hydrogen, the medium from which we will derive our kinematical data. Single dish observations of the system have been published by Roberts (1968), by Rots (1980), and by Fisher and Tully (1981). Early aperture synthesis maps of the hydrogen were made by Siefert, Gottesman, and Wright (1975) and by Gottesman (1982). These investigations produced images with resolutions of 72" and 36", respectively. As will be seen in Chapter II, the observations reported here improve on the latter figure by another factor of two. This was possible because of the exceptional sensitivity and resolution of the National Radio Astronomy Observatory's Very Large Array, which has only recently become available for 21 cm work. The resolution achieved here gives information on the kinematics and structure of NGC 3359 on a scale which is a small fraction of the galaxy's diameter. Therefore, an excellent opportunity is at hand to obtain data of sufficient resolution to test the predictions of the hydrodynamical models. That is our fundamental goal in the present work.

The remaining chapters of this dissertation form the body of our work on NGC 3359. Chapter II presents the observational data, with a brief introduction to the techniques used in obtaining them. In Chapter III, we will derive the distribution and kinematics of the neutral, atomic gas in the galaxy from these data. We will also determine certain

global properties of dynamical importance, and comment briefly on the relation of our results to other published data on the galaxy. Chapter IV introduces some further data, namely, near infrared surface photometry, which will be helpful in determining the driving potential in NGC 3359. A theoretical framework for using these data is also developed. The hydrodynamical modeling and its results are covered in Chapter V. Finally, Chapter VI gives a brief summary of the major conclusions of this work.

CHAPTER II
RADIO OBSERVATIONS

Neutral Hydrogen as a Kinematical Tracer

The first step in understanding the dynamics of the gas in a particular barred spiral galaxy must be to determine its present kinematical state. In this study of the barred spiral NGC 3359, we use observations, at radio wavelengths, of the column density and velocity distributions of neutral hydrogen for this purpose.

For the purposes of dynamical modeling, observations of any of several components of the program galaxy would suffice. Although the codes used are gas-dynamical, any component whose kinematics are closely correlated with those of the gas could be used as a tracer. Thus, optical observations of the hydrogen alpha line could be used, for example, because the H II regions whose velocities are measured are associated with young, early-type stars. The dynamics of these stars as a group will be quite similar to those of the gas from which they have only recently formed. The choice of neutral, atomic hydrogen gas, however, offers several important advantages. These result from the nature of the radiation from neutral hydrogen in the interstellar medium (ISM).

Interstellar space, in a disk galaxy, is filled by a tenuous medium having a number of characteristic phases, of varying density and temperature (McKee and Ostriker, 1977; Spitzer, 1978). One of the dominant phases consists of cool, intermediate density gas (approximate temperature and density: 100 K, 20 atoms cm^{-3}) whose primary constituent

is atomic hydrogen in the ground state. This ground state has hyperfine splitting into two levels separated by only 6×10^{-6} eV. The upper of these levels is occupied when the spin angular momentum vectors of the proton and the electron in the atom are aligned parallel to one another; in the lower state, they are antiparallel. The energy difference corresponds to a temperature of 0.07 K, far below the ambient temperature of the medium. Also, the radiative transition between the two levels is forbidden. Consequently, the time scale for collisional de-excitation is much shorter than that for radiative de-excitation, even in the low densities typical of the ISM. However, the large column density of hydrogen atoms along a typical line of sight through a galaxy makes the radiation observable. This line, at a wavelength of 21.1 cm, was predicted by van de Hulst (1945) and confirmed a few years later (Ewen and Purcell, 1951; Muller and Oort, 1951; Christiansen and Hindman, 1952). Observations of this line have since seen wide application in studies of our own and external galaxies. The latter work has been reviewed by Roberts (1975), by Sancisi (1981), and by Bosma (1981a, b; 1983), among many others.

The low temperature of the gas in the ISM causes the thermal Doppler width of the 21 cm line to be very small, although the line is broadened, by macroscopic turbulence, to velocity widths of several km/s (Kerr and Westerhout, 1965). The observed total line widths for external galaxies, therefore, are controlled by their rotational velocities. Because of this rotation, the optical depth in the line at any given velocity (characteristic of the bulk rotation of the galaxy at a given point) remains low for most lines of sight (Mihalas and Binney, 1981, p. 488).

Let us assume that the H I is optically thin, in this sense. Should this assumption be false, the observed brightness temperature T_B will approach the physical, spin temperature of the gas, which is typically about 100 K, as we have said. For our observations of NGC 3359, the highest brightness temperature recorded was 42 K, averaged over the beam. This implies an optical depth of about 0.55, which would lead to an underestimate in the column density of about 20% at this isolated peak. The underestimate would be less serious elsewhere. The assumption of small optical depth is nonetheless retained, since the optical depth structure, or "clumpiness," of the medium is not known. Then a simple integration, over velocity, of the observed brightness of the image suffices to determine the column density of the gas at any point (e.g., Mihalas and Binney, p. 489):

$$N_H(x,y)=1.8226 \times 10^{18} \int_{-\infty}^{+\infty} T_B(x,y)\, dv \ . \qquad (2\text{-}1)$$

If the velocity v is in km/s and T_B in K, equation (2-1) gives N_H in atoms cm^{-2}. Similarly, the first moment with respect to velocity gives a measure of the mean velocity at the point:

$$\overline{v}(x,y) = \frac{\int_{-\infty}^{+\infty} v T_B(x,y)\, dv}{\int_{-\infty}^{+\infty} T_B(x,y)\, dv} \ . \qquad (2\text{-}2)$$

The simplicity of these results is the first great advantage of neutral hydrogen observations for work of this kind. The other accrues from the way in which the gas is distributed in space. First, especially if one excludes regions furthest from the galactic centers, it is observed to be

among the flattest and thinnest of the disk components of both our own (Jackson and Kellman, 1974) and other galaxies (van der Kruit, 1981; also, compare Sancisi and Allen, 1979, to van der Kruit and Searle, 1981). Therefore, one may determine, with a reasonable degree of confidence, the two-dimensional position in the plane of the galaxy of any observed emission. Secondly, this phase of the intersellar medium is pervasive enough that the emission is seen by radio telescopes as being continuously distributed, even for relatively nearby galaxies. The main alternative source of kinematical data is optical spectroscopy, but in practice only small regions of ionized gas near the hottest stars can be detected by that technique (though the resolution is very good). The ubiquity of H I across the disk component allows one to assign a mean surface density and velocity to each point in the disk, whereas the distribution of the optical spectra is necessarily patchy.

The only other Population I component which gives rise to prominent spectral features is molecular gas, which may have mass comparable to, or even greater than, the H I. Unfortunately, the molecular hydrogen, which is presumably its dominant constituent, is very difficult to detect. The most easily detected transitions, those of carbon monoxide, fall in the millimeter-wave region, where observational difficulties are greater than at longer wavelengths, and where the first aperture synthesis arrays (see below) are only now becoming operational. Also, molecular gas is less broadly distributed in galactocentric radius than H I in most galaxies. It is usually far more concentrated to the inner disk and, often, even to the nucleus (Morris and Rickard, 1982).

Elementary Aperture Synthesis Theory

NGC 3359 was observed in the 21 cm transition using the Very Large Array (VLA) radio telescope of the National Radio Astronomy Observatory. The VLA is the largest and most sensitive of the radio arrays which operate by the principle of aperture synthesis, a technique for obtaining good sensitivity and excellent resolution in this wavelength regime. We will give a brief description of aperture synthesis below. Fuller discussions may be found in Fomalont and Wright (1974), Hjellming and Basart (1982), and Thompson and D'Addario (1982).

The fundamental building block of an aperture synthesis telescope is the two-element interferometer, as shown in Figure 2-1. Consider the case where each dish of this interferometer tracks a celestial source, whose direction is given by the unit vector $\underline{s}_0$. Let their separation in meters be written as $\underline{B}$. Then, for wavefronts from this source, there is a delay τ in arrival time at antenna 1 relative to antenna 2:

$$\tau = \frac{1}{c} (\underline{B} \cdot \underline{s}_0) \quad , \tag{2-3}$$

where c is the speed of light. Therefore, for monochromatic radiation of angular frequency ω , the signal received at antenna 2 is proportional to cos ω t, and that at antenna 1 is proportional to cos ω (t + τ). After detection and amplification, the signals from the two are multiplied together. In the simple scheme discussed so far, the result of this multiplication would be

$$R(t) \propto \cos(2\omega t - \omega\tau) + \cos(\omega\tau) \quad . \tag{2-4}$$

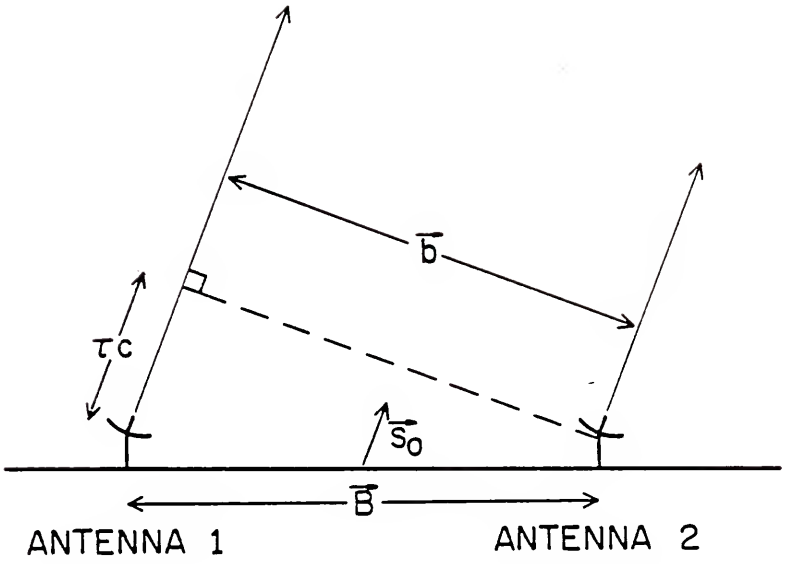


Figure 2-1. Geometry of the two-element interferometer. The unit vector $\underline{s}_0$ points toward the source being observed. Vector $\underline{b}$ is the component of the mutual separation $\underline{B}$ which is perpendicular to $\underline{s}_0$. Note that $\underline{b}$ and $\underline{s}_0$ need not lie in the plane of the paper. The additional propagation length to antenna 1 causes a relative delay for wavefronts traveling at velocity c.

When this signal is sent through a low-pass filter, only the second term remains. This is because the first term varies at a frequency twice that of the emission being observed, the second only at the rate of the sidereal change in the projected baseline $\underline{B}\cdot\underline{s}_0$ (in wavelengths), giving a period of order 1 second for 21-cm emission and typical VLA spacings.

In practice, the instrument must be made more sophisticated in a number of ways. The most fundamental is caused by the fact that astronomical radiation is not monochromatic, and neither are most receivers, so that a band of frequencies $\Delta\omega$ is always observed. The geometrical delay τ then corresponds to a different number of wavelengths for each frequency in the band, and the unavoidable integration over all these frequencies in the multiplier will cause the signal to be severely depressed in amplitude. The solution is to introduce a delay τ' into the propagation path from antenna 2 which almost cancels τ, at least at the band center ω_c, enabling coherence over broader bandwidths. This procedure is called delay tracking. Suppose the band contains signals between frequencies ω_l and $\omega_u = \omega_l + \Delta\omega$. Let τ_0 be the net difference in delay between paths 1 and 2: $\tau_0 = \tau - \tau'$. The change in phase, for frequency ω_i, over this delay is $\tau_0\,\omega_i$, and therefore the phase difference over the whole band at the end of the path is $\tau_0\,\Delta\omega$. If this difference amounts to 2π, all coherence is lost; the condition for satisfactory coherence is evidently

$$|\tau_0| \ll \frac{2\pi}{\Delta\omega} \quad . \qquad (2\text{-}5)$$

Of course, the effects of a finite bandwidth must still be taken into account by integrating the output signal over Δ .

Secondly, before multiplication the signals must be reduced to a lower frequency (usually called the intermediate frequency) than that received from the source. This is achieved by a heterodyne system, in which both antenna outputs are mixed with the signal from a monochromatic local oscillator (LO). A single sideband, of the two created by this mixing, is selected by appropriate filtering. This step is normally implemented before delay tracking. These considerations complicate the intermediate calculations considerably. However, if the band passes are stable, symmetrical, and identical in the two channels, and if delay tracking is acceptably good, one still obtains the remarkably simple expression (Hjellming and Basart, 1982, p. 2-15):

$$R(t) \propto \cos[\omega_0 \tau + \Phi(t)] \quad . \tag{2-6}$$

Here ω_0 is the frequency of the local oscillator, and Φ (t) the phase difference between the LO signal as received at the two mixers of the heterodyne system.

The formal extension of this result to a system of many antennas is straightforward. One simply replaces the local oscillator by a network of slaved local oscillators, one at each antenna, under the control of a master LO and rigorously synchronized. (The electronics system in use at the VLA actually uses a number of LOs in each signal path to optimize the frequency characteristics of each component in the system. However, there is no conceptual difference between this scheme and the use of a single LO.) Also, delay tracking is implemented for all antennas

relative to a single point (for example, the geometrical center of the array). In order to avoid the introduction of errors into the data by incorrect tracking, the positions of all of the individual telescopes must be known very accurately. This calibration is ordinarily performed when the configuration of the array is changed. It is carried out by observing point sources at a variety of hour angles and declinations, to obtain many measurements of $\underline{B} \cdot \underline{s}_0$ for different orientations of $\underline{s}_0$. $\underline{B}$ can then be estimated quite accurately for each antenna pair.

Of course, the delay tracking is strictly accurate only for a source at the pointing center $\underline{s}_0$. Suppose emission is detected from a source at some other position, $\underline{s}$. Let us write the instantaneous phase of the signal from the jth antenna, relative to the master LO, as Φ_j. This includes the phase shift associated with the delay tracking. Because $\underline{s}_0$ is the reference point for delay tracking, the correlated signal from the antenna pair j and k becomes

$$R_{jk}(t) = A_j(t)A_k(t)\cos\left[\frac{\omega_0}{c}\,\underline{B}\cdot(\underline{s}-\underline{s}_0)+\Phi_j(t)-\Phi_k(t)\right]. \qquad (2\text{-}7)$$

All of the quantities on the right-hand side of this equation should be quite slowly varying. The quantities A_j and Φ_j are called the instrumental amplitude and phase, respectively, of the jth antenna, and the expression

$$G_j = A_j \exp(i\Phi_j) \qquad (2\text{-}8)$$

is often referred to as the complex gain.

It is common to speak of the correlated signal obtained in this way as the "complex visibility function" $V_{jk}(t)$. This is simply the representation of equation (2-7) in complex-variable notation:

$$V_{jk}(t)=A_j(t)A_k(t)\exp\{i[\frac{\omega_0}{c}\underline{B}\cdot(\underline{s}-\underline{s}_0)+\phi_j(t)-\phi_k(t)]\}. \qquad (2\text{-}9)$$

By writing equation (2-9) as the product of independent terms for the two antennas, we have implicitly assumed that "antenna based" calibration is appropriate, as explained below. However, the quantities which are actually measurable are those specified by the equation

$$V_{jk}(t)=A_{jk}(t)\exp\{i[\frac{\omega_0}{c}\underline{B}\cdot(\underline{s}-\underline{s}_0)+\phi_{jk}(t)]\} \quad . \qquad (2\text{-}10)$$

The calibration of the data base consists of determining the complex gain of each antenna, as a function of time, from the observed quantities of equation (2-10). Observations of strong point sources, of well-known flux density and position, are used for this purpose. For a source of this kind, equation (2-10) becomes particularly simple. The assumption one normally makes is that $\underline{s} = \underline{s}_0$ for these calibrator observations; if this is not true, a phase gradient across the array will be introduced. Clearly, it is important that the positions of all calibrators be known as accurately as possible. With this assumption, equation (2-10) is simply

$$V_{jk}(t)=A_{jk}(t)\exp[i\phi_{jk}(t)] \quad . \qquad (2\text{-}11)$$

If this visibility were properly calibrated, it would be equal to the flux density S of the calibrator. Thus, one can rewrite equation (2-11) in terms of an effective complex gain for the jk baseline, G_{jk}:

$$V_{jk}(t)=G_{jk}(t)S \quad . \tag{2-12}$$

As we have stated, the usual assumption when dealing with VLA data, which is almost always valid (Hjellming, 1982), is that the complex gains of the antennas are separable. Then

$$G_{jk}(t)=G_j(t)G_k^*(t) \quad , \tag{2-13}$$

where the asterisk denotes the complex conjugate. Combining the expressions (2-8), (2-11), (2-12), and (2-13), we obtain the results

$$A_j(t)A_k(t) = \frac{A_{jk}(t)}{S}, \tag{2-14a}$$

$$\Phi_j(t)-\Phi_k = \Phi_{jk}(t) \quad . \tag{2-14b}$$

At each time for which calibration data are available, one has this pair of equations for each of the N(N-1)/2 baselines, so that the system of equations for the N antenna-based complex gains is highly over-determined. Therefore, a least-squares solution is appropriate. The residuals of $A_j(t)$ and $\Phi_j(t)$ in this solution are called the closure errors in amplitude and phase, respectively, for antenna j. Obviously,

the calibrator observations must be spaced closely enough in time so that the time dependence of $G_j(t)$ is known. One can then interpolate to find its value during the observations of the program object. For the VLA, the system is ordinarily stable enough so that a calibrator scan once every 30-45 minutes is adequate.

Before leaving the topic of calibration, we should mention the calibration of the antenna pointing. Like the telescope position calibration, this function is performed at regular intervals by the observatory staff. The errors in pointing for each antenna, as a function of altitude and azimuth, are determined by point source observations, combined with the known primary antenna power patterns. Corrections for these errors are then applied, in the on-line observing computers, to data taken subsequently for scientific programs. This is necessary for observing programs where highly accurate positioned information is sought, but is of little importance in the present work.

Now let us consider the response of the two-element system to an extended source of emission. As before, $\underline{s}_0$ gives the direction of the point on the sky which is tracked by the antennas; that is, the maximum of the single-dish power pattern P_0 is kept aligned with the changing position of $\underline{s}_0$. Let the vector $\underline{s}$ give the position of some emitting point within the field of view, not located at $\underline{s}_0$. Since the primary beam of each antenna will restrict this field to a relatively small solid angle about $\underline{s}_0$, $\underline{s}$ may be written as

$$\underline{s} = \underline{s}_0 + \underline{r} \quad , \tag{2-15}$$

where $\underline{r}$ is a vector perpendicular to, and much smaller than, the unit

vector $\underline{s}_0$. The complex visibility associated with this patch of emission will differ from that given in equation (2-9) for two reasons. First, we must now include explicitly the product of the source intensity distribution $I(\alpha, \delta)$ with the normalized, primary power patter P_0 $(\alpha-\alpha_0, \delta-\delta_0)$. (Here the equatorial coordinates (α_0, δ_0) give the position of $\underline{s}_0$.) This ensures that integration, over the solid angle of the source, of the apparent intensity yields the apparent flux density. The second, and more important, effect is that the delay τ is now a function of the vector $\underline{s}$, not $\underline{s}_0$. If we let $\underline{s}$ play this role in equation (2-9), recalling the definition of τ in equation (2-3), we obtain for the contribution from a small emitting patch of solid angle $d\Omega$:

$$dV_{jk}(\alpha,\delta,t)=A_j(t)A_k(t)I(\alpha,\delta)P_0(\alpha-\alpha_0,\delta-\delta_0)d\Omega \quad x$$

$$x \quad \exp\{i[\tfrac{\omega_0}{c}(\underline{B}\cdot\underline{r})+\Phi_j(t)-\Phi_k(t)]\} \quad . \qquad (2\text{-}16)$$

Before proceeding to simplify this expression, let us consider the geometry of the interferometer pair as viewed from the object under observation. Suppose, for simplicity, that the orientation of the baseline $\underline{B}$ is east-west in the earth's latitude-longitude system, as illustrated in Figure 2-2. From the viewpoint of our source, it is much more useful to describe this separation in the so-called (u,v) plane. This coordinate system has one axis (the v axis) parallel to the earth's rotational axis, while the other (u) is perpendicular to it and is broadside to the source, as shown. Clearly, if the object were located at the zenith as seen from the interferometer, then one could identify

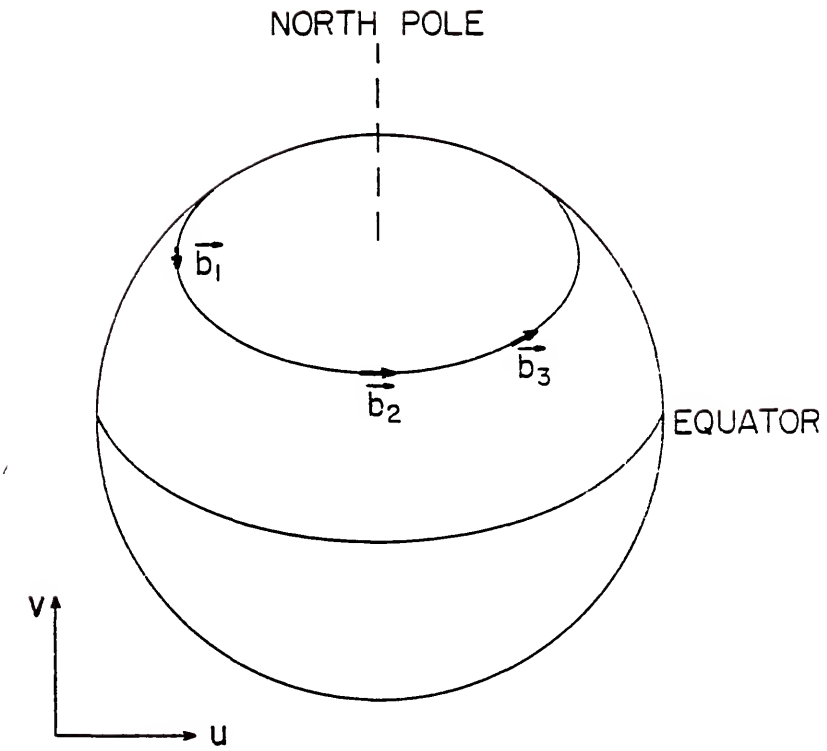


Figure 2-2. The (u,v) coordinate system. As viewed from a celestial object, an interferometer which is oriented east-west on the rotating earth has the projected baselines $\underline{b}_1$, $\underline{b}_2$, and $\underline{b}_3$ at three different times.

the u and v directions with east and north, respectively, on the earth's surface. In general, however, the transformation between the two systems depends on the interferometer's location on the rotating earth as seen by the source, and is therefore a function of time. It is apparent from Figure 2-2 that both the orientation and the length of the projected baseline seen by the source will change as the earth rotates, for a fixed pair of antennas. Thus, without the need to move either dish, one samples the visibility function with a wide variety of effective spacings. This is the basic principle of earth-rotation aperture synthesis.

Figure 2-2 suggests that it may be helpful to recast the observing geometry in terms of $\underline{b}$, the projection of $\underline{B}$ in the (u,v) plane. This quantity is shown both in Figure 2-2 and in Figure 2-1. In the latter, it is obvious that the vector baseline $\underline{B}$ may be written

$$\underline{B} = \underline{b}+\tau c\underline{s}_0 \quad . \tag{2-17}$$

Therefore,

$$\underline{B}\cdot\underline{r} = \underline{b}\cdot\underline{r}+\tau c\underline{s}_0\cdot\underline{r} \quad .$$

Since $\underline{s}_0$ and $\underline{r}$ are mutally orthogonal vectors,

$$\underline{B}\cdot\underline{r} = \underline{b}\cdot\underline{r} \quad , \tag{2-18}$$

and so equation (2-16) becomes

$$dV_{jk}(\alpha,\delta,t)=A_j(t)A_k(t)I(\alpha,\delta)P_0(\alpha-\alpha_0,\delta-\delta_0)d\Omega \quad x$$

$$x \quad \exp\ \{ i[\ \frac{\omega}{c}(\underline{b}\cdot\underline{r})+\Phi_j(t)-\Phi_k(t)]\ \} \quad .$$

(2-19)

Integration over the solid angle of the source gives

$$V_{jk}(t)=A_j(t)A_k(t)\exp\{\ i[\Phi_j(t)-\Phi_k(t)]\ \} \quad x$$

$$x \quad I(\alpha,\delta)P_0(\alpha-\alpha_0,\delta-\delta_0)\exp[i(\ \frac{\omega}{c}\underline{b}\cdot\underline{r})]d\Omega \quad .$$

(2-20)

This expression gives the response of each two-element interferometer in the array to an extended source of emission. Of course, in an actual observation a great deal of noise, generated both by the atmosphere and in the instrumental system, also will be detected, and this has been neglected in our analysis. However, this noise should not be correlated between antennas, and so, over a long enough period of time, it should average to zero for each antenna pair. In practice, statistical deviations from this ideal, together with the variations in the amplitude and phase gain of each antenna, will generally cause a small, but nonzero, correlation of this noise component for each baseline. If these effects are more or less equal for all the visibility records, as they should be, the noise in the final map will be spatially random. Noise which does not satisfy this condition can create map artifacts, and should be edited out of the data base, if possible. An example is interference, although the rejection of uncorrelated signals gives

aperture synthesis data some degree of protection from many types of interference.

The factors outside the integral in equation (2-20) are not of astronomical interest, as they contain only calibration factors in amplitude and phase. The integral itself is accordingly sometimes referred to as the visibility function; let us denote it by V'. By inspection, V' is the Fourier transform of $I'(\alpha,\delta)$, where I' is the product of the true brightness distribution I and the single-dish power pattern P_0. If the source brightness is specified in a coordinate system (x,y), which is centered on (α_0,δ_0) and parallel to the (α,δ) system at the source, V' can be written in the form of a standard two-dimensional Fourier integral (Fomalont and Wright, 1974, p. 261):

$$V'(u,v)=\int\int I'(x,y)\exp[i2\pi(ux+vy)]\,dx\,dy \quad . \tag{2-21}$$

This equation assumes that the field of the view is small enough that the curvature of the sky plane can be neglected.

The quantities u and v in this expression can be identified with the components of $\underline{b}$ in the (u,v) plane, provided that the latter are measured in wavelengths of the emitted radiation so that they are dimensionless. Equation (2-21) implies another interpretation of u and v: they play the role of "frequencies" in the spatial Fourier "spectrum" V' of I'. Consequently, they are usually called spatial frequencies. In light of this interpretation, we see that a single measurement of the complex visibility function of two antennas at a given time, corresponding to a particular projected baseline and hence a particular (u,v), gives a single Fourier component of I'. Because of the similarity theorem of Fourier transforms (Bracewell, 1965, pp. 101-104), features which have large extent in the (x,y) domain will have small extent in

the (u,v) domain, and vice versa. Therefore, in order to obtain high angular resolution, one must have wide spacings present in the array. Conversely, the detection of large-scale structure requires the inclusion of low spatial frequencies (i.e., small spacings are required).

Suppose we define one antenna of our pair as the reference position and plot the relative position of the other in the (u,v) plane. As time goes by, this point will sweep out an arc whose shape depends on the orientation of the baseline $\underline{B}$ on the earth and the declination of the source. Since the choice of the reference antenna is arbitrary, the arc which is diametrically opposite will also be traced out. In effect, two Fourier components are measured at once. However, no more information is gained because I' is a real function, and therefore its Fourier transform is Hermitian (Bracewell, p. 16). Hence it is completely defined by its values over half the plane. An example of the arcs produced for a particularly simple geometry is shown in Figure 2-3. When many antennas are present in an array, and all their relative separations are plotted in this way, the entire (u,v) plane becomes covered with these tracks, indicating that very many of the Fourier components of the source brightness have been measured. The array has then approximately synthesized the Fourier sampling of a single filled aperture equal to the greatest separation in the array. This is the origin of the term "aperture synthesis." However, spacings less than the physical diameter of the dishes are missing from the synthesis. The importance of this point is discussed below.

Intensive coverage of the (u,v) plane is necessary, because the quantity of interest is actually, of course, the intensity I in the sky plane. The primary beam P_0 is usually known well enough so that I can

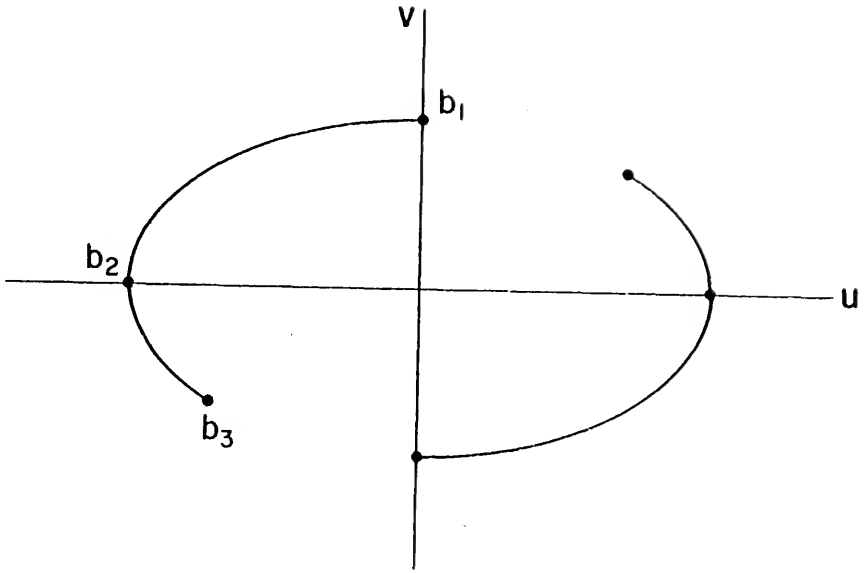


Figure 2-3. The tracks swept out in the (u,v) plane, over the course of several hours, by a single interferometer pair. The points labeled b_1, b_2, b_3 correspond to the three interferometer positions shown in Figure 2-2. This coverage is typical of a northerly source observed by an east-west interferometer.

be recovered from I' if the latter is available. However, I' must be calculated by inverting the Fourier transform of equation (2-21):

$$I'(x,y)=\iint V'(u,v)\exp[-i2\pi(ux+vy)]\,du\,dv\ . \qquad (2\text{-}22)$$

Since V' is measured only at a finite number of points, and each measurement includes observational noise, I' cannot be determined uniquely, nor without error. We will return to a discussion of this inversion later in the present chapter.

We have now described all of the fundamentals of this observing technique which bear directly on the present research, except for the acquisition of high-resolution spectral data. Good spectral resolution of the signal received by an aperture synthesis instrument requires that one of several possible methods be used to divide the emission into a number of independent, narrow-band spectral channels which can be processed separately. While all these methods are conceptually equivalent, in practice the necessity of holding the number and complexity of electronic components to a manageable level has dictated the use of correlation receivers. A thorough introduction to these devices, with some attention to practical design considerations, is given by D'Addario (1982). The mathematical analysis and the electronic implementation of this type of system are quite complex, as Fourier transforms in both the time/frequency and space/spatial frequency domains are encountered. Here we will only outline the basic principle of their operation, beginning with a qualitative demonstration of the underlying reasoning.

In our discussion of delay tracking, the compensating delay served the purpose of restoring the time coherence of the wavefront received

at the two antennas. Indeed, in the idealized case of monochromatic incident radiation, if one had perfect, continuously adjustable delay tracking, and no atmospheric or ionospheric distortions of the wavefront, the delay τ' would restore full coherence to the wavefront. Except for the contribution due to background noise, the outputs of the two receivers would then be perfectly correlated for a point source at the tracking center $\underline{s}_0$. Now suppose that a relatively large additional delay τ_l is inserted, after τ', in this blissful arrangement. Obviously, the effect is to destroy the coherence of the two signals at the inputs to the multiplier, resulting in zero signal detected after correlation. The only exception to this will be for a very narrow range of frequencies centered on some frequency ω_i, for which the shift corresponds to a complete cycle of 2π in phase. In effect, the introduction of the lag τ_l has acted as a narrow-pass filter, isolating a single spectral channel from the band $\Delta\omega$. The use of many such lags allows one to divide the signal into many channels. The frequency resolution of the system is controlled by the maximum lag. The most common correlator scheme, which is used at the VLA, employs lags of both positive and negative sign in equal numbers, and spaced at equal intervals $\Delta\tau_l$. If there are a total of N_l of these, the frequency resolution is $(CN_l\Delta\tau_l)^{-1}$, where C, of order unity, depends on the weighting used in the frequency Fourier transform (Baldwin et al., 1971).

Mathematically, there is no distinction between τ_l and the other various delays which have been considered previously. We can rewrite equation (2-21) to allow for its presence:

$$V'(u,v,\tau_l)=\iint I'(x,y)\exp\{i[\omega\tau_l+2\pi(ux+vy)]\}\ dx\ dy\ .$$

(2-23)

If we integrate over the bandwidth, we find the integrated visibility V_I':

$$V_I'(u,v,\tau_l)=\iiint I'(x,y)F(\omega)\exp\{i[\omega\tau_l+2\pi(ux+vy)]\}\ dx\ dy\ d\omega\ ,$$

(2-24)

where F( ) is the bandpass function. Because of the symmetry of the lag spectrum, only the real part of V_I' need be Fourier transformed (Hjellming and Basart, 1982, p. 2-37). When we take the Fourier transform with respect to the lag τ_l, we obtain

$$\mathrm{Re}(V_I')\exp(i\omega\tau_l)d\tau_l=\iint I'(x,y)F(\omega)\exp[i2\pi(ux+vy)]\ dx\ dy\ ,$$

(2-25)

where the right-hand side must now be evaluated at one of the narrow-band frequences ω_i determined by the lag spectrum. Therefore, the extension of ordinary aperture synthesis techniques to spectral line work involves three principal complications, the first two of which are resolved in the design of the instrument and are not of immediate concern to the user: (1) the need for a complicated correlating receiver with provision for a large number of digital lags; (2) the need to Fourier transform the lag spectrum once per integration period, as implied by equation (2-25); and (3) the need to calibrate the band pass $F(\omega)$. The latter is accomplished by observing a strong continuum source, which is assumed to exhibit no spectral variation over the quite narrow total bandpasses normally used in spectral line observations.

VLA Observations of NGC 3359

The NRAO Very Large Array was used to observe the 21 cm line in NGC 3359. The VLA is an earth-rotation aperture synthesis instrument, consisting of 27 fully steerable dishes of 25 m aperture apiece (Thompson et al., 1980). Nine antennas are deployed along each of the three arms of a Y-shaped pattern. The arms are oriented 120° apart, approximately in the north, southeast, and southwest directions. This arrangement causes the set of baseline vectors at any given time to cover a wide range in azimuth, providing good (u,v) coverage. The radial spacing along an arm increases outward from the array center. This gives denser sampling of the (u,v) plane near its origin, which results in better sensitivity and sidelobe suppression. We will return to this point in our discussion of tapering.

The spacings of the antennas can be changed to allow observations at different resolutions, but remain fixed during a single observing run. There are four standard sets of antenna positions. All have roughly the same relative spacings, but the overall scale of the array is changed from one to another. The choice among these configurations involves a compromise between resolution and sensitivity requirements of the observing program. The similarity theorem dictates that the highest resolutions can only be achieved with large separations. On the other hand, it also requires the inclusion of short spacings if broad structure is suspected to be present. Even more importantly, high resolution observations have relatively poor brightness sensitivity. Although the minimum detectable flux density ΔS depends only on the system temperature, bandwidth, integration time, and total collecting area, for mapping a resolved source it is the detectable brightness which is important.

This is related to ΔS by (e.g. Kraus, 1966, p. 102)

$$\Delta B = \Omega_{bm}^{-1} \Delta S \quad , \tag{2-26}$$

where Ω_{bm} is the solid angle of the "synthesized beam," in other words, the power pattern of the array as a whole. The synthesized beam is thus equal to the normalized brightness distribution observed for a point source. The complementary relation expressed by equations (2-21) and (2-22) indicates that the beam is smaller when larger (u,v) spacings are used. Consequently, the sensitivity to extended structure is degraded whenever the resolution is improved, and vice versa.

For this project, the two smallest available configurations, the so-called C and D arrays, were used. Since narrow-band spectroscopy divides the available signal power from the galaxy into many channels, sensitivity tends to be poor. The C configuration was expected to be the largest array with adequate brightness sensitivity. This is demonstrated by our results. The best peak signal-to-noise ratio for any of the maps is about 20. The next largest or "B" array synthesizes a beam of only one-tenth the solid angle of that of the C array. Thus, we could have achieved a "two sigma" detection, at best, with the B array. The C configuration contains spacings from a maximum of about 3 km, or roughly 15,000 wavelengths at 21 cm, down to approximately 100 m, but the inner spacings are rather sparsely sampled. Data from the smaller D array were desired as well, to ensure that no low-amplitude, large-scale emission was missed. The separations in the D array vary from about 40 m up to just under a kilometer, overlapping the inner C spacings fairly heavily. The total range of available spatial frequencies, therefore,

gives sensitivity to structure on scales from about 14" to 17'. If spatial wavelengths larger or smaller than this are present within a single spectral channel, they will not be detected. Naturally, a real object is not likely to be composed of a single spatial frequency. It is, perhaps, more relevant that a source whose brightness distribution is Gaussian will be at least 50% resolved, if its full width at half power lies between about 7" and 8' (see Figure 10.A4b of Fomalont and Wright, 1974). Of course, it is certain that fine-scale structure is indeed missed in this way. The lower limit, in practice, was increased from 14" to about 18" because of the taper employed (see below). Since the (u,v) coverage was fairly uniform at all azimuths, these limits should not depend strongly on the shape of the source.

These spacings are unprojected, and are therefore upper limits. The physical size of the antennas places a lower limit on the available (u,v) spacings. When the projected separation of two antennas becomes less than the dish diameter of 25 m, one of them is partially blocking the other's view of the source. This "shadowing" effect can be corrected in the calibration and editing of the data. However, a more insidious result of this situation is "cross-talk," in which the shadowed antenna detects signals from the electronics of its neighbor. In practice, it has been found that in every case where even slight shadowing occurs, there is a quite noticeable deterioration in signal quality, attributable to cross-talk. All such data were simply discarded. The problem only arises when the elevation of the source is low, so for NGC 3359 (δ=+63°.5) very few data are involved.

The observations of NGC 3359 were obtained on 27 January 1983, with the C array, and on 21 June 1983, with the D array. In most spectral

line applications, it is not possible to use all the antennas because of the limitations of the correlator. The correlator must multiply the signals from 2n delay lines, for each of N(N-1)/2 baselines, where n is the desired number of channels and N the number of antennas. The details of the particular correlator used, therefore, place an upper limit on the product nN^2. The choice of n is dictated by two conflicting astronomical requirements. The larger n is, the better one's spectral (and hence velocity) resolution, but the poorer the sensitivity. At a minimum, n must be large enough to cover the entire velocity width of the 21 cm line. For NGC 3359, past experience and inspection of previously published global 21 cm profiles indicated a choice of approximately 20 km/s for the single channel width. Rots (1980) had measured a width of 260 km/s, at the 25% level, for the global spectrum. To be certain of detecting any faint emission at extreme velocities, it is necessary to allow a margin beyond this. It is also desirable to have a "baseline" of a few signal-free channels. Since n is restricted to being a power of two, so that the Fourier transform of the lag spectrum can be calculated with Fast Fourier Transform (FFT) techniques, n=32 was chosen. During the January observations, this necessitated restricting the number of antennas used to 21. Because of the northerly declination of NGC 3359, the southeast and southwest arms are the most important for getting good coverage of the (u,v) plane. Of the six antennas to be omitted, four were selected from the northern arm, and one from each of the others. By June, the correlator had been upgraded, and only two antennas had to be omitted. These were simply chosen on the basis of their recent malfunction history.

In each observing run, the data were calibrated in amplitude and phase, as described in the previous section, using a somewhat complicated calibration scheme. The purpose of this calibration is to correct the slow variations in the system response to a point source of constant flux. These variations are of two types: those arising somewhere in the instrument, and those imposed by fluctuations of the atmosphere. The latter are usually more severe at the VLA. Consequently, it is important that the calibrator source be near the program source in the sky, and especially that it be at nearly the same elevation. However, the number of available calibrators is small, since few bright radio sources are both unresolved and constant in flux. For NGC 3359, it proved helpful to bracket it in declination between two different calibrators. A third calibrator had to be used to complete this scheme, on account of the differing hour angles of the calibrators and NGC 3359. Finally, the bright source 3C 286 did double duty as the "primary" calibrator, i.e., the fundamental standard used to determine the flux densities of the other or "secondary" calibrators, and as the bandpass calibrator. In the C array observations, 3C 386 was observed twice, once on either side of transit. This made it possible to optimize the choice of switchover time between secondary calibrators, while ensuring that all of them had primary calibration. It also guarded against losing the primary calibration for any antennas which happened to be experiencing technical malfunctions during the time of primary calibration. Unfortunately, the hour angle made it impossible to do this for the June observations. In both cases, the basic observing cycle alternated thirty-minute integrations on NGC 3359 with much shorter scans of the calibrators. The duration of the latter

depended on the flux density of the calibrator, varying from three minutes for 3C 286 to eight for secondary calibrator 1031+567.

There are three relevant integration times for these data. The one of ultimate interest is, of course, the total amount of time spent observing the program source. The integration time of immediate concern in the on-line computer system, however, is the time for which the output signals from the multipliers are averaged, before the FFT of the lag spectrum is taken. The latter step produces a single estimate of the multi-channel spectrum of the source for each baseline. Subsequently, this signal is itself averaged for a slightly longer time to reduce the volume of data to manageable levels. These times were 20 and 40 seconds, respectively, for both observing runs on NGC 3359.

The implementation of the calibration procedure outlined in the previous section is conceptually straightforward, in the case where there are no unusual problems in the data. Primarily for calibration purposes, a "pseudo-continuum" channel is generated along with the narrow band channels. This channel is so called because it contains the broadband signal from the central 75% of the original intermediate frequency bandpass; it would be more accurate to term it the continuum-plus-line channel. In the present case, this "channel 0" has a bandwidth of approximately 4.7 MHz, compared to the frequency separation of about 98 kHz for the spectral line channels. Therefore, the sensitivity to the calibrators is almost seven times better in this channel than in the individual line channels. The calibration is determined for channel 0, and then applied to the line channels, along with the bandpass calibration. The detailed procedure is as follows.

First one forces the flux density of the primary calibrator to assume its well-known value at the frequency of observation. A solution for the amplitude and phase of each antenna in the array, as a function of time, is then computed, using all the scans of the various calibrators. The flux densities of the secondary calibrators are used as free parameters in the amplitude solution. When this has been done, baselines with large closure errors in the amplitude or phase solutions can be identified. For the observing mode being discussed here, one often has a few such errors of greater than 10% in amplitude or 10° in phase. It is best to delete such baselines from the data for the span of time in which they are troublesome, and repeat the solution. After iterating this procedure until the data seem acceptable, one can use the flux density of the primary calibrator to determine those of the secondary calibrators. The latter are said to be "bootstrapped" fluxes. Their errors are indicative of the overall stability of the array-atmosphere system during the observing run. The fluxes of all calibrators, and the total on-source integration time for each, are listed for both observing runs in Table 2-1. The source names used for the secondary calibrators are the IAU designations. The integration time is also given for NGC 3359 itself, for comparison.

Once this stage has been reached in the calibration process, the observer must decide whether the particular solution for the antenna gains, which he now has in hand, is satisfactory. Since the solution uses a particular reference antenna to establish the phase reference for the array, it is well worth one's while to try to select one which is particularly stable. This is done by listing the phases and amplitudes of the other antennas relative to the reference antenna, after the

TABLE 2-1
CALIBRATION PROPERTIES

Data Base	Object	Type[a]	Flux Density, Jy	Integration Time, min
Jan. 1983,	3C 286	PC	14.76[b]	22
C Array	0836 + 710	SC	4.172 ± .051	20
	1031 + 567	SC	1.883 ± .004	73
	1311 + 678	SC	2.529 ± .015	21
	NGC 3359	PO	-	456
June 1983,	3C 286	PC	14.755[b]	13
D Array	0836 + 710	SC	4.268 ± .032	12
	1031 + 567	SC	1.938 ± .013	55
	1311 + 678	SC	2.539 ± .013	28
	NGC 3359	PO	-	290

[a]PC = Primary Calibrator, SC = Secondary Calibrator, PO = Program Object.

[b]Assumed flux density, found by interpolating the results of Baars et al. (1977) at the frequency of observation.

solution has been made. The hope is to find an antenna whose variations, compared to the ensemble of all the other antennas in the array, are slow, as small as possible, and not monotonic functions of either space or time. If the first antenna chosen does not meet these criteria, one can usually improve the solution by choosing another reference antenna and repeating the entire process from the beginning.

When a good solution has been found, it is applied to the entire data set, including the observations of the program source, by a simple running mean, or "boxcar," interpolation of the computed amplitude and phase gains of the individual telescopes. Inspection of the quality of the solution at this point can identify more baselines which may merit removal. Finally, the bandpass is calibrated by assuming a flat spectrum, over the small total spectral-line bandwidth, for the primary calibrator. Then the calibration is complete, and the data may be Fourier inverted to give maps.

Unfortunately, both the C array and the D array observations of NGC 3359 were affected by peculiarities which complicated the calibration and editing considerably. The spectrum of the bandpass calibrator, 3C 286, in the C data was afflicted by a rather sharp drop, mimicking a spectral line, of about two channels in width. The drop in flux was about 2%, many times the mean noise across the band. Subsequently, the variation was found to be caused by a hardware problem in the Fluke synthesizer, which controls the precise frequency of observation. It was not possible to isolate the bad data in the (u,v) plane , and the effect eventually had to be corrected in the mapping process.

The D array observations were rather adversely affected by the presence of interference from the sun. The sun is a bright source at

21 cm, and its sidelobes can create difficulties, especially for the type of observing program described here. The sidelobes will drift through the field being observed, but normally when one is well away from the position of the sun in the sky, they will be washed out by bandpass smearing. For spectral line observations, of course, this helpful effect of the bandwidth is greatly reduced. Also, since the sidelobes are large-scale features, they are most deleterious at the shortest baselines, so that the D array is the most susceptible of the VLA configurations. Although the sun was some 71° from NGC 3359 in the sky during the D observations, strong interference was nevertheless found in the data. Its most obvious manifestation was the presence of very large fluctuations in the amplitudes and phases of the individual baselines after applying the calibration solution. That noise ceased abruptly at sunset. This conjecture as to the nature of the interference was also confirmed by its being tied to specific baselines, rather than to antennas, and primarily to the shorter baselines. It was decided, again, that the problem was best dealt with in the map plane. The only solution to this type of interference is to remove baselines from the data set, but if one simply deletes all the short baselines from the run, the (u,v) coverage is badly degraded. The approach must be to reject the worst offenders until the effect is no longer visible in the maps. While still in the calibration and editing process, however, the amplitudes and phases were studied carefully, and a list of recurrent problem baselines was prepared. These provided a starting point for flagging after the data were transformed into the map plane. It was found that most of the baselines of less than 60 m had at least sporadically large errors of this type. The worst of these was the second shortest baseline in the array, at

40 m unprojected separation. For that antenna pair, the problem was so pervasive that the baseline was removed in the editing mode to allow a good calibration solution.

There are several decisions to be made in the Fourier inversion of the visibility data, for each channel, to give the single channel maps. These concern primarily the compromise between sensitivity and resolution referred to in the previous section, and they arise in the assignment of certain weights to the data before transforming. Obviously, the calibrated visibility data are very poorly distributed in space, for the purposes of computing their Fourier transform. As the number of individual points is on the order of 250,000 in the present instance, FFT techniques are a necessity. These require that the data be arranged in an m x l rectangular grid, where m and l are both powers of two. In the present work, we have used a square grid (m = l), as the sampling is quite similar in the u and v directions. The visibility data, on the other hand, lie on elliptical arcs, with spacings between measurements determined by the rate of change of the projected (u,v) separation for each baseline. This situation is handled by computing an approximate representation of the data on the desired square grid, where each tabulated value is found by averaging nearby observed points in some way. To do the averaging, we convolve these points with a function C(u,v) whose Fourier transform is rather flat to some radius and then falls off rapidly. However, it should not have high sidelobes beyond the map area, which disqualifies the simple "two-dimensional boxcar," or "pillbox," function. (This requirement helps suppress "aliasing," onto the map, of sources which lie outside the field of interest; Sramek, 1982.) In our maps, the convolving function used was the product of an exponential and a sinc function, i.e.

$$C(u,v) = C'(u)C'(v)$$

where

$$C'(u)=\exp\left(\frac{-|u|}{2.52\,\Delta u}\right)^2 \cdot \frac{\sin(\pi u/1.55\Delta u)}{(\pi u/1.55\,\Delta u)} \quad , \qquad (2\text{-}27)$$

and Δu is the (u,v) "cell size," that is, the spacing of the square grid to be formed. The values of the numerical constants have been chosen to optimize the detailed shape of the Fourier transform of this function. Sramek gives a thorough discussion of the factors affecting the choice of C(u,v), as well as the weighting functions discussed below. After convolution, the map is then sampled by the two-dimensional "shah" function (Bracewell, 1965, p. 214) to obtain the gridded values, and the FFT can be computed.

Before the transform is performed, one may wish to assign some additional weights to the gridded data. In general, there are two kinds of weighting employed, the first of which is global in nature, the second local. The global weighting is called tapering, and consists of multiplying the weights of all points by a factor which decreases at greater distances from the origin of the (u,v) plane. The purpose of tapering is to decrease the small scale sidelobes of the array by reducing the importance of the measurements at large spacings, which determine the high spatial frequency structure of the map. Tapering is usually effective at reducing sidelobes because these outer portions of the (u,v) plane are less densely filled with data, and hence less well determined. Another way of viewing the situation is in terms of the "dirty beam," which is the response of the array to a point source. For

extended observations with the VLA, the dirty beam usually consists of a small, Gaussian core, with broader sidelobes at the level of one to ten percent. The shape of these sidelobes is simply the Fourier transform of the unsampled spacings in the (u,v) coverage (Ekers, 1982, p. 12-10). This statement is strictly true if we include, in the unsampled spacings, those at radii from the outer edge of the sampled aperture to infinity, whose Fourier transform is the Airy diffraction disk. The low density of observed points in the outer (u,v) plane is therefore the direct cause of the strong inner sidelobes, which is why tapering to reduce the importance of these measurements improves the sensitivity. However, it is equally obvious that it does so at the expense of the resolution. A simple Gaussian taper is normally chosen, if one is used at all.

One can also assign a weight to each cell, based upon some measure of the expected signal-to-noise ratio within that cell. In one scheme, the weight is simply proportional to the number of records in the cell. This is called "natural weighting." It is useful for observations where the signal-to-noise ratio needs improvement, but it tends to lead to a significant loss of resolution because it weights the center of the (u,v) plane very heavily. The other common approach, "uniform weighting," simply assigns equal weight to all nonempty cells. When this weighting method is adopted, the beam characteristics are controlled by the taper (Sramek, 1982).

Finally, one must choose the (u,v) cell size Δu. From a consideration of the gridding in the (u,v) plane, it is obviously desirable to have the product of Δu with the number of cells on a side, m, be neither so large that the outer cells are all empty and the inner ones heavily undersampled, nor so small that many points at large

spacings are thrown away. Furthermore, Δu is related to the cell spacing Δx in the map plane by

$$\Delta u \Delta x = \frac{1}{m} \qquad (2\text{-}28)$$

where Δx is in radians. For the VLA, an empirical relation which produces good sampling is that the synthesized beam should be three to four times Δx.

After the transform has been calculated, producing a dirty beam and its "dirty map," one can try to remove faulty data, such as those alluded to in the discussion of calibration, by analyzing the image in the map plane. As pointed out by Ekers (1982), many serious errors are difficult or impossible to detect in the (u,v) plane. Conversely, many errors which seem significant in the (u,v) plane produce no detectable features in maps, and so need not be pursued further. The two most serious difficulties in the NGC 3359 observations--the hardware-related bandpass glitch in the C array data, and the solar interference in the D array data--behave in the former way, for the most part. In both cases, the most seriously affected of the single-channel maps had readily visible artifacts. These consisted of fairly high-frequency, concentric rings in the first instance, and of low-frequency, linear ripples in the second. The solution in each case was to isolate a group of possible culprit baselines, as found from the (x,y) wavelength of the image defects, and to delete each of them in turn, checking whether the image was then noticeably improved. This method worked especially well for the solar interference. The removal of a second baseline solved the problem

entirely. That no more drastic remedy was required, in view of the apparent problems in the calibrated amplitudes and phases mentioned earlier, probably can be attributed to the comparatively short duration of most of those (u,v) errors. For the C array data, a similar procedure was followed, and again only two baselines were deleted. In this case, although the other affected channels were corrected quite well, the most severe example, that at a central velocity of 1119.4 km/s, could not be wholly restored, and some very low-amplitude ripple remains in the map. However, the level of this spurious emission is well below the threshold which would be detected as a real signal by the routines used to calculate mean densities and velocities. Therefore, the contamination of the final integrated maps by this feature should not be too important.

After one has tranformed the visibility data into a set of single-channel maps, these often suffer noticeably from contamination by the sidelobes of the dirty beam. An equivalent way to state this is that the map is flawed, owing to an unrealistic estimate of the brightness at unsampled spatial frequencies (Högbom, 1974; Cornwell, 1982). In particular, the straightforward application of the FFT upon the visibility records gives an estimate of zero brightness at all unmeasured (u,v) positions, wherever they occur. In one sense, this estimate is the truest to the available data, since it assumes absolutely nothing about the source which is not measured. For this reason, it is called the principal solution. On the other hand, it is obviously implausible, and it would be very desirable to get a better estimate of the true brightness distribution at the unmeasured positions. To put it another way, we wish to remove the effects of the sidelobes of the dirty beam.

There are several ways to attempt this. In the case of NGC 3359, we have used the well-known "CLEAN" program, devised by Högbom (1974) and modified by Clark (1980). This program, in the form used at the VLA, has been discussed extensively by Cornwell (1982) and by Clark (1982). The algorithm of CLEAN performs an explicit correction for the effects of the dirty beam. In essence, CLEAN simply subtracts the entire dirty beam, including sidelobes, at the position of greatest brightness in the map. Then, the remaining map is scanned for its brightest point, where the procedure is repeated, and so forth. This process is continued until it is decided that the remaining emission on the map is simply noise. The best estimate of the true brightness is then considered to be the set of points that CLEAN has selected, each with its proper brightness, but without sidelobes. When CLEAN has been applied properly to a map, the "blank" areas of that map should have no visible sidelobe structure, only random noise. This criterion, judged fairly subjectively for a few example maps, was used to choose an iteration limit for CLEAN and all channels were then processed in a uniform manner. An empirical check on the validity of this process is provided by measuring the rms noise of an apparently blank region of sky. This value should be nearly the same in each channel, and should not exceed by more than factor of about two the following empirically predicted value (Rots, 1983), which holds for spectral line maps made with natural weight:

$$\sigma_p = a[N(N-1)T_i\Delta\nu]^{-1/2} \quad \text{mJy per beam.} \qquad (2\text{-}29)$$

Here N is the number of antennas used, T_i is the total on-source integration time in hours, $\Delta\nu$ is the narrow bandwidth in kHz, and a is a constant which is equal to 620 in the 21 cm band. Taking N = 23,

T_j=12.4, and $\Delta\upsilon$=98, we have for our observations σ_p=0.8 mJy per beam solid angle. The measured value of about 1.1 mJy per beam is quite acceptable.

The CLEAN algorithm is complicated in several ways by the requirements of stability and accuracy of the final or "clean" image. First, only a small fraction--the so-called loop gain--of the brightness at the selected point is actually multiplied by the dirty beam, and subtracted, at each step. Second, it is best to restrict the area which CLEAN searches for flux to be cleaned, in order to restrict the number of degrees of freedom available in fitting the data (Cornwell, 1982). The most common practice is to set a few rectangular boxes, often consisting of the inner quarter of the field, augmented by any other areas of obvious signal. CLEAN is allowed to find real flux only within these boxes. Finally, the "clean beam" must be mentioned. As explained above, the end result of CLEAN is that the observed flux density has been distributed into a set of discrete spikes of varying intensity. This is not felt to be a very good representation of the true brightness distribution in a typical, extended astronomical source. Some sort of spatial smoothing of these spikes must be performed. This is accomplished by convolving them with a "clean beam," consisting of an elliptical Gaussian function, fit to the inner portions of the dirty beam. The convolved images, hereafter designated "clean maps," will form the basis for our discussion of the neutral hydrogen characteristics of NGC 3359. Note that the resolution of the clean map is controlled directly by the dimensions of the clean beam (and so, ultimately, by those of the dirty beam).

Besides improving on the sidelobe characteristics of the principal solution, there is a further advantage to the use of CLEAN. Most of the gaps in the (u,v) coverage result from the vagaries of the detailed sampling available, and their importance is slightly lessened by the presence of observations at the same absolute separation, but different azimuthal orientation. There is a significant exception, however. At the very center of the (u,v) plane is a hole of some finite radius, which cannot be filled in. For the larger array configurations, this radius is set by the array scale, but when D observations are included, the size of the hole is reduced until a physical limit is reached. That limit is at least one antenna diameter, because of shadowing problems; the presence of cross-talk will increase the limit slightly beyond this. The point of greatest interest is that at the origin of the (u,v) plane. The visibility at this point is simply the Fourier transform of the total flux in the mapped field, but it cannot be measured. Clark (1982) has discussed the resulting difficulties in detail. Their root cause is that, in these circumstances, the dirty map will have an average value of zero, despite the fact that all real flux is positive. This is achieved by balancing the relatively small region of actual emission on the map by a large area of slightly negative "brightness." Normally, one sees this as a slight, negative "bowl" in which the source sits. Since CLEAN provides an estimate of the brightness at points in the (u,v) plane not actually sampled, it may be able to interpolate successfully across the central hole, if it is not too large. As Clark mentions, the positive mean value of the clean beam is instrumental in producing this effect. This is one of the best justifications for its use. CLEAN succeeded quite well in performing this task, in the case of NGC 3359. Of course,

these data do contain D array spacings, so that the case is a rather favorable one.

One further application of the CLEAN program is quite useful in 21 cm line work. At that wavelength, there are many unresolved continuum sources in a typical field which are bright enough to pose a problem of possible confusion with the neutral hydrogen emission. The most satisfactory way to eliminate such sources (van Gorkom, 1982) is to estimate their contribution in the (u,v) plane and to subtract them there, before inverting the visibility data to obtain dirty maps. This estimate can be provided by CLEANing a continuum map of the field, with the algorithm allowed to find flux only in a few very small regions centered on the discrete continuum sources. Six continuum sources were removed from the NGC 3359 field in this way. Their estimated positions and flux densities are listed in Table 2-2.

The mapping parameters used for the combined C and D visibilities for NGC 3359 were as follows. The data were gridded onto a 512 x 512 (u,v) plane, giving, of course, output maps 512 x 512 pixels in extent. The pixel size in the (x,y) plane was 6 arc seconds. Uniform weighting was selected, as the brightness sensitivity of the data was quite good. A Gaussian taper was applied, having a scale, from the origin to the 30% level, of 10,000 wavelengths. (This Gaussian function, therefore, has a dispersion of 6450 wavelengths or 1.36 km.) These values were chosen by experiment, using one channel map with a very strong signal and another with only weak H I, in order to find an acceptable compromise between sensitivity and resolution.

Since care was taken, in the choice of bandwidth parameters, to leave a baseline of observed channels with no expected signal, the

Table 2-2
CONTINUUM SOURCES IN THE NGC 3359 FIELD

Source	Right Ascension (1950.0)	Declination (1950.0)	Flux Density, mJy[a]
1	$10^h\ 44^m\ 07\overset{s}{.}19$	+63°33'38".5	5.62
2	10 45 08.88	+63 28 52.9	24.5
3	10 44 48.60	+63 22 08.8	103.3
4	10 42 43.65	+63 22 41.7	19.8
5	10 41 27.87	+63 11 06.2	25.3
6	10 42 27.90	+63 34 11.4	3.54

[a]These values have been corrected for the effects of the primary beam.

highest and lowest velocity channels in the data base were strictly continuum maps. For safety, the first three seemingly line-free channels, on either side of the band center, were assumed to be possible line maps. This left five maps on either side of the line emission which were not processed further. Although the remaining channel maps are often called "line maps," it should be remembered in this case that they still contain any continuum signal from NGC 3359 itself.

The dirty line maps were subjected to the CLEAN algorithm, with a loop gain of 0.2, and 2000 iterations were used. The clean beam was 18.″04 x 17.″64 in size (full width at half maximum or FWHM), with its major axis at a position angle of -53.°4 (astronomical convention). Thus, as expected for a source of such northerly declination, the beam is nearly round. More importantly, the linear resolution achieved is more than twice as fine as that reached for this galaxy by Gottesman (1982) using the Green Bank three-element interferometer. It is also somewhat better than the typical resolutions of about 25" previously achieved at the VLA for H I in barred spirals by Gottesman, et al. (1984) for NGC 3992 and NGC 4731, and by van der Hulst, et al. (1983) for NGC 1097 and NGC 1365. This is primarily because the H I in NGC 3359 is so bright that, in choosing the mapping parameters, there is little need to sacrifice resolution for greater sensitivity.

The mean rms noise of the dirty maps is 1.16 mJy per beam solid angle, with very little scatter from one map to another. That of the clean maps is 1.11 mJy per beam. This value can be converted to a brightness temperature via the following conversion for the clean maps, which is based on equation (2-26):

$$\Delta T_B = \frac{1222.2}{\nu^2 \theta_1 \theta_2} \Delta S \quad , \qquad (2\text{-}30)$$

or ΔT_B = 2.13 K for the present observations, where we have specified ΔS in mJy per beam, ν in GHz (1.4156 for these data), and θ_1, θ_2 are the full widths at half maximum, in arc seconds, on the major and minor axes of the elliptical, Gaussian clean beam. On a single map, the value for the noise level varied by 2-3%, depending on the particular background area measured.

Figures 2-4 (a) to 2-4 (u) are contour representations of the single-channel clean maps. The dirty map of the 1016 km/s channel is shown, for comparison, as Figure 2-5. As might be expected from their respective noise statistics, there is almost no difference between the dirty and clean maps. This is not too unusual when uniform weighting has been used, as the sidelobes generated are ordinarily less extensive than for natural weighting. Table 2-3 gives information on the signal and noise properties of the dirty and the clean maps for each channel.

We will, for the most part, defer our consideration of the astronomical significance of these results to the next chapter. There, we will discuss the integrated maps of density and velocity for NGC 3359, prepared from the single-channel maps. However, two points are worthy of mention before passing on to the next stage of our analysis.

First, the asymmetry of the spiral arms noted in Chapter One is reflected in Figures 2-4. The channel maps at velocities slightly lower than the band center velocity of 1016 km/s exhibit more spatially

Figure 2-4. Single-channel clean maps of the neutral hydrogen in NGC 3359 made with a velocity separation of 20.7 km/s and an effective velocity resolution of 25.2 km/s. Coordinates of epoch 1950.0 are shown. The contour levels are multiples of the approximate rms noise of 1.15 mJy/beam: the first contour is at three times this value, the next at five times, and succeeding contours at increments of 2.5 times the noise level. The mean heliocentric velocity of each map is given in its upper left-hand corner. The synthesized beam (full width, half maximum) is shown to scale in the lower left of Figure 2-4 (a).

Figure 2-4 consists of Parts (a) through (u).

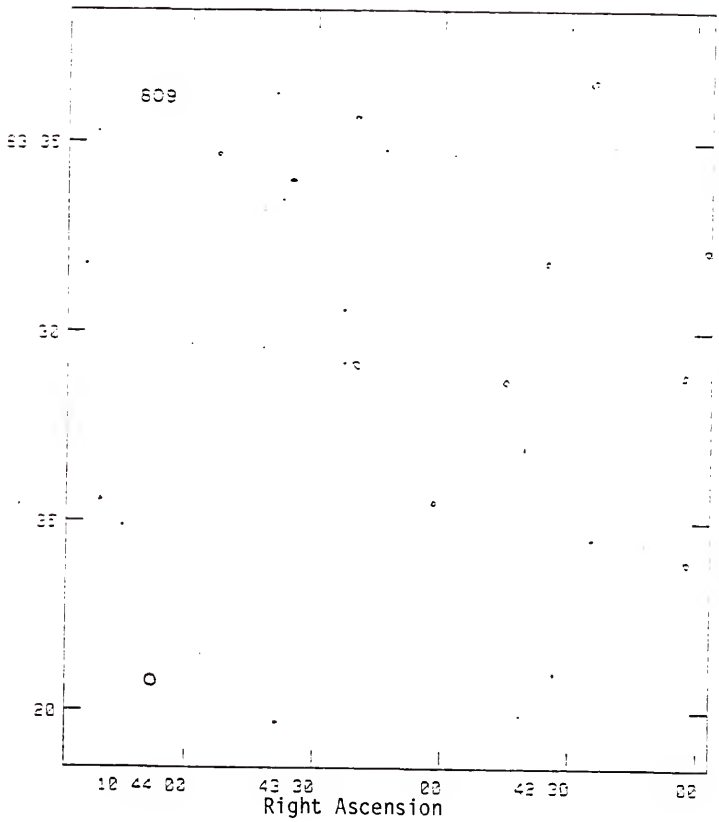


Figure 2-4 (Part a)

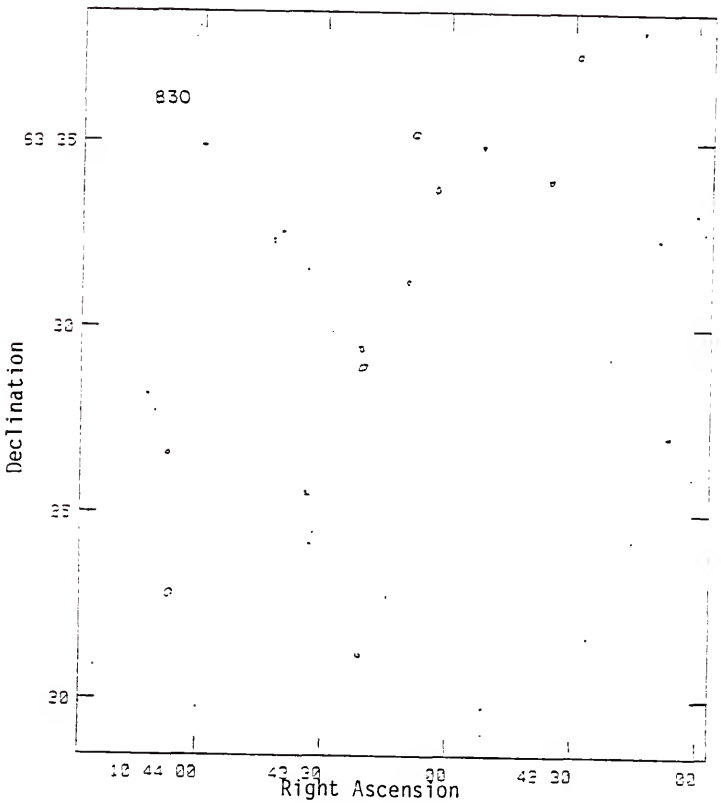


Figure 2-4 Continued (Part b)

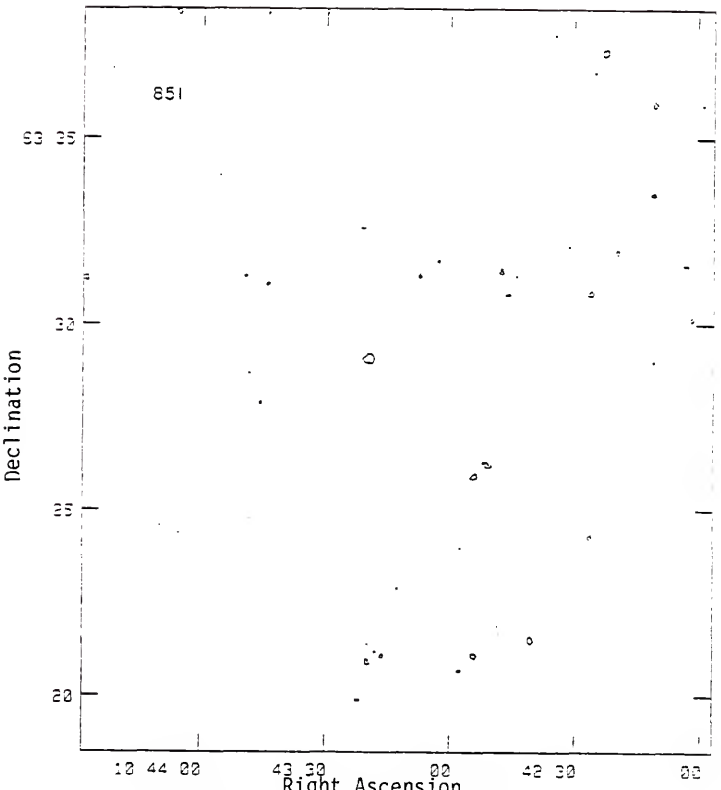


Figure 2-4 Continued (Part c)

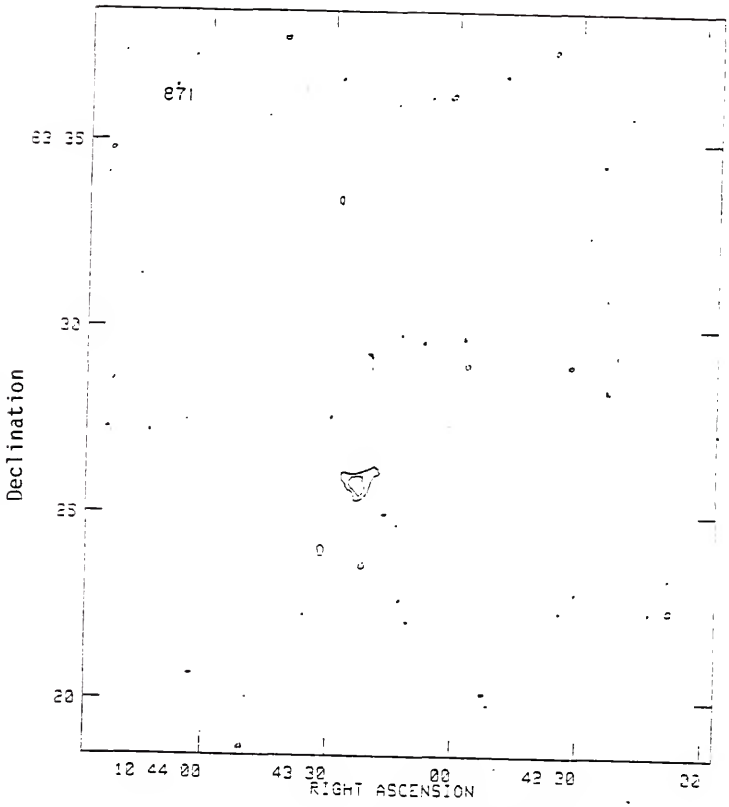


Figure 2-4 Continued (Part d)

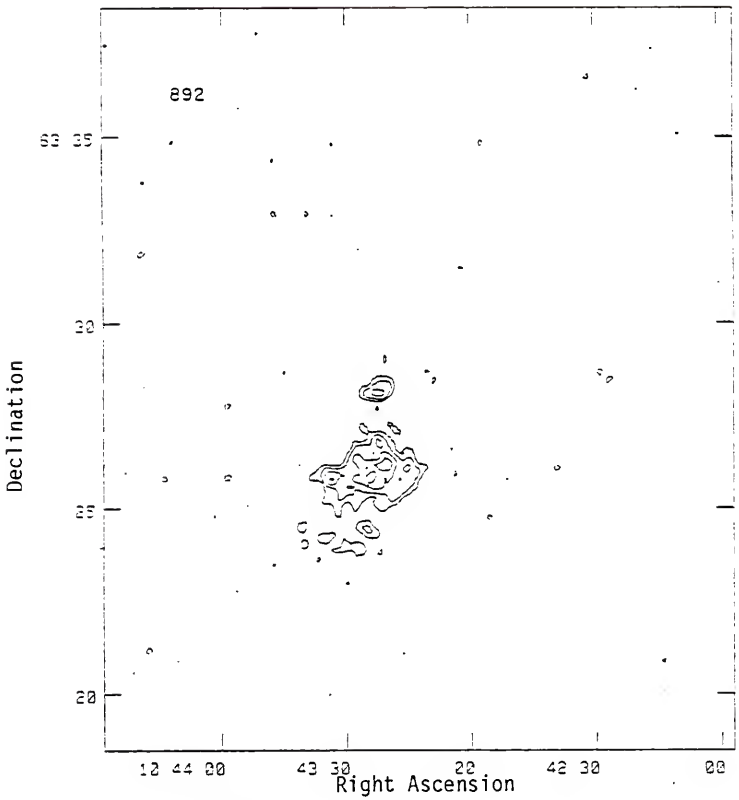


Figure 2-4 Continued (Part e)

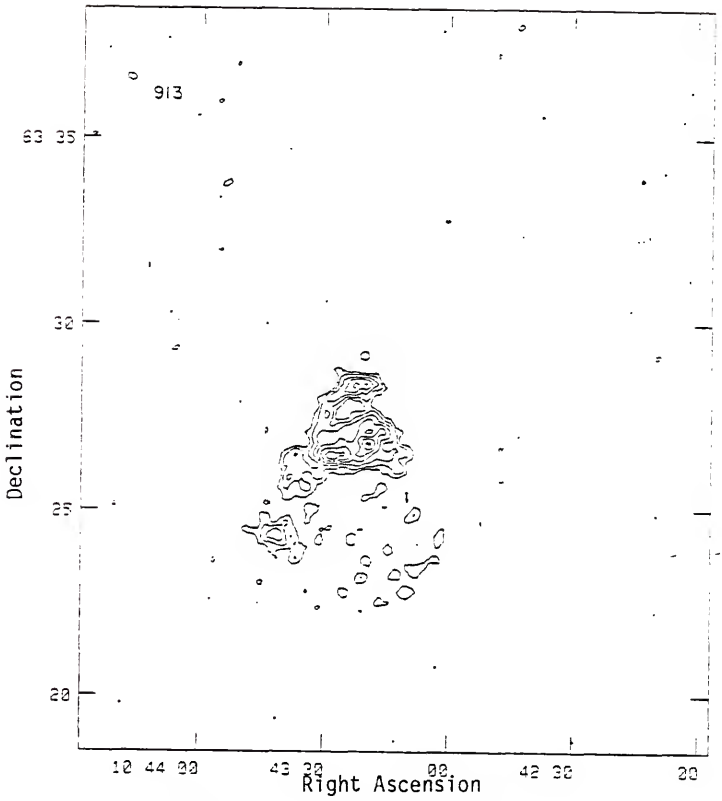


Figure 2-4 Continued (Part f)

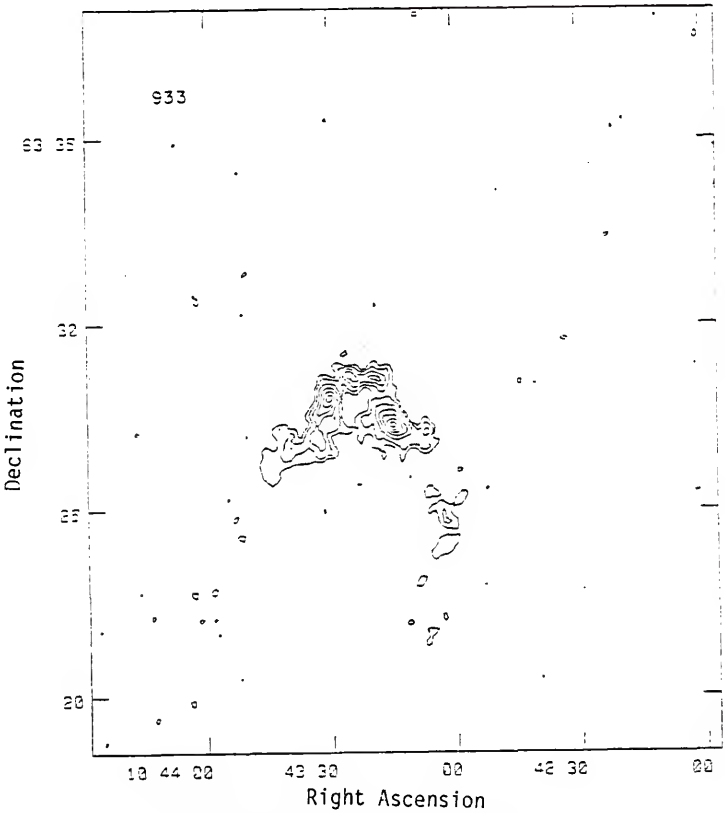


Figure 2-4 Continued (Part g)

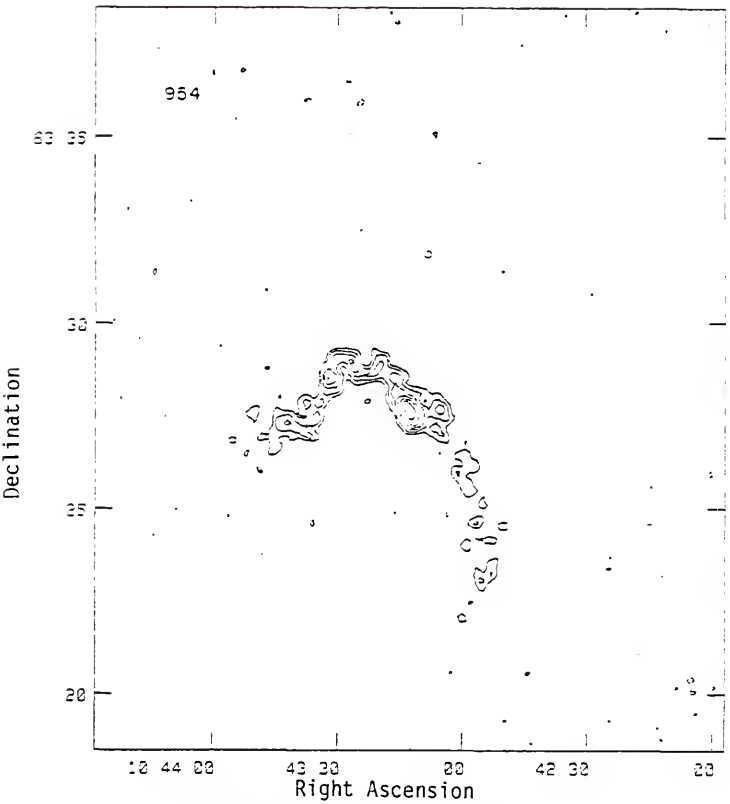


Figure 2-4 Continued (Part h)

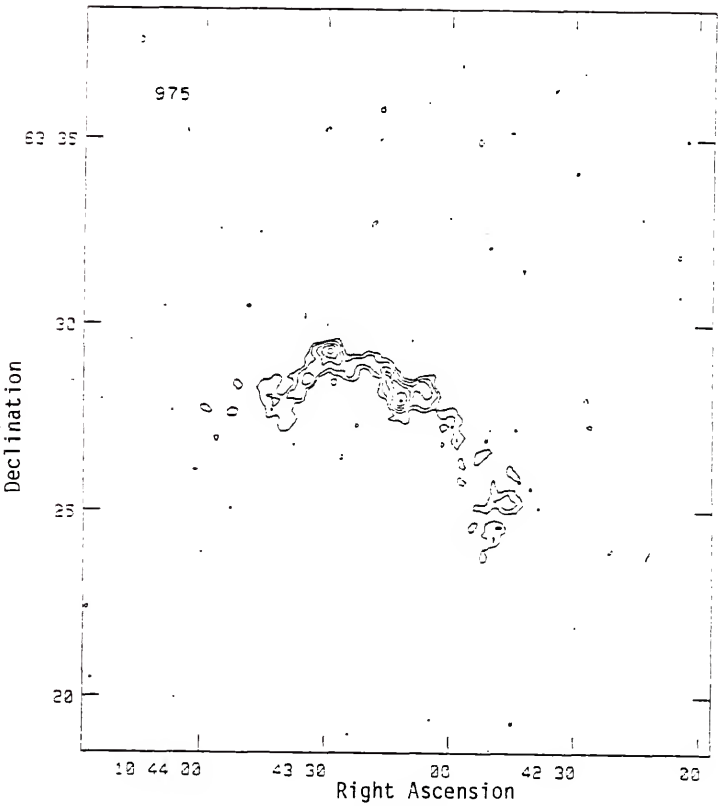


Figure 2-4 Continued (Part i)

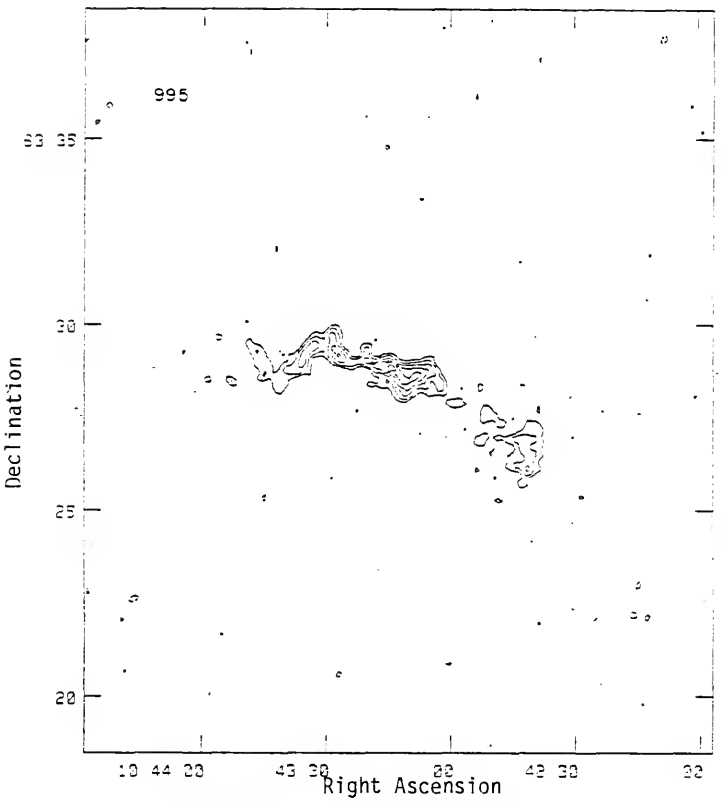


Figure 2-4 Continued (Part j)

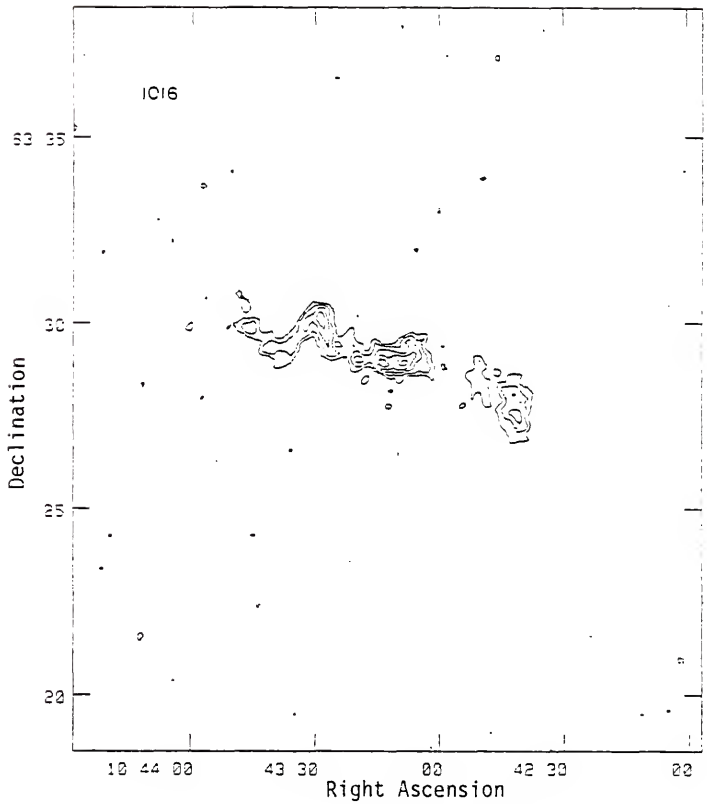


Figure 2-4 Continued (Part k)

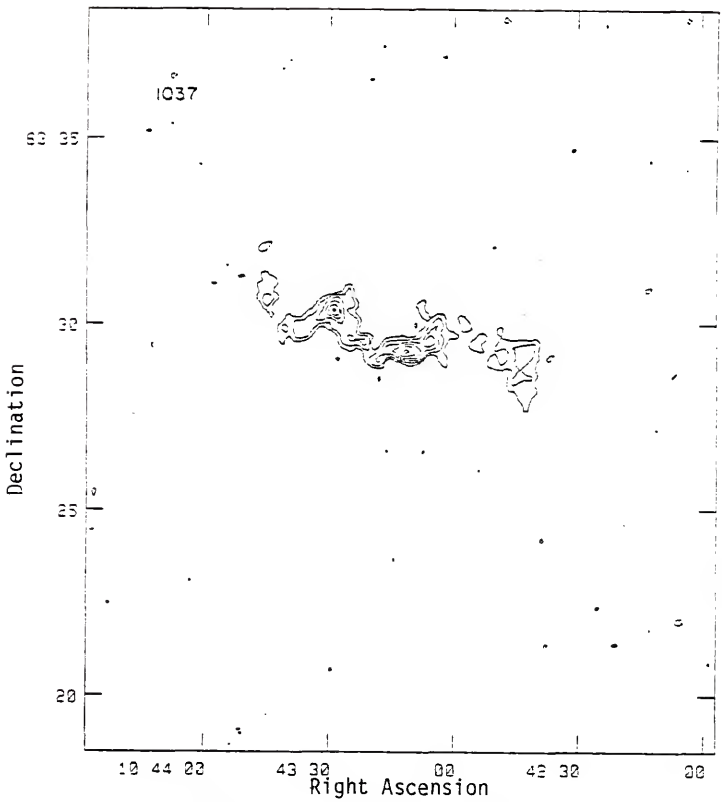


Figure 2-4 Continued (Part 1)

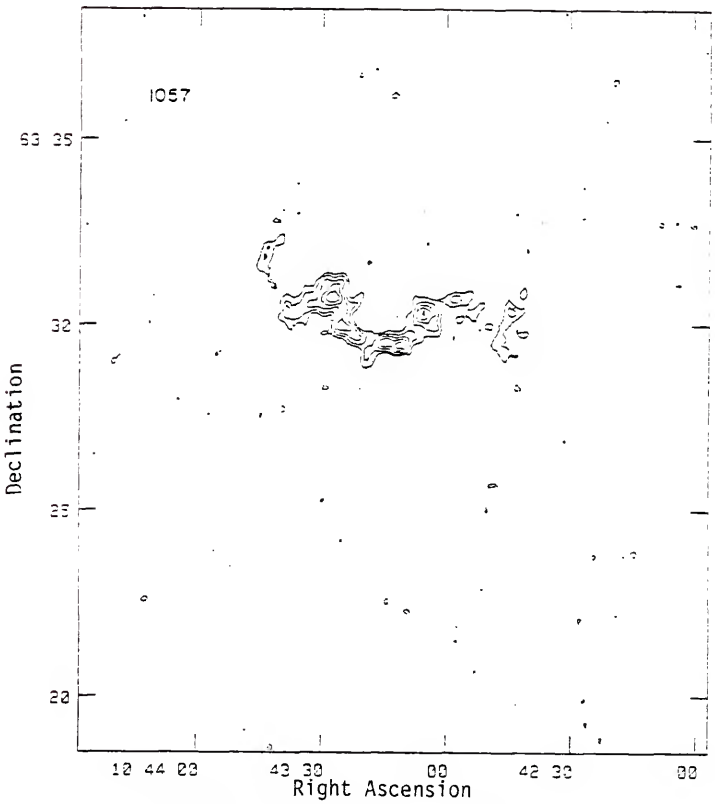


Figure 2-4 Continued (Part m)

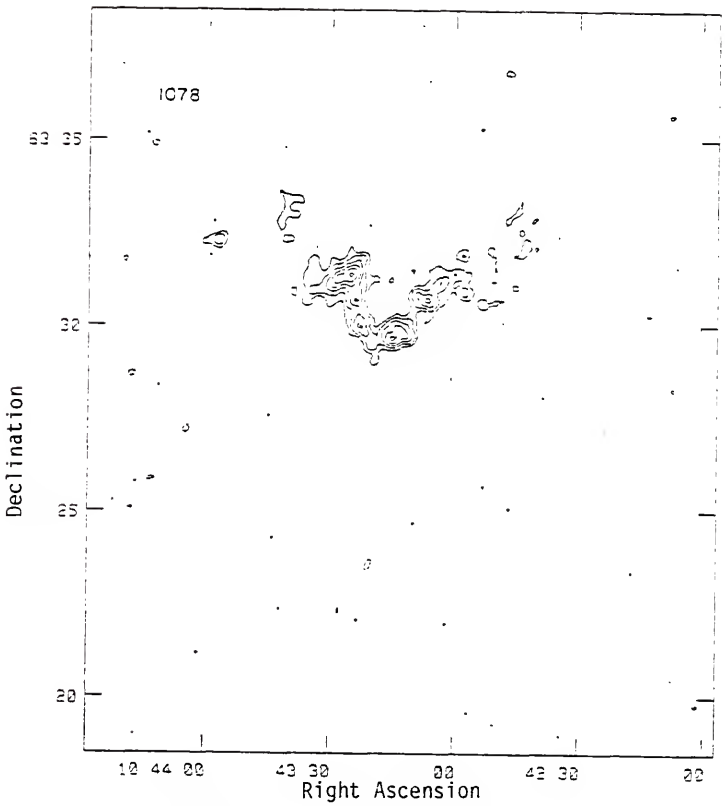


Figure 2-4 Continued (Part n)

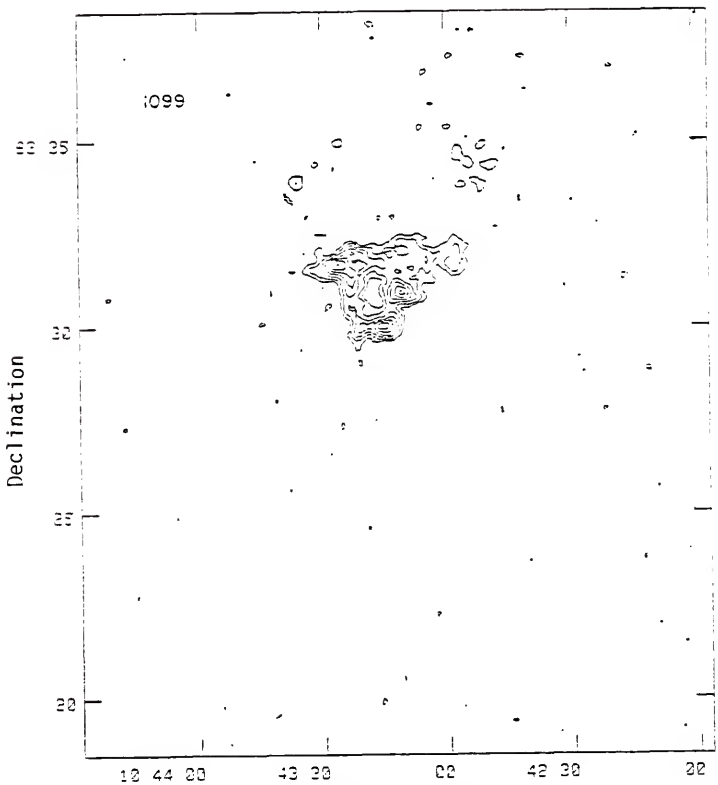


Figure 2-4 Continued (Part o)

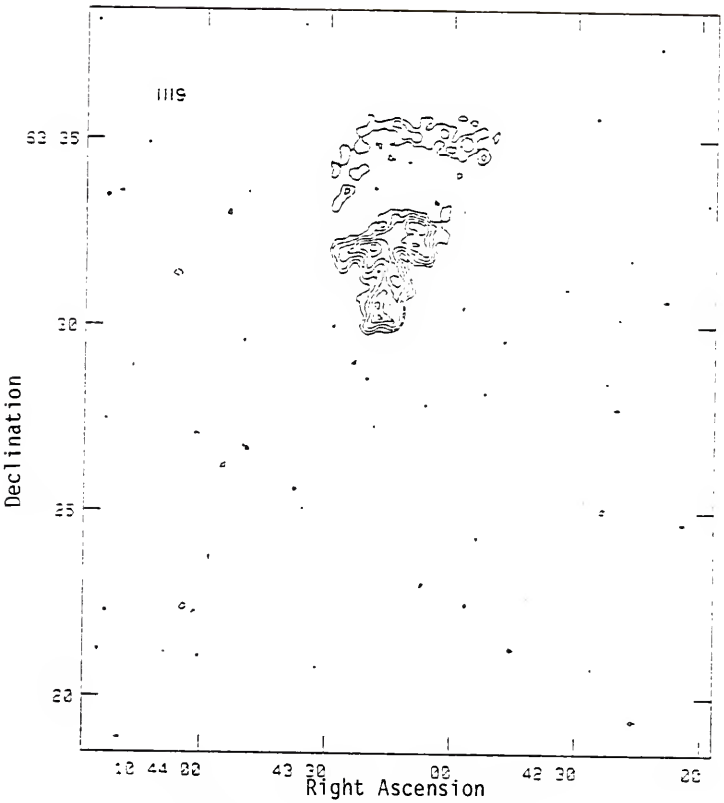


Figure 2-4 Continued (Part p)

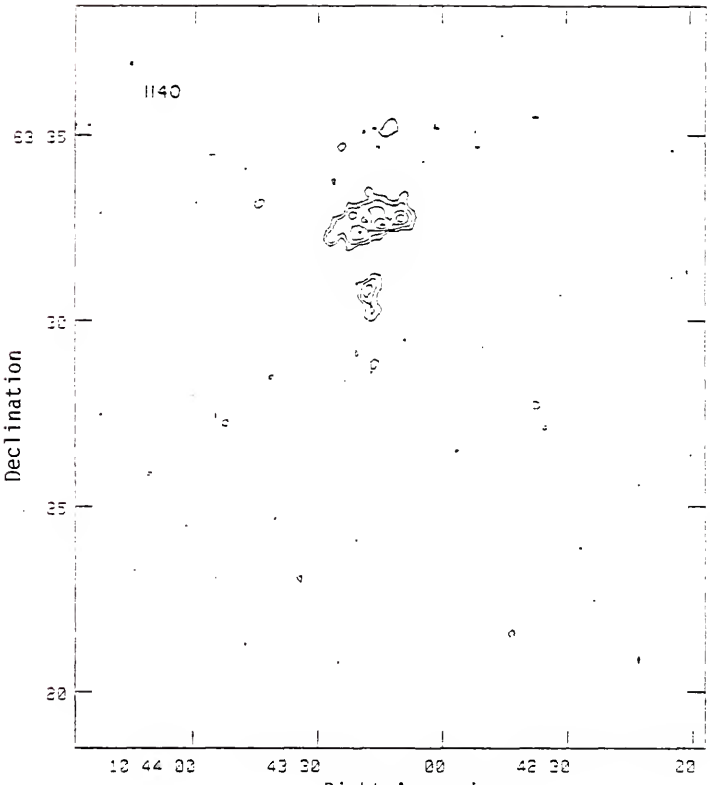


Figure 2-4 Continued (Part q)

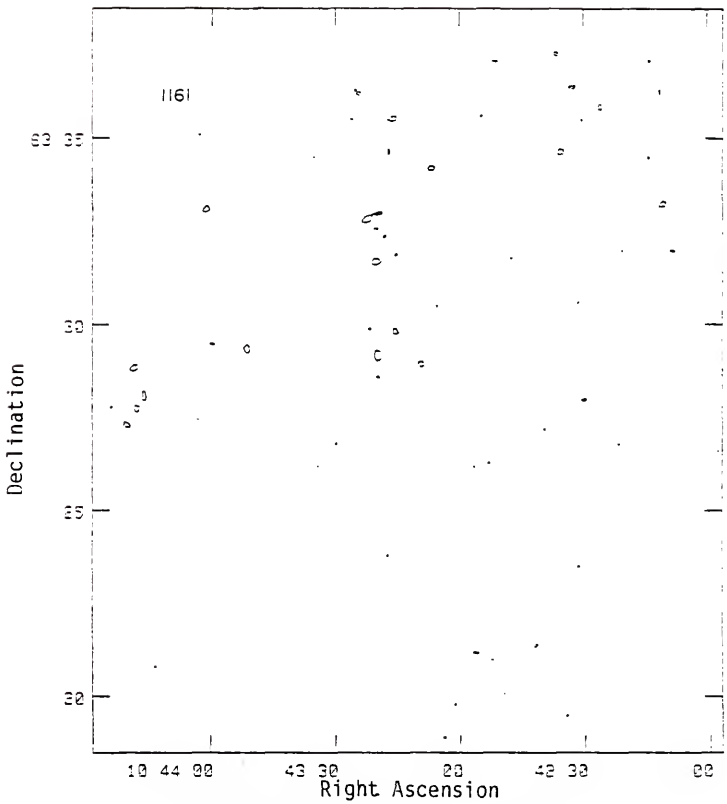


Figure 2-4 Continued (Part r)

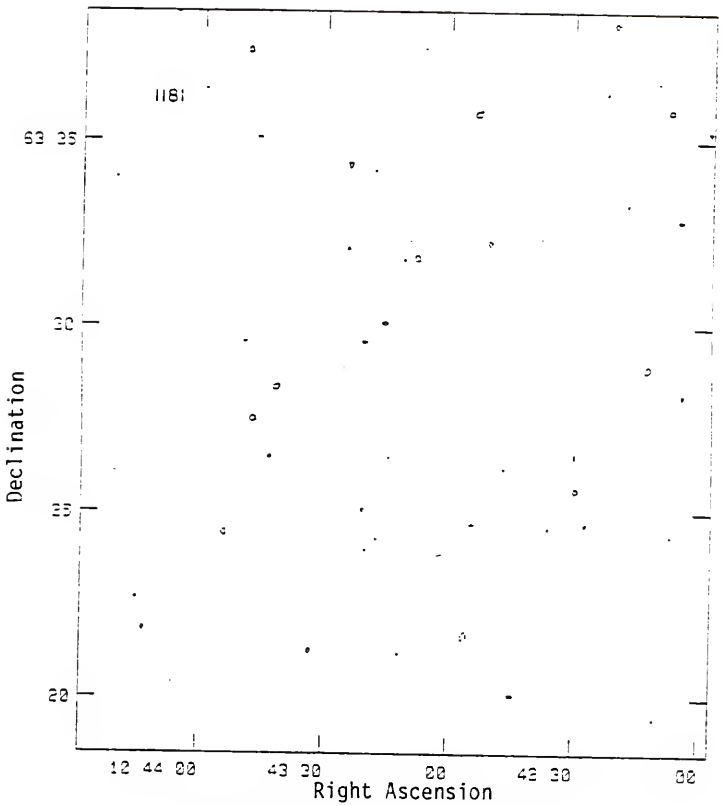


Figure 2-4 Continued (Part s)

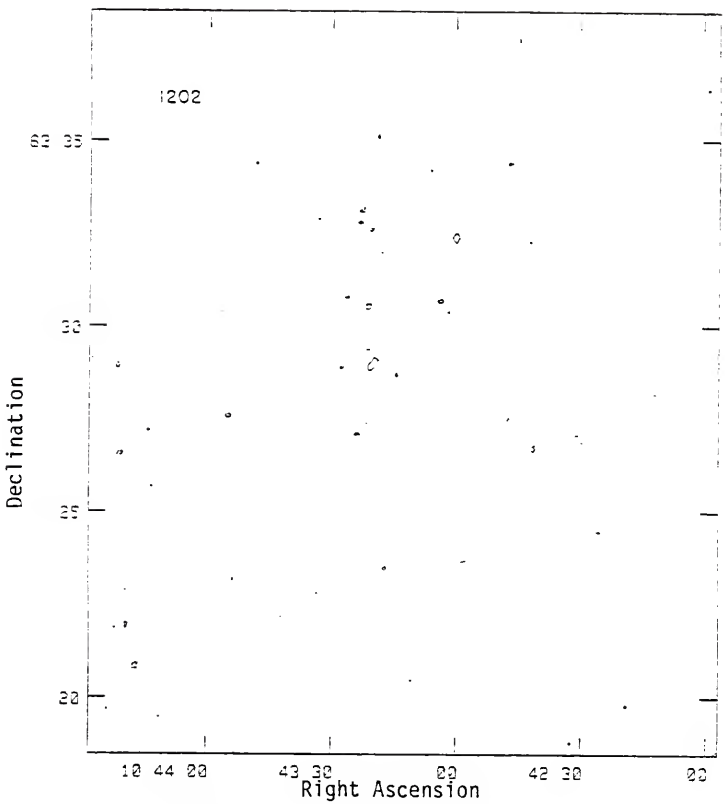


Figure 2-4 Continued (Part t)

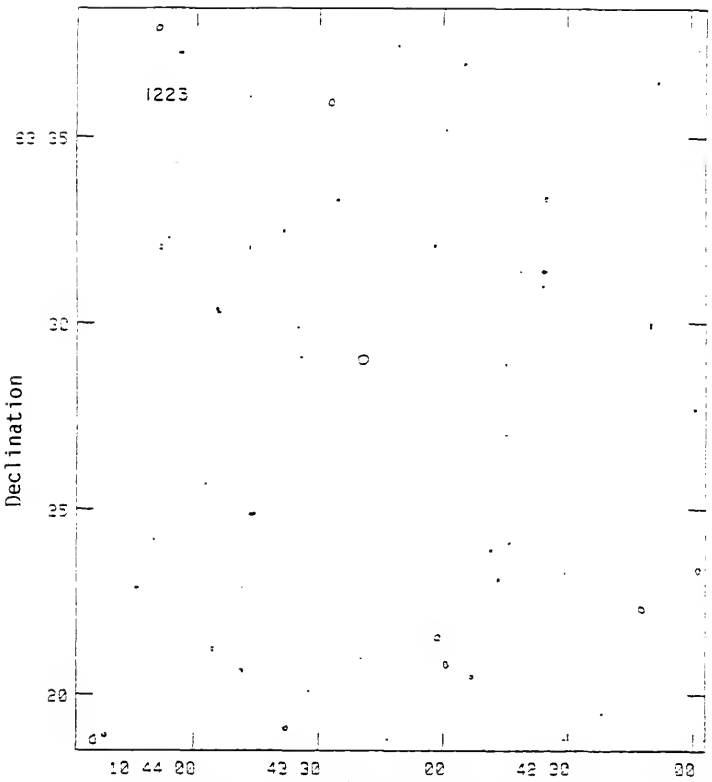


Figure 2-4 Continued (Part u)

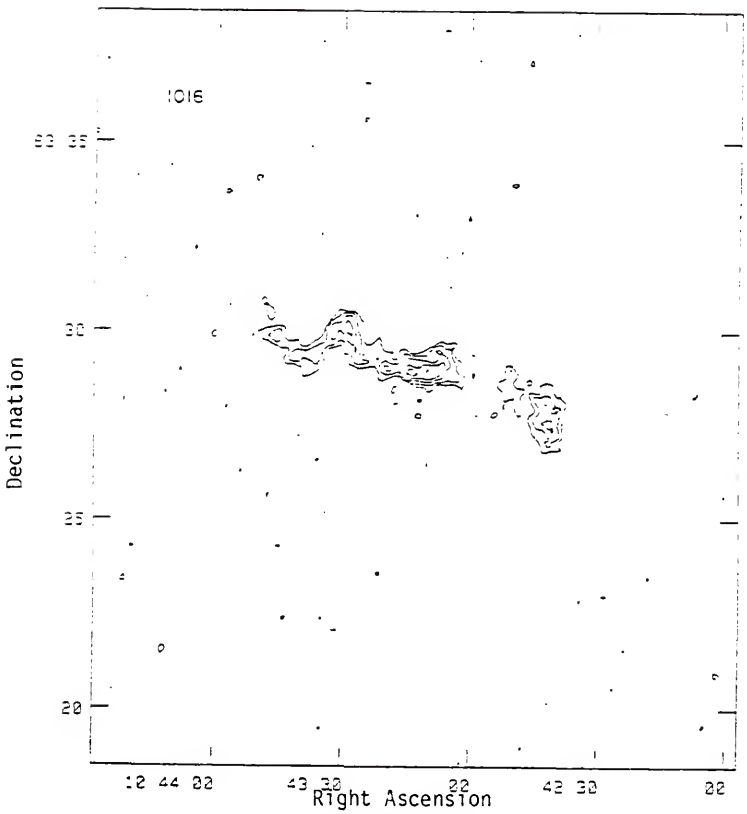


Figure 2-5. The single-channel dirty map of the H I in NGC 3359, at the center velocity of 1016 km/s. This contour map should be compared to the clean map of Figure 2-4 (k). The map is labeled with epoch 1950.0 coordinates. The contour levels are the same as those used in Figure 2-4.

extended emission than their counterparts at higher velocities. These former maps are those which contain the signal from the southern half of the galaxy. While the neutral hydrogen associated with the dominant arm is also seen in these maps, the outlying emission is well beyond the radius of that arm. As will be discussed in the next chapter, in the integrated density map this emission seems to form ragged, outer arms or a partial ring. Here we simply point out that the hydrogen at this large radius is asymmetric in the same sense as the optical arms.

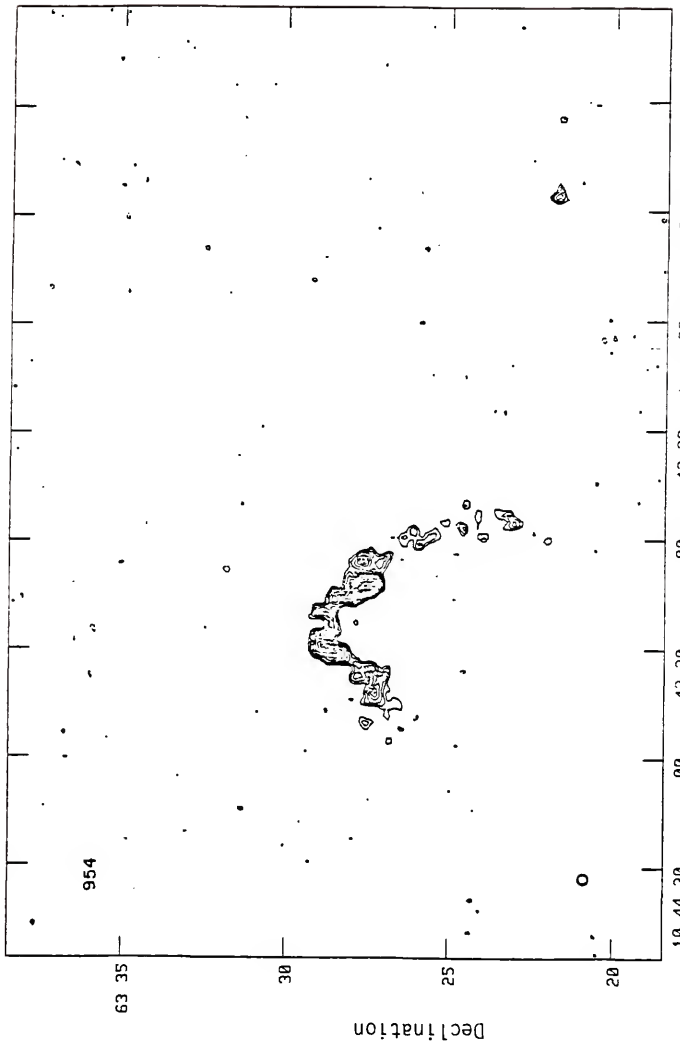
Second, one major surprise emerged when the C and D observations were concatenated to give a single, high-sensitivity data set for mapping. At a distance of some fifteen minutes of arc from the center of NGC 3359, and at least seven minutes from any other detected H I, a small, isolated 21 cm feature was found in two adjacent spectral channels. This object seems to be a previously unknown, low-mass satellite galaxy to NGC 3359. Support for this interpretation will be given in the next chapter. It is not listed in the Master List of Nonstellar Optical Astronomical Objects (Dixon and Sonneborn, 1980).

The object is clearly visible in Figures 2-6 (a) and (b), which are simply the relevant clean channel maps with a greater area of sky shown than in Figures 2-4. (An area centered on this feature and about 12' on a side was searched for emission in the CLEAN program, along with the inner quarter of the field centered on NGC 3359.) The emission from the main body of NGC 3359 has been "overexposed" by the choice of contour level in Figures 2-6, so that the weaker signal from the isolated cloud can be seen. Finally, we include as Figures 2-6 (c) and (d) the adjacent map on either side of this pair, at the same scale. The satellite H I clearly decreases drastically in brightness at these latter velocities. The velocity width of its spectrum is evidently quite small.

Figure 2-6. Single-channel clean maps of neutral hydrogen in NGC 3359, showing a larger area than in Figure 2-4. Each map is labeled with its heliocentric velocity in km/s. Contour levels are identical to those of Figures 2-4 and 2-5. Coordinates are for 1950.0.

(a) and (b). Two channel maps containing emission from the newly discovered satellite galaxy of NGC 3359.

(c) and (d). The two channels adjacent to these; no emission from the satellite is detectable.


954
63 35
30
25
20
Declination
10 44 30
00
43 30
00
42 30
00
41 30
00
Right Ascension

Part (a)

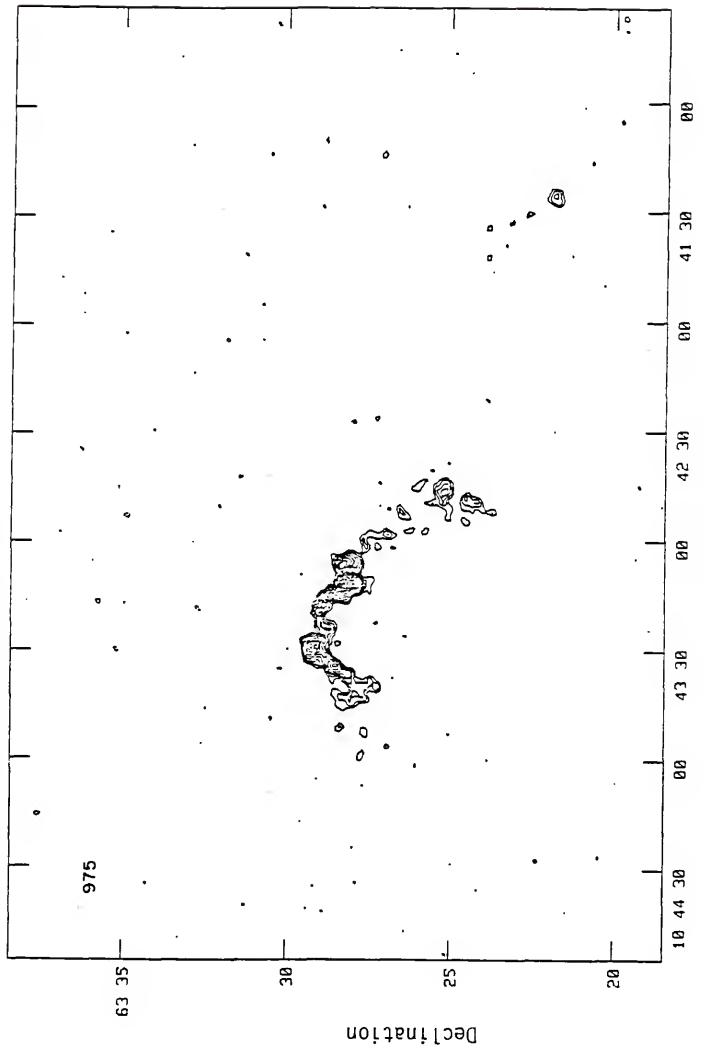


Figure 2-6 Continued. Part (b).

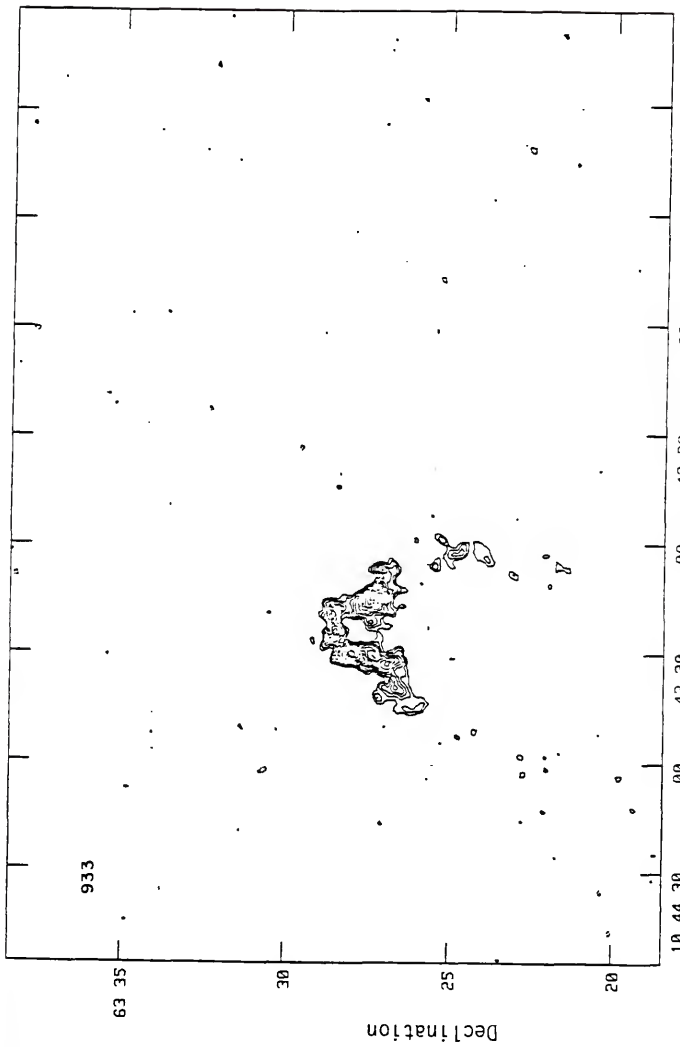


Figure 2-6 Continued. Part (c).

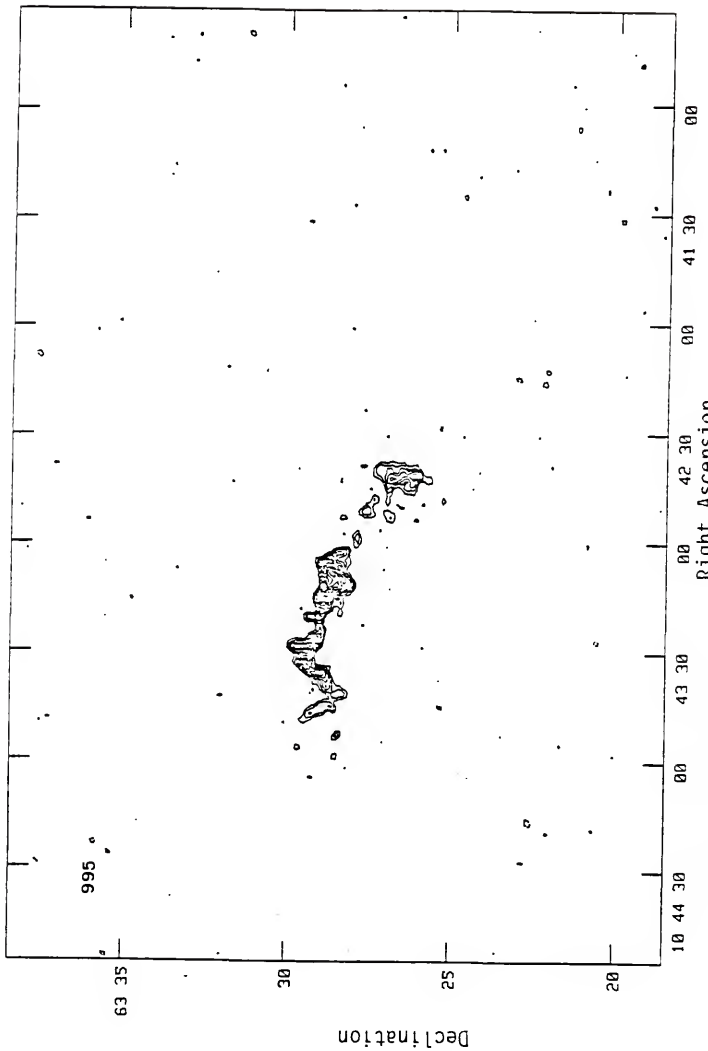


Figure 2-6 Continued. Part (d).

TABLE 2-3
SINGLE CHANNEL SIGNAL AND NOISE CHARACTERISTICS

Channel Number[a]	Center Velocity (km/s)	RMS Noise of Dirty Map (mJy/beam)	Peak Brightness of Dirty Map (mJy/beam)	RMS Noise of Clean Map (mJy/beam)	Peak Brightness of Clean Map (mJy/beam)
6	1222.8	1.15	5.29	1.11	5.35
7	1202.1	1.15	5.79	1.11	5.50
8	1181.4	1.14	5.12	1.09	5.12
9	1160.8	1.17	5.51	1.14	5.42
10	1140.1	1.16	14.38	1.11	15.05
11	1119.4	1.18	18.31	1.13	18.83
12	1098.7	1.19	19.93	1.13	20.22
13	1078.0	1.13	17.77	1.10	17.80
14	1057.4	1.15	16.84	1.12	16.98
15	1036.7	1.15	18.30	1.11	18.28
16	1016.0	1.17	16.31	1.13	16.65
17	995.3	1.14	18.83	1.10	19.29

18	974.6	1.17	17.41	1.13	17.87
19	954.0	1.15	18.94	1.09	19.82
20	933.3	1.18	21.69	1.11	21.92
21	912.6	1.16	20.42	1.13	20.80
22	891.9	1.15	13.58	1.11	14.15
23	871.2	1.17	8.28	1.12	8.42
24	850.6	1.16	5.50	1.10	5.42
25	829.9	1.16	5.66	1.12	5.45
26	809.2	1.16	5.74	1.07	5.55

[a]Ordered by increasing frequency. Channels 1-5 and 27-31 are continuum channels and are therefore omitted.

CHAPTER III
DISTRIBUTION AND KINEMATICS OF NEUTRAL HYDROGEN

Techniques of Spectrum Integration

In the last chapter, we discussed the acquisition, calibration, and processing of 21 cm visibility data obtained with the VLA. As the end product of this data reduction, we obtained a set of single-channel, clean maps, with a channel separation of 20.7 km/s. The distribution and kinematics of the neutral hydrogen in NGC 3359 can be determined from the ensemble of these maps, using equations (2-1) and (2-2). These equations show that the mean column density and velocity at a given point in the galaxy are the zeroth and first moments, respectively, of the brightness temperature with respect to velocity. One can picture the single-channel maps as being stacked, one atop another, in order from lowest frequency to highest. This mental construct is often referred to as the data "cube" in discussions of spectral line data (van Gorkom, 1982; Ekers and van Gorkom, 1983), although the third dimension--that identified with the frequency or channel number--is usually much smaller in extent than the other two, which represent the number of pixels in right ascension and declination, respectively. The mean density and velocity are then found by summation along the "vertical" dimension, at each (x,y) position. In the column density integration, all the temperatures have equal weight; for the mean velocity computation, one weights by the velocity at each channel. In the absence of observational noise, this procedure would be straightforward. In practice, however, one must use a more sophisticated

algorithm than a simple integration over the full extent of each spectrum, owing to statistical fluctuations in the noise over the limited number of spectral channels.

The difficulty in high spatial resolution 21 cm work is that, typically, the signal-to-noise ratio is rather low, often well below five for extended regions. Under these circumstances, there will be many points at which the relative contribution of noise to the integrals of equations (2-1) and (2-2) is not negligible. Some investigators (e.g., Weliachew, Sancisi, and Guelin, 1978; Sancisi and Allen, 1979; Gottesman, 1982) have decided on appropriate limits of integration by visual inspection of every spectrum, but this is not practical with the present observations of NGC 3359, where computer programs were used to search approximately 40,000 spectra for H I emission. One method for rejecting noise in an automated integration procedure is to impose a "cutoff" temperature, and to ignore all points lying below this level when doing the integration. This technique was used in early 21 cm mapping by Rogstad and Shostak (1971), Weliachew and Gottesman (1973), Rots and Shane (1975), Winter (1975), and others. However, the method is prone to systematic errors in its determination of the mean velocity field. Consider the typical case (such as the present observations of NGC 3359) where, as we have said, the signal-to-noise ratio is low over much of the object, and where the product of the number of velocity channels and the number of pixels surveyed is large (of order 10^6 for NGC 3359). In such a case, one must expect that there will be many pixels at which the largest positive noise spikes are of comparable strength with the peak of the signal. One cannot exclude these spikes from the integration without excluding, as well, much of the signal from the low surface brightness portions of the

galaxy. If the noise spikes are included, the consequences are quite serious, especially for the map of mean velocity. The resulting error is systematic, because the expectation of the "velocity" of these noise spikes will be the center of the passband (usually coinciding with the systemic velocity), regardless of the spatial location of the pixels in which they are found. Therefore, the entire velocity map will be biased toward the central velocity of the system.

The problem of noise contamination is a familiar one in 21 cm work. An illuminating account of its consequences, and a suggested cure, are given by Bosma (1978, 1981a). We will rely primarily on the method he introduces. This method, commonly called the "window" integration scheme, tries to discriminate against the inclusion of noise spikes in an automated integration routine, by providing the routine with some a priori information concerning how one expects a noise-free spectrum to look. Specifically, one requires that the real emission be confined to a single range of contiguous channels. If there is an isolated local maximum at a far different frequency from the rest of the signal in a spectrum, it is likely to be a noise spike. This criterion reflects the physical argument that there should not be wildly disparate velocities in the gas in a single small region of the galaxy; even allowing for the finite spatial resolution introduced by the telescope beam, a spectrum which shows maxima separated by a span of several signal-free channels is implausible. An exception to this statement must be made for galaxies that are seriously enough warped that some lines of sight pass through the warped disk twice, as in M 31 (Emerson, 1976; Unwin, 1980a, b; Brinks and Shane, 1984) and M 33 (Rogstad, Wright, and Lockhart, 1976). No systematic pattern of double-peaked spectra, indicative of such a warp, is apparent in NGC 3359.

The window technique, therefore, defines an acceptance gate, or "window," in velocity, about the peak of the spectrum. An iterative process is used to determine the velocity width of this window, which is initially very narrow. At each step, the "continuum" is defined as the mean level of points outside the window. An iteration consists of tentatively expanding the window to include the two channels currently on either side of it, and comparing the new value of the continuum to the previous value. When these two numbers agree to within a specified tolerance, the integration is considered to have detected all of the real signal, and the procedure is stopped for that pixel. The tolerance depends on the rms noise of the single-channel maps, and on the number of points remaining in the empirically defined continuum. There are also two signal-to-noise criteria imposed directly on the data. If a spectrum contains no channels whose temperature is at least twice the rms noise of the single-channel maps, the spectrum is not integrated. Points which pass this test must then give a value for the column density integral of at least three times the expected rms uncertainty in that quantity, or be discarded.

This method gives maps of column density and mean velocity which are free of the systematic biases of the cutoff technique. It has been used on 21 cm observations of the external galaxies, for instance, NGC 4151 (Bosma, Ekers, and Lequeux, 1977); NGC 2841, 3198, 5033, 5055, and 7331 (Bosma, 1981a); and NGC 3992 and 4731 (Gottesman et al., 1984). However, while the method is sounder than the cutoff scheme, it fails to use all the available information in the data cube, and, hence, to produce the best possible estimate of the desired moments at each point in the galaxy. Specifically, no reference is made to spatially adjacent

points, which, just as we have done with the adjacent spectral channels, can be analyzed to help determine whether a given signal is real. As we shall see, the most serious effect of this omission is that, in areas of low surface brightness, the best possible signal-to-noise ratio is not achieved. In the present instance, we have been able to measure the rotation curve of NGC 3359 reliably to a considerably greater radius than otherwise possible, by supplementing the automated window routine with a procedure based on visual inspection of spectra. The procedure is outlined below.

The key to this method lies in the observation that a random noise spike has a characteristic spatial signature. It is uncorrelated with the temperatures measured in the same channel at adjacent pixels, although in practice it will not be confined entirely to a single pixel, because of the finite size of the clean beam which is used to generate all of the single-channel maps. No feature can be smaller than this Gaussian beam, and so the distinguishing feature of noise spikes is that they decay from their peak values at precisely the same rate as the clean beam. Such structures are quite easy to identify, if they occur far enough in velocity from any real emission so that there is no confusion. Noise spikes that fall in channels where there is real signal, of course, cannot be identified. However, the H I lines are typically only a few channels wide, and so the fraction of the passband over which one can correctly isolate large noise spikes as such is fairly high. In the present observations, we estimate it to be somewhat above one-half. The signal-to-noise ratio can thus be improved by the rejection of spurious signals. The reverse situation is at least equally important: one can identify regions of broad, low brightness emission, by searching for

areas where the spectral features are correlated over scales much larger than the beamwidth. The combination of these two effects makes it possible to extend accurately determined densities and velocities to a significantly larger portion of the object. A search for these effects in the data could be automated, but in practice would require a great deal of computer time. Instead, the scheme detailed below was adopted.

First, the single channel maps of NGC 3359 were split into four quadrants, to reduce the number of integrations necessary in each step. Each quadrant was then integrated using the window method, with the signal-to-noise criteria mentioned above. The spectra were also Hanning smoothed in frequency, to improve further the signal-to-noise properties of the integrated maps. For comparison, the same integrations were computed by the cutoff technique. The cutoff adopted was twice the rms single channel noise, and the same "three sigma" requirement on the integral was enforced as in the window method. The density and velocity maps generated by the former technique were subtracted from those produced by the latter, to give density and velocity difference maps in the four quadrants. These maps were among the primary tools used to find regions of the galaxy where errors were likely. Another very useful tool was to plot the point-by-point velocities, determined by the velocity-window program, after subtracting at each point the mean rotational velocity (properly deprojected) of the annulus in which it lies. In such a plot, either a large absolute value, or a large difference in value between neighboring points, can indicate a spurious velocity. This same procedure was not useful when applied to density maps; in that case, the most prevalent cause of large gradients was crossing the edge of the real emission in the galaxy, as defined by the window routine.

After these programs had identified a given region as meriting closer scrutiny, a Gould plot was produced, showing the individual spectra over the entire region. It was found that plotting each spectrum in a field one inch square gave an acceptable compromise between alleviation of eye fatigue and the need to see large areas at once, in order to spot large-scale features. Particular pixels with suspected errors were compared to their neighbors, and a decision was made to accept or to change the current values of the density, velocity, and continuum. These values were changed only if the presence of one or two noise spikes was clearly demonstrable. The changes made were of two kinds; either the entire spectrum was rejected as noise, and the density and velocity set to zero, or a new window was defined around the channels believed to contain real emission, and the integration repeated.

In addition, virtually the whole of the periphery of NGC 3359 was searched to see whether there might be any weak, extended emission, as described previously. A good deal of H I was detected in this way. Although this weak, outer emission clearly had been suggested by the automated window program, until the various corrections resulting from the visual inspections of the spectra had been made, the ratio of the number of points with accurate velocities to the number of spurious points was rather low. Consequently, the rotation curve could not be determined accurately. The combination of the rejection of spurious data, and the inclusion of weak but real emission, thus gave a very substantial improvement in the accuracy of the last few points on the rotation curve. The radius of the last well-determined point on the rotation curve increased from just over four arc minutes to more than six. As an indication of the effectiveness of the various procedures used,

the automated window routine resulted in just over 3000 spectra passing both signal-to-noise criteria. An estimated 5000 spectra were subjected to at least a cursory visual examination, and nearly 1000 sets of values were changed, of which the overwhelming majority were pixels previously defined not to have signal. Thus, the extension of the window technique by inspection of the individual spectra proved to contribute a significant number of points to the H I density and velocity maps, especially at large radii.

An alternative approach to the procedure described here, suggested by van Gorkom (1982), would be to convolve the single-channel maps to a lower spatial resolution, and use the lower resolution data to determine the extent of the window for each spectrum. One would then apply these windows in performing the integrations on the original, high resolution data. This method, like the one we have used, would be effective in enhancing the sensitivity to broad, low-level emission. However, because of the computer time required, primarily for the convolution of each of the 21 rather large single-channel maps, it was not considered in the present investigation.

Before proceeding to discuss the results of the corrected integrations, it may be of interest to mention briefly the most common effects found in the spectra. By far the greatest number of points added by the visual-inspection procedure contained very narrow line emission, generally only one or two channels wide. The window program is only able to find the signal, in this case, if there are no noise spikes of even moderate strength. In fact, it is a fairly accurate generalization to state that the automated procedures are defeated by spectra which are complex in any way. The human "data processor" is able to recognize, and to

correct for, a slightly greater degree of complexity, although these generalizations are hard to quantify. When the spectrum became so complex that it was no longer obvious where to place the limits of the window, the results of the window program (usually zero, indicating that the program had detected no signal) were retained, even if it seemed probable that there was some real signal present. This situation arose in $\leq 5\%$ of the pixels which were studied closely. The program also can be confounded by particular distributions of rather small noise spikes. The situations where there happen to be small positive spikes on either extreme end of the spectrum, or several moderately negative spikes grouped together anywhere, give particular trouble.

Neutral Hydrogen Distribution in NGC 3359

Figure 3-1 is the map of the integrated neutral hydrogen column density in NGC 3359, derived by the procedure detailed in the last section. The field shown includes only the main body of emission, omitting the area around the small satellite galaxy, which will be discussed later in this Chapter. The densities plotted in Figure 3-1 are those observed, with no correction applied for the effects of convolution with the synthesized beam, nor for the directivity of the individual antennas. These densities are derived under the assumption of low optical depth. The maximum density observed is 3.51×10^{21} cm^{-2}. The rms noise in this map is approximately 1.3×10^{20} cm^{-2}, assuming that a typical spectrum contains about five channels with signal. The contour levels of Figure 3-1 are at intervals of 10% of the peak density. As the density distribution is somewhat chaotic in places, with many local maxima and minima, it is rather difficult to interpret Figure 3-1 by itself. A grey tone display of the data is, therefore, given as well, in Figure 3-2.

Figure 3-1. Contours of neutral hydrogen surface density in NGC 3359. The contour interval is 10% of the peak density, which is 3.51×10^{21} cm^{-2}. North is at the top in this plot, east to the left. The positions of the foreground stars used as fiducial marks throughout this dissertation (see Figure 1-1) are indicated by crosses, as is the position of the galaxy center. The latter is also indicated by marks along the outer border. The tick marks around this border are spaced 0!1 apart. The entire field shown here is 11!9 x 16!4. The 18" synthesized beam is shown to scale in the lower right of the figure.

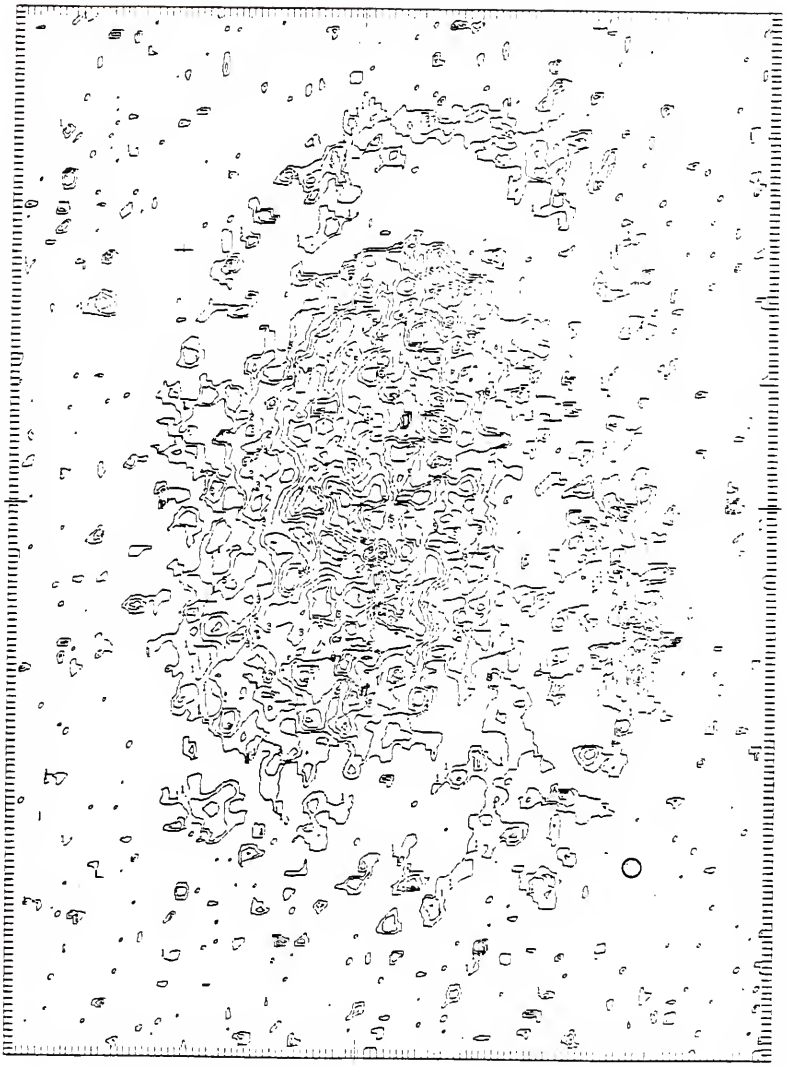
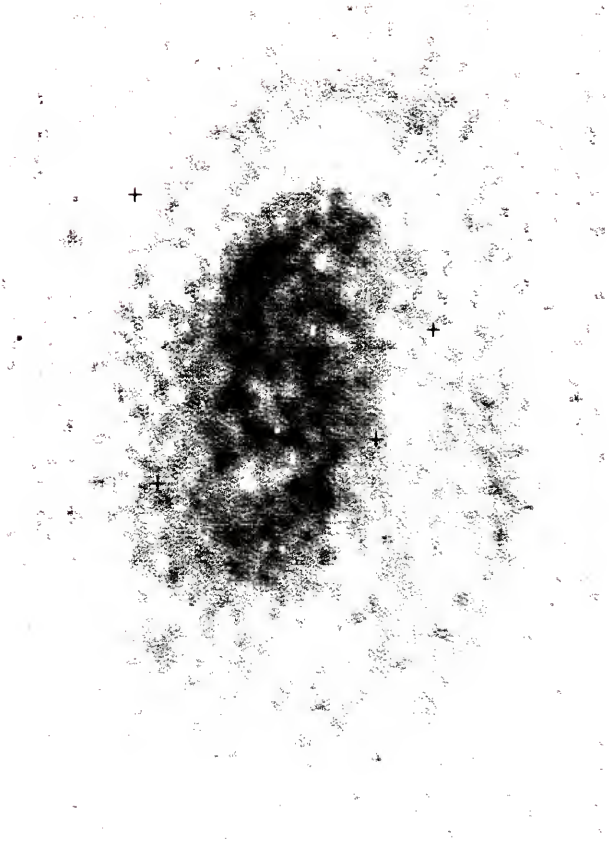


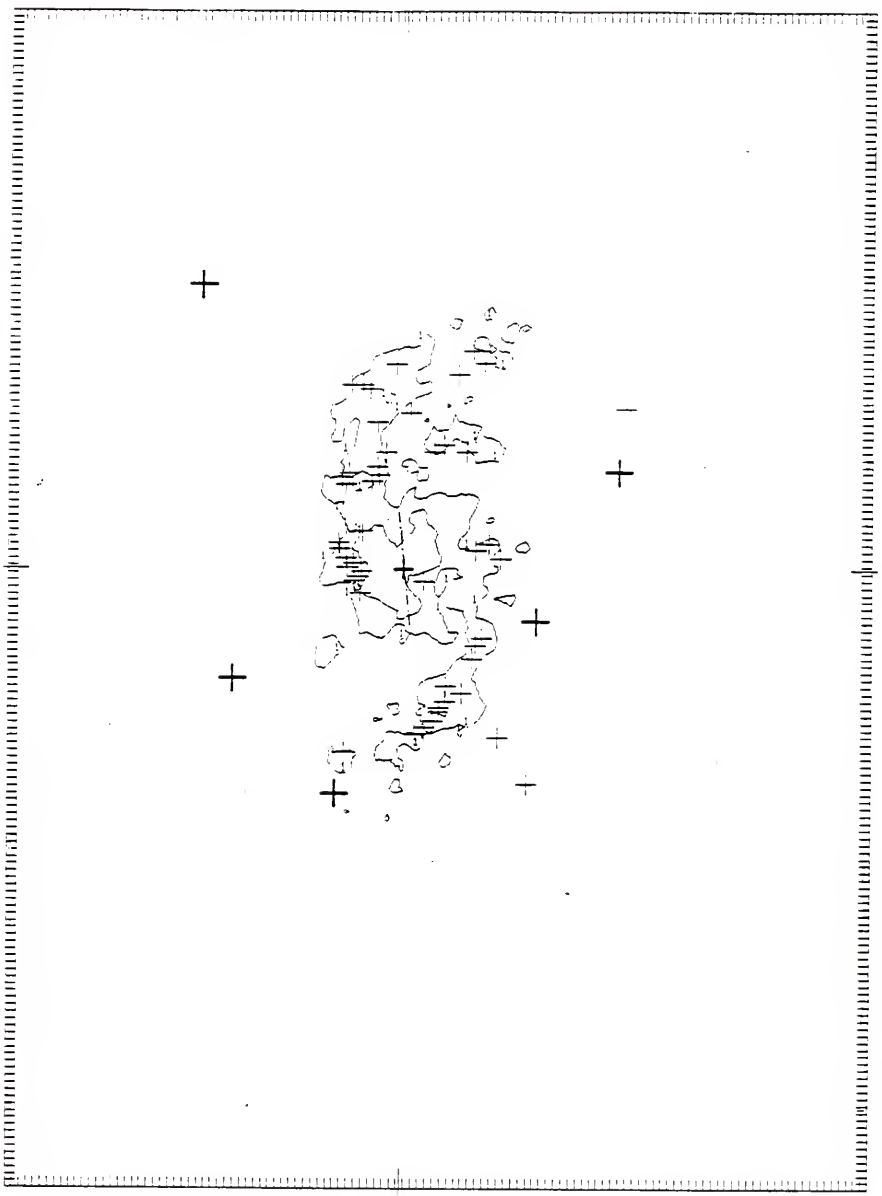
Figure 3-2. The same integrated H I map as in Figure 3-1 is shown here in grey tone format. The positions of foreground stars, and of the galaxy center, are once again marked by crosses. This figure aids identification of local maxima and minima in Figure 3-1.



Perhaps the most striking characteristic of the integrated density map of Figures 3-1 and 3-2 is the patchy distribution of the density peaks. The spiral structure can be traced clearly, but it does not completely dominate the gas distribution in the classic "grand design" style, as seen, for example, in the somewhat later-type barred spiral NGC 4731 (Gottesman et al., 1984). (Note, however, that the relative resolution for the present data is finer by about 30%.) In this clumpy distribution of H I peaks, NGC 3359 resembles the well-observed Sc I galaxy, M 101 (e.g., Allen and Goss, 1979), except that NGC 3359 may be a more extreme case. Nevertheless, the galaxy is definitely classed as a grand design spiral, in the sense of showing two principal spiral arms which can be traced continuously over a large radial extent. This classification was established by Elmegreen (1981) on the basis of blue and near infrared plates, and is confirmed by the H I data presented here. In order to illustrate more clearly the large-scale structure of the integrated hydrogen map, Figure 3-3 presents this map once more, this time showing only a single contour, that at 1.8×10^{21} cm^{-2}. The positions of all the H II regions in Hodge's (1969) catalog are printed as light crosses on this map, to facilitate comparison with optical images, and the location of the optical bar is sketched as well.

The H II and H I are strongly correlated in position. The two principal spiral arms are prominent in both types of emission, along with some irregular features. We have made a crude determination of the pitch angle of the spiral, by measuring the loci of a number of points which lie on fairly well-determined parts of the spiral pattern. The measurements were then used to estimate the pitch angle α, assuming a logarithmic spiral of the form

Figure 3-3. Distributions of H I and H II in NGC 3359. The contour is the 1.8×10^{21} cm^{-2} level of neutral hydrogen, from the same data as illustrated in Figures 3-2 and 3-3. The light crosses mark the locations of H II regions, from the catalogue of Hodge (1969). Darker crosses indicate the fiducial star positions, and the galaxy center. The location and extent of the optical bar are given by the dash-dot line. The field shown here is precisely the same as in Figure 3-1.



$$\theta(R) = \frac{\ln(R/R_0)}{\tan \alpha}. \qquad (3\text{-}1)$$

R_0 is a radius where the spiral arm crosses the major axis, in our case at around 0!9. The measurements, and the resulting values of α, are presented in Table 3-1. The first two columns give the polar coordinates in the sky plane, r and ϕ, the latter having as its origin the line of nodes (see below). These are converted to coordinates (R,θ) in the galaxy plane by the transformations

$$R = \frac{r(1 - \cos^2 \phi \sin^2 i)^{1/2}}{\cos i}, \qquad (3\text{-}2)$$

and

$$\tan \theta = \frac{\tan \phi}{\cos i}. \qquad (3\text{-}3)$$

The fit of the observations to a logarithmic spiral is imperfect, not surprisingly. The western arm is a better approximation to this form, especially beyond the inner portions, where small errors in estimating (R/R_0) can be important. This arm is described reasonably well as a logarithmic spiral of pitch angle 19°. The eastern arm, the less clearly organized of the two, appears to show a steady increase of pitch angle, although the measurements at small radii may be contaminated by the complex emission not clearly associated with the principal arm. The

TABLE 3-1
LOGARITHMIC SPIRAL FIT TO FIGURE 3-3

	r	ϕ	R	θ	α
Western Arm	0.92	0°	0.92	0°	-
	0.92	42	1.20	55	15.3
	1.47	126	2.08	114	22.2
	1.65	141	2.09	127	20.2
	2.08	161	2.24	151	18.6
	2.43	168	2.51	161	19.5
	2.80	188	2.84	193	18.5
Eastern Arm	0.87	0°	0.87	0°	-
	0.78	48	1.06	61	7.3
	0.87	93	1.38	92	13.9
	1.39	148	1.66	135	13.9
	2.54	164	2.69	156	21.5
	3.38	189	3.45	194	21.2

pitch angles derived for the two arms in the outer regions are fairly similar. Since these values refer to the cumulative winding from the center out, this agreement indicates a reasonable degree of large-scale symmetry between the two arms.

The western spiral arm, which is the dominant spiral feature in both the blue and the near infrared passbands, is also the most coherent structure in the integrated 21 cm map. However, many of the small-scale details of the gas density seem to have little to do with the spiral pattern. The arm-interarm contrast (defined as the ratio between the highest and lowest observed densities in the region of the disk well beyond the bar) is also quite moderate, being approximately 2.5 in the densest patches of the spiral arms, and as low as 1.6 in many places. Of course, this ratio is reduced somewhat by beam smearing, but that effect is minimized by the relatively good angular resolution. The 18" synthesized beam (0.95 kpc at the assumed distance of NGC 3359) is shown to scale in Figure 3-1. The arm-interarm contrast is somewhat higher, on the average, when calculated for the stronger, western arm. The actual H I column densities in this arm, however, are no higher than those found scattered elsewhere across the disk. Owing to the patchiness of the density distribution, it is very difficult to calculate its width. One can trace a broad ridge of emission, whose outer boundary is at a level of about 1.6×10^{21} cm^{-2}, and which seems subjectively to trace the general trend of the spiral structure. The typical width of this ridge varies from 45" to 70", implying a width, after a crude deconvolution with the beam, on the order of 2 to 3 kpc. At slightly higher levels of column density, the feature is broken up into a succession of knots at scales down to the limit of resolution.

The spiral structure away from the western arm is somewhat difficult to trace in detail. The area north of the center shows particularly

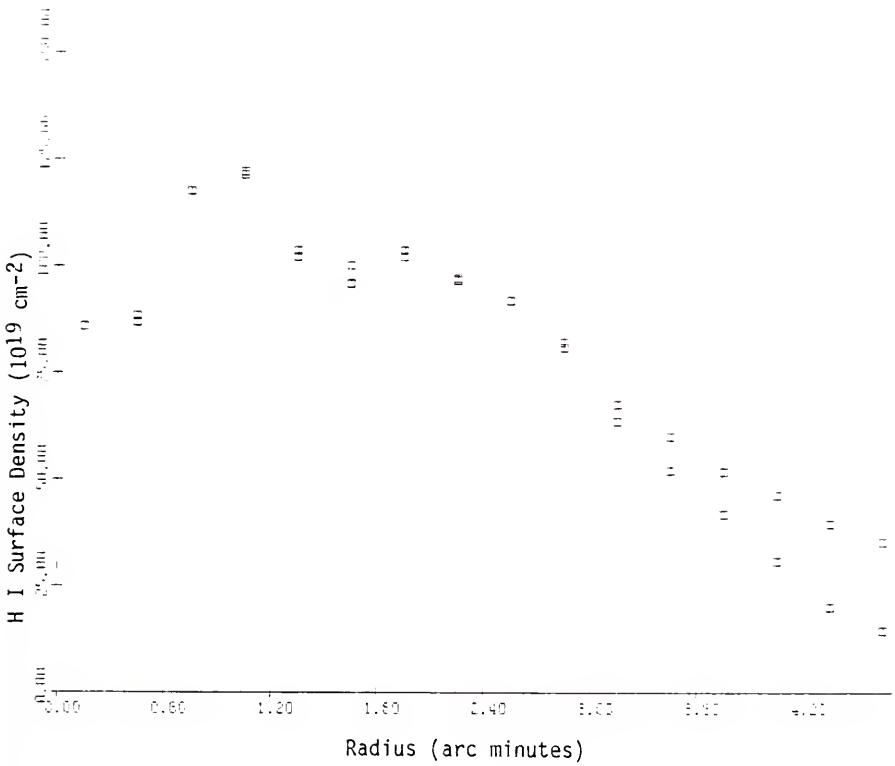
complex structure. The eastern, and weaker, of the two principal spiral arms can be traced fairly readily, although it is less obvious than the western arm. This is partially on account of the presence of additional regions of enhanced density. The most prominent of these regions is located in the northwest quadrant of the galaxy, slightly beyond the end of the bar. It occupies a position roughly symmetrical with the interarm region on the southern side of the galaxy. The broad density maximum formed by this feature more or less abuts, on its northeast side, that associated with the weaker arm. It would seem on dynamical grounds, however, that this anomalous enhancement is quite unlikely to be the continuation of that arm, as it is found at a deprojected radius well interior to the arm. The arm would thus have to curve rather sharply inward, if the two are in reality a single feature. It is more probable that their apparent merging is an effect of the finite resolution. On the Palomar Sky Survey print of the galaxy, several short arm-like spurs are visible within this enhancement (cf. Figure 1-1), and these may well be related to the H I structure. Several H II regions are also located in these spurs, as seen in Figure 3-3, suggesting that the appearance of this small region of the galaxy is due to a local enhancement of the gas density, with associated quasi-spiral structure. If the complicating emission of this feature were removed, the H I column density distribution in NGC 3359 would be fairly symmetric, when comparing the two principal spiral arms.

Toward the nucleus of the galaxy, the gas distribution might be characterized as forming an irregular, slightly elongated ring, of radius approximately 1', or 3.2 kpc at the assumed distance of 11 Mpc (Table 1-1). Alternatively, the appearance of a ring could be caused entirely

by the continuation to small radii of the two principal H I arms, which blend together in this area because of the finite beam. In that case, the apparent elongation is simply caused by the nonzero pitch angle of the arms, which first emerge clearly as distinct entities from two diametrical points on this ring. The region interior to the ring coincides fairly closely with the zone that would be swept by the bar in a rotation about its center; the bar radius, when deprojected, is approximately 0.'85. The H I surface density in that zone is, for the most part, somewhat lower than in the ring; without correcting for resolution effects, the ratio of surface densities between the ring and the interior zone is in the range from 1.3 to 1.8. However, the depletion is not uniform throughout the zone. On its western side, there is an area where the density is as high as at any point in the ring itself. Thus, we have another instance of moderate departure from detailed bisymmetry in the galaxy. The ring, and the slight depletion in its interior, can be seen plainly in Figure 3-4, which plots the angle-averaged, deprojected surface density of H I against radius.

There are several noteworthy phenomena in the outer parts of NGC 3359. For the most part, the neutral hydrogen is confined to an inclined disk of radius approximately 4.'1, or 13.1 kpc. This disk is bounded, around most of its circumference, by a fairly regular and distinct edge. The diameter of the H I disk is, therefore, 20% larger than the isophotal major axis, D_{25}, given by de Vaucouleurs, de Vaucouleurs, and Corwin (1976). At the edge of this disk, the mean H I surface density decreases sharply from about 3×10^{20} cm^{-2} to something below 2×10^{20} cm^{-2}. Unfortunately, it is difficult to determine the mean location of Bosma's (1981b) proposed H I "isophotal" level, 1.82×10^{20} cm^{-2}, very precisely from these data. When adjusted for the inclination of NGC 3359, this level is 2.9×10^{20} cm^{-2}. Virtually all of the

Figure 3-4. Mean deprojected H I surface density in NGC 3359, as a function of radius. Units of the ordinate are 10^{20} cm^{-2}. Each plotted point is obtained by averaging the surface density in an elliptical annulus of constant deprojected radius. When two joined points are shown, they indicate upper and lower bounds in a region where a low signal-to-noise ratio prevails.


H I Surface Density (10^{19} cm^{-2})
0.00
0.60
1.20
1.80
2.40
Radius (arc minutes)

pixels within the H I disk exceed this column density until one reaches its edge, where the density goes to zero abruptly. No doubt this is primarily a result of the signal-to-noise characteristics of the spectra. Beyond this radius of 4!1, calculation of the mean surface density is misleading, as the emission arises solely from a few scattered patches of lumpy brightness.

Beyond the radius of the H I disk, some intriguing H I features may be seen in Figures 3-1 and 3-2. Their general appearance is of segments of a broken ring, or of ragged, fragmentary outer arms, the two most prominent of which emerge from the edge of the disk at position angles of approximately 70° and -40°, relative to the nucleus. The latter of these could conceivably be a weak extension of the eastern principal arm, but the former emerges from the disk at too small a deprojected radius to be a continuation of the dominant principal arm. There is also a feature, in the southwest quadrant, whose status is less clear. It could be a separate entity, located approximately symmetric to the "outer arm" segment which originates from position angle 70°. Alternatively, it might be simply an enhancement in a faint background medium which joins the H I disk to the other "outer arm" feature seen on the western side of the galaxy, which we have already mentioned. As we shall see in the following section, the rotation curve suggests that all of this material probably lies in the same plane as the main H I disk.

These ragged features are quite complex, when examined in detail. It must be emphasized that the "empty" regions, which separate the islands of H I in these features from one another, and from the main H I disk of the system, have been searched exhaustively for any spectral line signal, by the procedures described earlier. To the limit of

detectability, none was found. This statement is equally true for a rather large area beyond the outer perimeter of the entire region of emission. It is primarily these gaps that give these outer arms their appearance of being faint features. Where emission is detected, the typical column density is about 20% of the peak value found anywhere in the galaxy, with many knots in which the ratio rises to 40% and even 50% of the peak. Thus, the neutral hydrogen in these regions is evidently distributed in a very nonuniform, "clumpy" manner.

The overall distribution and intensity of this matter are, rather obviously, asymmetric. However, probably no particular importance should be attached to this observation; in dealing with features as irregular as these, large-scale symmetry is not to be expected. It is more interesting to speculate on the question of whether their true nature is more aptly characterized as being armlike or ringlike. Though neither interpretation can be ruled out by these observations, we tend to favor the armlike hypothesis, largely on the basis of the northeastern feature. That feature clearly shows an increasing deprojected radius along its length, implying a nonzero pitch angle. As a final note on this outer zone of neutral hydrogen, we note that the material in the southwestern outer arm is at the greatest galactocentric radius observed to have H I emission in this study, with the exception of the satellite galaxy; assuming it is in the same plane as that of the H I disk, it extends to a distance of 28 kpc from the nucleus.

The column densities of Figures 3-1 and 3-2 can be integrated in solid angle to give the total H I mass in the galaxy. The result is $M_{HI} = 3.85 \times 10^7 \, M_\odot \times D^2$, where is the distance to NGC 3359 in Mpc. With the assumption D = 11, the total H I mass is then $4.66 \times 10^9 \, M_\odot$. As we shall see shortly, a rough correction for the zero spacing

flux increases the total hydrogen mass to approximately 4.9×10^9 $M_\odot$. For comparison, Rots (1980) found 5.8×10^9 $M_\odot$, and Fisher and Tully (1981) 6.1×10^9 $M_\odot$, while the early interferometry by Gottesman (1982) detected only 1.8×10^9 $M_\odot$, primarily because he had no spacings shorter than about 100 m. Apparently, some neutral hydrogen may have been missed in this survey, despite our efforts to detect outlying regions of low surface brightness. This could be caused by the presence of a significant amount of H I in a component of half-power diameter approximately 8' or larger in a single channel, for which our sensitivity is significantly reduced on account of the (u,v) coverage. Inspection of the channel maps (Figure 2-4) suggests that this may be of some relevance near the minor axis. However, it is more probable that the missing emission is simply at a surface brightness too low to be detected with our signal-to-noise criteria.

The continuum flux was estimated from the results of the spectrum integration, as described in the previous section. The map resulting from this procedure is shown as Figure 3-5. This method of determining the continuum is somewhat prone to systematic errors, as discussed by Gottesman et al. (1984). In particular, spectra which fail to have a signal of twice the rms level of the single-channel maps in at least one channel are not integrated, and, hence, no estimates of their continuum fluxes are available. At points which do pass this test, the continuum is computed as the mean of the channels outside the window containing line emission. Because of the special effort made to detect broad, low-level H I in NGC 3359, the systematic bias of the continuum map should be minimized. However, as a glance at Figure 3-5 shows, no continuum emission was detected, except from the inner portions of the galaxy. The

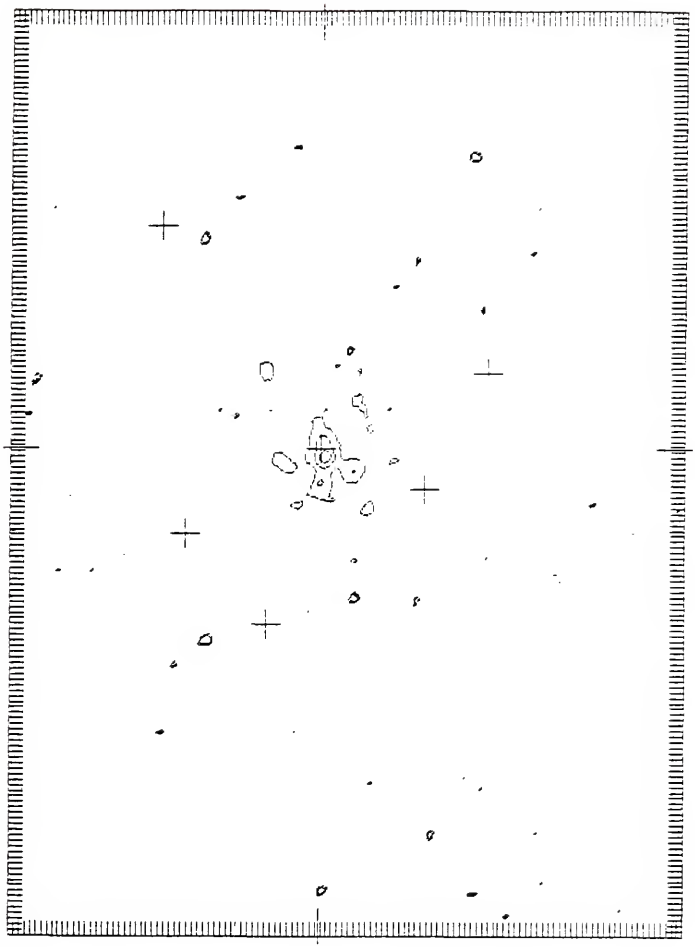


Figure 3-5. Continuum emission in NGC 3359. The contour interval is 1.2 K, or about twice the rms uncertainty. Negative contours are drawn as double lines. The correction for the zero-spacing flux, discussed in the text, has not been applied to these data. Crosses mark the positions of the foreground stars used as fiducial marks.

rms uncertainty in this map, assuming the noise is random, is 0.58 K, or 0.30 mJy per beam. The total continuum flux detected is 15.7 mJy.

There is a slight negative bias in the continuum map, indicating that the problems associated with the lack of a zero spacing flux have not been completely eliminated by the use of the CLEAN program. Investigation of several independent regions in the galaxy reveal that the mean value is about -0.25 K, with an uncertainty of about 0.05 K. There is no significant variation in this level across the map, implying that the effect is due to the missing point in the center of the (u,v) plane, as discussed in Chapter II. The magnitude of this DC offset is an order of magnitude below the rms noise level of the individual single-channel maps, which we saw in Chapter II to be 2.13 K. There will be two minor consequences of this offset in the integrated density map. First, the signal-to-noise criteria used to discriminate real line emission from noise will be slightly more stringent than we have stated; in practical terms, this effect is more than compensated for by our painstaking inspection of individual spectra. Second, the detected densities, and, hence, the total mass, are slightly underestimated. If we assume that an offset of 0.25 K should be added in every channel of the spectra, and that an average spectrum contains five channels of signal, the density should then be increased at each point in the galaxy by about 5×10^{19} cm^{-2}, which is less than half the rms uncertainty in the density map, and less than 20% of the lowest contour level in Figure 3-1. When integrated over the entire extent of the galaxy, the total mass associated with the offset is 2.4×10^8 $M_\odot$. Adding this corection to the directly determined mass of 4.66×10^9 $M_\odot$ gives a total neutral hydrogen mass of 4.90×10^9 $M_\odot$.

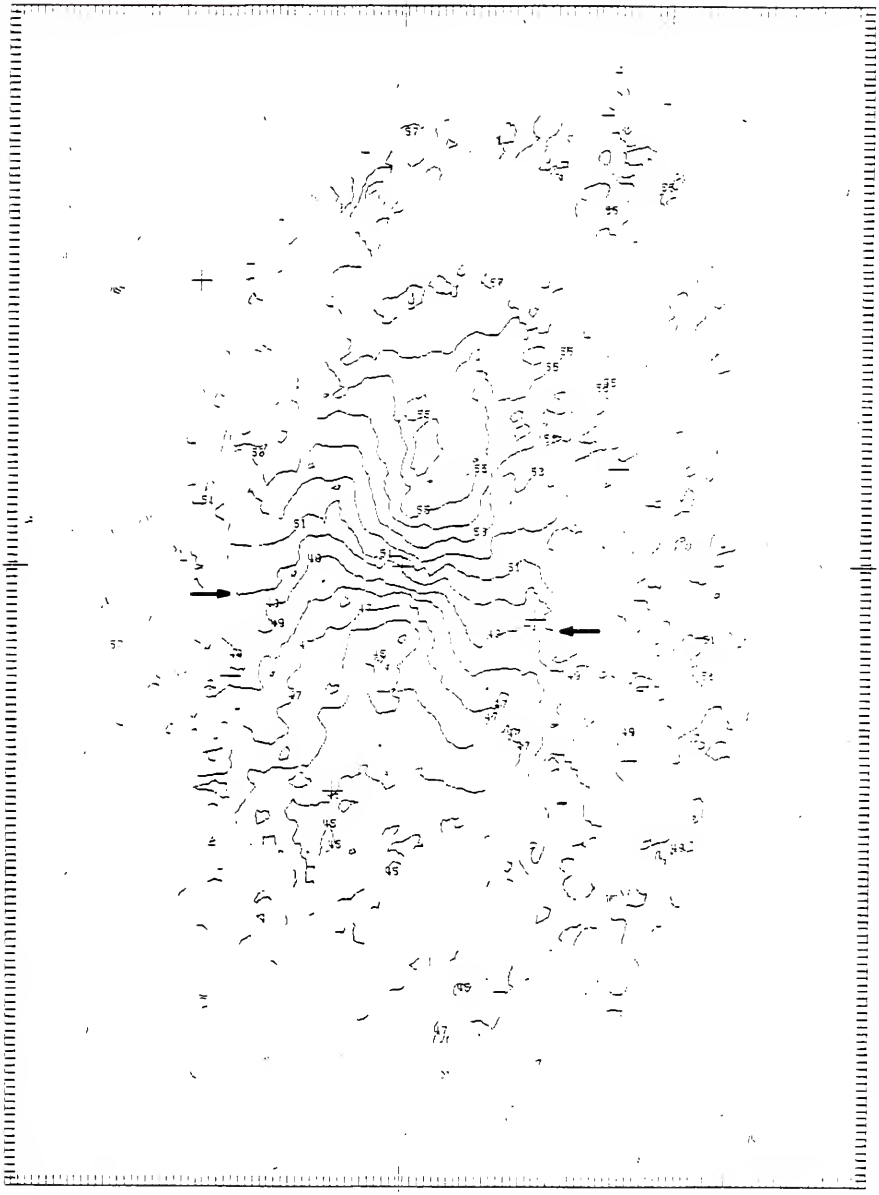
The brightest continuum source in NGC 3359 was found to be the nucleus, and there is some evidence for a continuum enhancement roughly coincident with the bar. The remaining emission seen in Figure 3-5 may be associated with the southwestern portion of the inner, broken ring, although the detection of this emission is only marginal. The continuum feature found northwest of the nucleus by Gottesman (1982) is not detected.

Kinematics of the Neutral Hydrogen

Figure 3-6 presents the temperature-weighted, mean velocities of the neutral hydrogen in NGC 3359, as determined by applying equation (2-2) to the observed spectra. The spectra were Hanning smoothed in velocity space before the integration. This convolution was performed because the original sampling rate of the data, 20.7 km/s, is of the same order as the filter bandwidth, 25.2 km/s. Additionally, the smoothing improves the signal-to-noise ratio before the integration. The velocity resolution of these smoothed data, therefore, is 41.4 km/s, but the channel separation is still 20.7 km/s, thereby giving optimal, Nyquist sampling. The contour interval in Figure 3-6 is 20 km/s.

Perturbations in the velocity field, associated with the spiral structure of the galaxy, are readily visible in this figure. In fact, the principal spiral features are considerably more striking in the velocity contours than in the surface density maps. The contours in the bar-dominated region are skewed somewhat toward the direction parallel to the bar. This is a common feature in hydrodynamical models of barred galaxies (Huntley, 1982; this dissertation, Chapter V), and has been observed, for instance, in NGC 5383 (Sancisi, Allen, and Sullivan, 1979), though not in the pronounced degree seen here. The kinematical

Figure 3-6. Contours of heliocentric, line-of-sight velocity in NGC 3359. The contour interval is 20 km/s. A number of contours are labeled with integers, which are the associated velocities in units of 20 km/s. The 1000 km/s contour level is indicated by the arrows. As in all the figures of this dissertation, north is at the top of the figure, east to the left. Foreground star positions are marked by the crosses, along with the center of the galaxy. The tick marks around the border are separated by 0.1.



effects of the two principal spiral arms are also quite obvious, appearing as distinct bends in the velocity contours at positions which coincide with those of the H I and optical arms. We interpret these distortions as indicative of streaming motions associated with the spiral arms, rather than as resolution-dependent, density-weighting effects (Siefert, Gottesman, and Wright, 1975), both because the features are fairly well resolved by our 18" beam, and because the sense of their correlations with the H I densities is opposite from that expected in the latter case. The locations of these distortions are symmetric about the center of the galaxy, when one compares positions in one arm against the other. Indeed, perhaps the most interesting characteristic of the velocity field is the symmetry, both of position and of degree, of this distortion wave. The western arm, which is significantly the stronger one, both in photographs and in H I column density, causes velocity perturbations no more severe than its apparently weaker counterpart. This finding strengthens the hypothesis that the asymmetry in NGC 3359 is in truth a rather mild one, and that the basic dynamical behavior of the galaxy can be understood in terms of a bisymmetric model. Certainly, Figure 3-6 shows little difference in the kinematical behavior of the gas associated with the two arms. As suggested earlier in this dissertation, it is possible that the apparent dominance of the western spiral arm is partly due to its being somewhat more coherently organized along its whole length, especially in optical images. This arm may, for example, have experienced a recent, coherent burst of star formation, as suggested by Hodge's observations of the H II regions which delineate it so well.

The velocities in the outer, ragged arms or broken ring are quite chaotic on a small scale, as suggested by the velocity contours drawn for them in Figure 3-6. On a large scale, they are dominated by the overall rotation field of the galaxy. As we shall see shortly, the rotation curve gives no indication of a change in the kinematics of the galaxy beyond the 4:1 radius of the H I disk.

In order to put our discussion of the kinematics of NGC 3359 on a more quantitative basis, it is necessary to derive certain global parameters of the rotation field. These arise from a decomposition of the velocity V_{obs} into the components V_{rot}, the rotation velocity, and V_{nc}, the noncircular term. If we also define V_{sys} to be the heliocentric, systemic velocity, the observed velocity at each point in the galaxy can be written as

$$V_{obs}(R,\theta) = V_{rot}(R)\cos\theta\sin i + V_{nc}(R,\theta) + V_{sys}, \qquad (3\text{-}4)$$

where i is the inclination, with respect to the sky, of the plane of the galaxy, in which the polar coordinates of the point in question are R and θ. The origin of θ is the line of nodes of the galaxy plane. In a galaxy with departures from circular rotation as obvious as those in NGC 3359, determination of i and of Φ_{lon}, the position angle on the sky of the line of nodes, cannot safely be made from an analysis of the velocities. The position angle Φ_{lon}, for example, is strongly dependent on radius, because of the noncircular velocities present. At the same time, because of the high degree of symmetry of the noncircular response seen in the two principal arms, any line through the center of the galaxy divides it into symmetrical halves, so that no unique choice can be made. Therefore, we have adopted values of i and of Φ_{lon}

determined simply by deprojecting the disk of the galaxy into a circle. De Vaucouleurs, de Vaucouleurs, and Corwin (1976) find i = 51°, upon applying this deprojection to an optical isophote map of the galaxy. Estimating the location of the H I disk from Figures 3-1 and 3-2, we arrive at a range of values which is quite similar, i = 48° to 57°. We measure Φ_{lon} to be in the range -5° to -12°. However, these measurements are rather uncertain, as the patchy outer features described in the previous section make the edge of the H I disk somewhat ill-defined. Gottesman (1982) derived, from his low-resolution studies of the velocity field, i = 51° and Φ_{lon} = -8°, and we have found no reason to prefer a significantly different set of numbers. Therefore, for the sake of making comparison easier, we have adopted these values. We did determine V_{sys} independently, by taking the mean velocity of the data within a deprojected radius of 4!5. The value we obtain, 1009 km/s, is slightly lower than Gottesman's value of 1024 km/s, but he suspected that his velocity coverage was incomplete. The value given in the Revised Shapley-Ames Catalog, which is the mean of a number of independent determinations, is 1008 km/s (Sandage and Tammann, 1981).

With these orientation parameters decided, we can proceed to derive the rotation curve of the galaxy. Two methods were adopted. First, an angle-averaging scheme, introduced by Warner, Wright, and Baldwin (1973) was employed. This method uses all the observed data to determine the mean rotational component of the velocities in a filled elliptical annulus, which represents a circular ring in the plane of the galaxy. All points are weighted by the cosine of the angle θ, so that those near the minor axis have lower weight. The velocities are deprojected, using our assumed orientation parameters. The rotation curve obtained

in this way is shown as Figure 3-7. The curve is well-defined, with rms errors below 20 km/s, for the entire extent of the data. However, beyond a radius of 6!75 (21.6 kpc), the data consist entirely of points well off the major axis. The line-of-sight velocities in these areas have only a negligible rotational component, so the rotational velocity cannot be determined on the basis of these data. Consequently, the rotation curve has not been plotted to radii beyond 6!75 in Figure 3-7. In fact, much of the formal error in the rotation curve is likely to be produced by the strong deviations from pure circular rotation, which are apparent in the velocity contour map.

The rotation curve was also determined using only points within 15°, measured on the sky, of the position angle of the line of nodes. As before, individual velocities are weighted by the cosine of θ, but the weights vary by less than 4% from unity. The result is plotted in Figure 3-8. The only significant change from Figure 3-7 is that the rms uncertainty attached to each point is higher, owing to the inclusion of fewer data. We shall refer only to Figure 3-7 in discussing the dynamical implications of the rotation curve.

One immediately apparent feature of this rotation curve is the lack of a truncation signature. These signatures have been described by Casertano (1983) and by Hunter, Ball, and Gottesman (1984). They arise when a flattened disk is truncated fairly sharply at some radius: the observed rotation curve shows an abrupt decrease just beyond the truncation edge. Such signatures may have been observed in the edge-on spiral galaxy NGC 5907, by Sancisi (Casertano, 1983), and in the barred spiral NGC 3992, by Gottesman et al. (1984). Even if the disk is embedded in a halo of mass comparable to its own, the signature may be detectable,

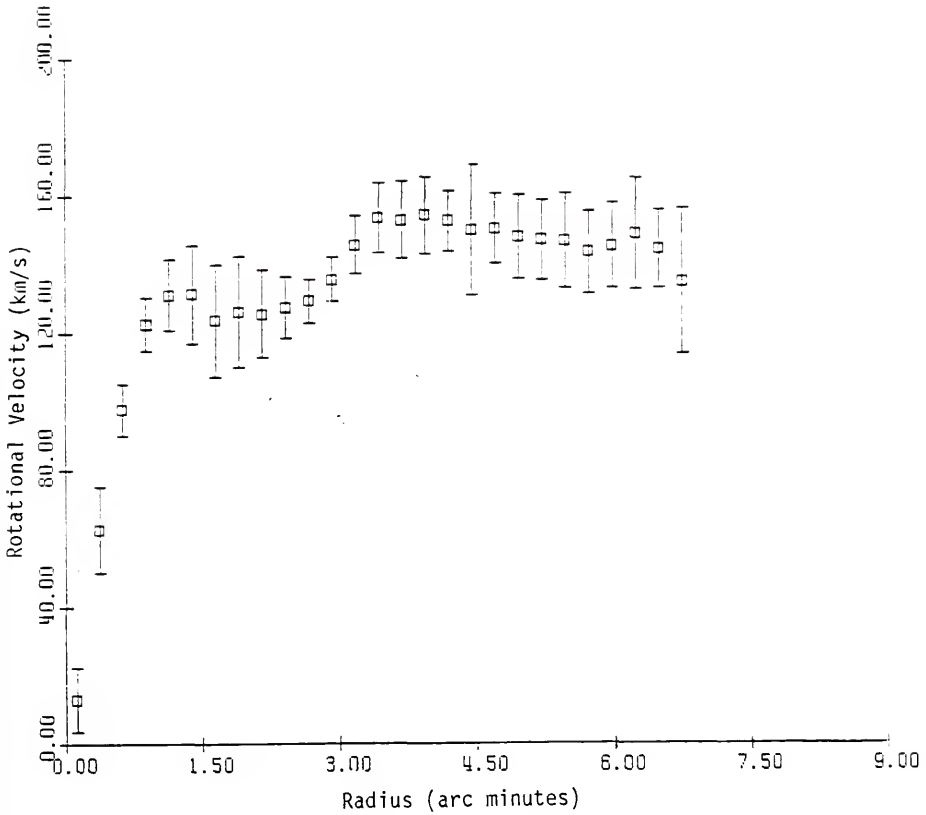


Figure 3-7. Angle-averaged rotation curve of NGC 3359. The line-of-sight velocities have been averaged in concentric, elliptical annuli which are circular in the plane of the galaxy. The deprojection parameters assumed in calculating this rotation curve are: line of nodes position angle, -8°; inclination angle, 51°; systemic velocity, 1009 km/s. Error bars indicating plus and minus one standard deviation are included. These formal errors are largely due to the presence of systematic, noncircular velocities.

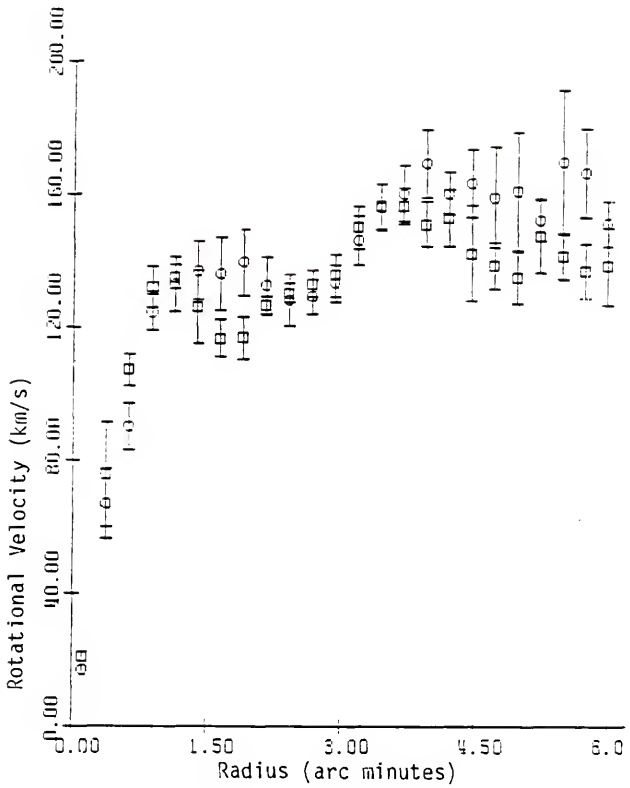


Figure 3-8. Rotation curves determined independently from velocities along the two halves of the major axis. Data from the northern end of the major axis are shown as circles, those from the southern end as squares. Points lying more than 15° from the line of nodes, as measured in the sky plane, have been excluded from this analysis. The deprojection parameters assumed are: major axis position angle, -8°; inclination angle, 51°; systemic velocity, 1009 km/s. Error bars of plus and minus one standard deviation are shown centered on each point.

especially for rotation curves which are flat, or rising, interior to the truncation. In NGC 3992, there is a well-defined H I disk of radius 3!8, with patchy, fainter emission beyond this radius. The apparent truncation signature is found precisely at the edge of the disk, to the resolution of the observations. As we have seen, the neutral hydrogen in NGC 3359 follows a large-scale distribution which is, qualitatively, quite similar to this, the disk radius being 4!1 in this case.

However, there is not the slightest suggestion of a truncation signature in Figure 3-7, either at 4!1 or at any observed radius. At radii larger than about 3!5 (10 kpc), the rotation curve is very close to being flat, with an extremely slow decline, and a mean value of about 150 km/s. The failure of the rotation curve to fall by any substantial amount, even at a radius more than 50% beyond the edge of the H I disk (which, itself, is larger than the photometric disk), obviously requires the presence of a good deal of dark matter. Judging by the absence of a truncation feature, some of this matter almost certainly must continue to be distributed in a disk, even at radii beyond the observed disk radius. At galactocentric distances in excess of 4!1, the determination of the rotation curve is based entirely on the matter in the ragged outer features mentioned previously. The smoothness with which the curve continues into this region, with no indication that a dynamically significant edge has been crossed, strongly suggests that, not only is there no abrupt truncation in the underlying disk, but this broken outer material is in the same plane as that defined by the H I disk.

The rotation curve of Figure 3-7 shows several distinct kinematical regimes. Interior to a radius of 0!75 (2.4 kpc) or slightly less, solid-body rotation is a fairly good approximation to the observed curve,

with a mean slope of approximately (160 ± 15) km/s/arc minute. Note that this is the region of the galaxy inhabited by the bar. The slope we have measured may be compared to the optical data of Cheriguene (1975), who measured the velocities in the inner 1' in the lines of H α and of [N II]. Beyond about 36", his rotation curve begins to show a suggestion of flattening. Restricting ourselves to radii smaller than this value, we have fit a least-squares line to Cheriguene's measurements. We obtain an estimate for the angular velocity of the solid-body portion of the rotation curve, using a deprojection based on our values of the orientation angles, of 169 km/s/arc minute, or 53 km/s/kpc. We have assumed the galaxy to be at a distance of 11 Mpc, as before. The optical velocities are, therefore, in quite good agreement with our radio data, as far as they go. Evidently, the loss of resolution in our data has not seriously affected the determination of this gradient.

The behavior of the rotation curve between about 1' and 3.5 is rather unusual. At the end of the bar, the rotation curve at first flattens sharply, and remains roughly constant, at about 130 km/s, from about 1.2 to 2.5. At that point, the rotation curve again begins to rise, attaining its maximum value of 153 km/s at 3.5, after which it once more levels off, with, as we have mentioned, a very gradual decline at greater radii. The rotation curve of Gottesman (1982) also shows the rise between 2.5 and 3.5 quite clearly (see his Figure 7).

Two principal explanations for this phenomenon suggest themselves. If a significant fraction of the mass interior to 1' is in the bar, and if the bar is truncated rather abruptly, as its optical image seems to imply, the effect on the rotation curve will be qualitatively similar to the truncation feature described above for truncated disks. Since the

rotation curve due to the axisymmetric components is still rising, the net result will be to depress the rotation curve below that which one would observe for these components alone. As one proceeds further out in radius, this effect diminishes at the same time that spiral arm perturbations begin to assert themselves, and the rotation curve rises rapidly. The crossing of the principal spiral arms occurs, along the line of nodes (where the observations are given highest weight), at about 1' and again near 3'. Noncircular velocities associated with the spiral structure are strong in NGC 3359, as we have seen in the isovel map, and may be sufficient to produce the distortions of the rotation curve. If this interpretation is basically correct, we should be able to reproduce these features of the rotation curve in the hydrodynamical models, to be discussed in Chapter V. The alternative is that the galaxy is warped, with the fundamental planes of the gas interior to 2!5, and exterior to 3!5, being inclined at different angles. We will not explore this possibility at any length, for the reason that, as the velocity field of Figure 3-6 shows, the velocity perturbations are very well correlated with the spiral structure. We feel that this makes it unlikely that a warp is also involved somehow in producing these same perturbations.

Beyond a radius of 3!5, the rotation curve exhibits no features other than a long, very slow decline. As an estimate of the slope of this decline, a least squares line fit to the data beyond 3!5 yields a slope of -1.0 km/s/kpc. (This slope is, clearly, very small, but as the points plotted in Figure 3-6 are almost completely independent of one another, we believe that the decline is probably real. It is difficult to evaluate its statistical significance, however, as the errors associated with these measurements are known to contain systematic terms--the

noncircular velocities--of uncertain and radius-dependent magnitude.) As we have discussed above, the near flatness of this curve indicates that a significant amount of matter must still be present between the edge of the H I disk, at 13.1 kpc, and our last measured point, at 21.6 kpc.

First order mass models of NGC 3359 can be fit to the rotation curve, for the purpose of determining its total mass. At the present juncture, we will only employ two simple models, deferring consideration of more exacting mass models to Chapter V. Our approach here will be to fit a "mean curve" which ignores the perturbations referred to above, as our interest at the moment is simply to derive the total mass interior to the last measured point. The first mass model employed consists of a Generalized Mestel Disk, or GMD (Hunter, Ball, and Gottesman, 1984), which is also a member of Toomre's (1963) family of infinitely flattened disks. Hunter, Ball, and Gottesman showed how these disks may be truncated, to avoid the divergence of the total mass, caused by their asymptotically flat rotation curves. Unless one is near or beyond the truncation radius, however, the rotation curve is indistinguishable from that of the infinite GMD, which we shall use here for simplicity, since no truncation signature is detected. The fit of the GMD to Figure 3-7 is not perfect, as the GMD's rotation curve is always rising. Nevertheless, this model fits better than any other known potential consisting solely of a single-component disk.

The circular velocity in the GMD is given by

$$V_0^2(R) = C_0^2[1-b_0(b_0^2 + R^2)^{-1/2}], \qquad (3\text{-}5)$$

and the mass interior to radius R by

$$M_0(R) = \frac{C_0^2 R}{G}\left[\left(1 + \frac{b_0^2}{R^2}\right)^{1/2} - \frac{b_0}{R}\right].$$

(3-6)

The rotational velocities of the model adopted as a preliminary fit to NGC 3359 are plotted as open circles in Figure 3-9. This model is parameterized by the length scale, b_0 = 0!5, and the amplitude, C_0 = 155 km/s. The mass interior to the last measured point, at 6!75 or 21.6 kpc, therefore, is 1.12×10^{11} $M_\odot$.

The second mass model considered was chosen on the basis of the apparent decline in the rotation curve at large radii. This behavior can be reproduced only by disks which are more centrally condensed than the GMD, but such disks generally have rotation curves which fall much more rapidly than observed for NGC 3359. We have used Toomre's n = 1 disk, whose rotation curve declines more gradually than those of other members of the family (except for the GMD), and embedded it in a spherical halo to keep the rotation curve from falling too steeply. The circular velocity associated with the disk alone is

$$V_1^2(R) = \frac{C_1^2 R^2}{b_1}\left(b_1^2 + R^2\right)^{-3/2},$$

(3-7)

and its mass, within radius R,

$$M_1(R) = \frac{C_1^2}{G b_1}\left[1 - b_1\left(R^2 + b_1^2\right)^{-1/2}\right].$$

(3-8)

In this case, the amplitude parameter C_1 has the units of velocity multiplied by distance.

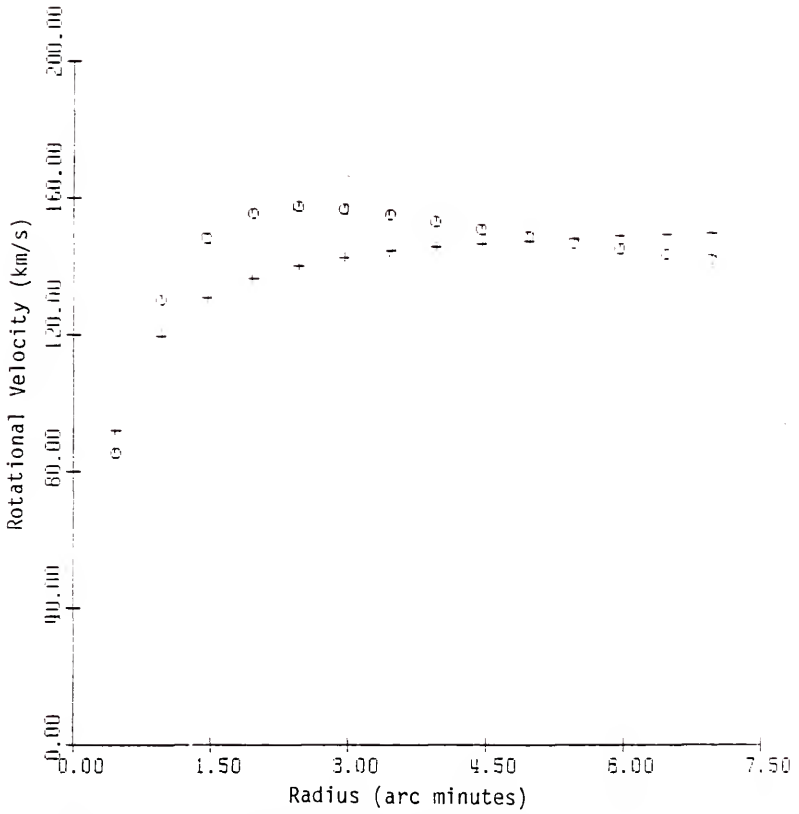


Figure 3-9. Rotational velocities of the first-order mass models used in this chapter to estimate the mass of NGC 3359. The crosses are the velocities associated with the Generalized Mestel Disk, while the circles are from a superposition of a Toomre disk, of index n = 1, and a spherical halo. Details of these mass models are given in the text.

The maximum value of the rotation curve due to this disk alone, which occurs at $R = \sqrt{2}\, b_1$, is given by

$$V_{max} = \left(\frac{4}{27}\right)^{1/4} \frac{C_1}{b_1} .$$

(3-9)

The halo employed is similar to those considered by Hunter, Ball, and Gottesman (1984): its density is constant out to some small core radius r_c, beyond which it declines as r^{-2}. The condition that the density be continuous at the boundary between these two zones leads to the expression for the mass interior to radius R:

$$M_H(R) = \frac{K}{G} \left(\frac{R}{r_c}\right)^3, \quad R \leq r_c;$$

(3-10a)

$$M_H(R) = \frac{K}{G} \left(\frac{3R}{r_c} - 2\right), \quad R > r_c.$$

(3-10b)

Since the mass is assumed to be distributed with spherical symmetry, the velocity associated with this halo is simply

$$V_H^2(R) = \frac{GM_H(R)}{R} .$$

(3-11)

The velocities V_1 and V_H add quadratically, since the resultant force is a linear superposition. The model used to represent NGC 3359 with these potentials is specified by the parameters: $b = 1.625$, $C_1 = 960$

(km/s) kpc; r_c = 0.5, K = 9000 $(km/s)^2$ kpc. The rotational velocities of this model are plotted as crosses in Figure 3-9. The resulting masses, inside radius 6.75, are $M_1 = 4.93 \times 10^{10}\ M_\odot$, and $M_H = 8.06 \times 10^{10}\ M_\odot$. Hence, in this particular model the halo mass is about half again that in the disk, within the radius of our last measured point. These numbers should be regarded only as roughly indicative of the possible disk-halo mass ratio, as the model is hardly unique. The combined mass, $1.30 \times 10^{11}\ M_\odot$, is quite similar to that derived using the earlier mass model. This second model is a slightly better fit to the data, especially if a bar of substantial mass is believed to be present as the cause of the flattening between 1.2 and 2.5. However, that is not surprising, as the number of parameters used to specify the mass distribution has been doubled. The agreement of the two estimates of the total mass inside 21.6 kpc, $1.2 \times 10^{11}\ M_\odot \pm 10\%$, is the principal result of these approximate calculations.

The global profile of the neutral hydrogen in NGC 3359 is plotted in Figure 3-10. The spectra from which this Figure was produced were first Hanning smoothed. No correction for bandwidth smoothing has been applied.

Table 3-2 summarizes our principal findings concerning the density and velocity fields of the H I in NGC 3359.

The Companion Galaxy to NGC 3359

In our presentation of the single-channel maps in Chapter II, we noted that in two of these maps--those at heliocentric velocities of 953 km/s and 974 km/s--emission is detected from a position about 15' to the southwest of the center of the galaxy. In the channel maps, this feature appears to be an isolated H I cloud, with no detectable bridge of neutral hydrogen joining it to the contiguous body of H I. In this section, we discuss this isolated cloud as a small satellite galaxy to NGC 3359.

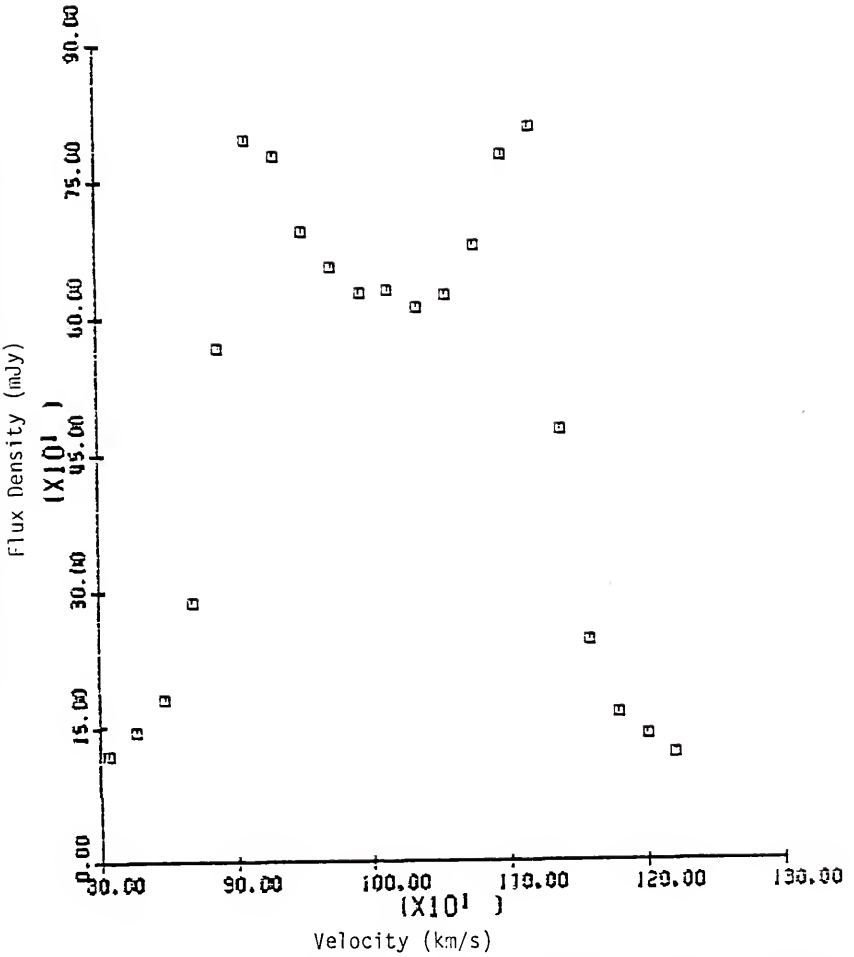


Figure 3-10. The observed global spectrum of neutral hydrogen in NGC 3359. These data have not been corrected for inclination, nor for the 41.4 km/s instrumental profile of the Hanning smoothed observations.

TABLE 3-2
SUMMARY OF NEUTRAL HYDROGEN OBSERVATIONS

Synthesized Beam, FWHP	18.″04 x 17.″64
	0.96 x 0.94 kpc
Position Angle of Beam	-53.°4
Channel Separation	20.7 km/s
Filter Width	25.2 km/s
Continuum Flux	15.7 mJy
Peak H I Column Density	3.51×10^{21} cm^{-2}
RMS Noise in Density Map	1.30×10^{20} cm^{-2}
Diameter of H I Disk	8.′2
	26.2 kpc
Radius of Greatest H I Extent	28 kpc
Inclination Angle	51°
Position Angle of Line of Nodes	-8°
Systemic Velocity	1009 km/s
Width of Global Spectrum[a]	326 km/s
Maximum Rotational Velocity[b]	153 km/s
Radius of Maximum Velocity	11.2 kpc
H I Mass, M_{HI}	4.9×10^9 M
Total Mass, M_T (within 21.6 kpc)	1.2×10^{11} M
M_{HI}/M_T	0.040
M_{HI}/L_B (Solar Units)[c]	0.45
M_T/L_B (Solar Units)[c]	11.1

[a]Full width at 20%, corrected for inclination and for a Hanning-smoothed filter width of 41.4 km/s.
[b]Corrected for inclination.
[c]L_B = corrected blue luminosity from Table 1-1.

The observed spectra in the area of sky containing this object were integrated in the same manner as those in NGC 3359 itself, to determine the column densities and mean velocities of neutral hydrogen in this feature. In order to ascertain whether the feature is indeed isolated from the H I of NGC 3359, the entire region separating the two was integrated and then searched for H I which might plausibly be real. To enhance the sensitivity to faint emission, the signal-to-noise criterion for the integral at each pixel was relaxed, from the requirement that it be three times the rms uncertainty, to allow it to be only twice this value. The manual inspection of spectra, described in the first section of this chapter, was also performed for this entire region of sky. The automated and manual searches agree in concluding that no H I is detected in a gap of more than 7' which separates the isolated cloud from NGC 3359.

The integrated density map of an area centered on the cloud is shown in Figure 3-11. The peak column density, after correcting for the attenuation of the signal by the VLA primary antenna beam, is 2.93×10^{21} cm^{-2}. Owing to the less stringent signal-to-noise criteria used in the integration routine, the map is quite noisy. The small cross marks the position $\alpha = 10^h41^m6$, $\delta = +63°21'.7$ (1950.0), which was a preliminary determination of the cloud's position, based on a single-channel map. The center of the integrated H I image is located about 0'.1 east and 0'.2 north of this position.

Figure 3-12 presents a tracing, of the outer contour level of the main patch of emission, onto an enlargement of the area from the Palomar Sky Survey. An examination of the Sky Survey prints at suitable magnification reveals the presence of a faint, extended image on the blue print.

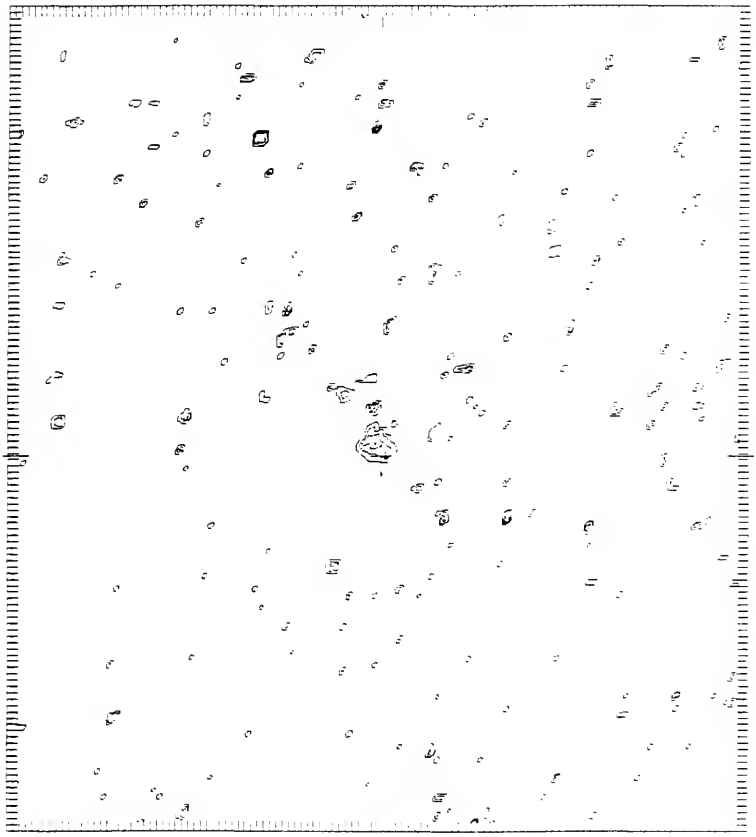


Figure 3-11. Contours of integrated H I column density in the area surrounding the isolated companion to NGC 3359. The contour levels are multiples of 4.8×10^{20} cm^{-2}. The position marked by the cross is at right ascension 10^h41^m6, declination +63°21'.7. North is at the top, east to the left, in this plot, and the tick marks around the border are spaced 0'.1 apart. Refer to Figure 3-12 for the location of this object, relative to NGC 3359.

Figure 3-12. Relative locations of NGC 3359 and its satellite. The central position of the companion is 14!8 southwest of the nucleus of NGC 3359. The outer contour of the H I in the satellite, from Figure 3-11, is traced at the correct position on an enlargement of the National Geographic Society--Palomar Sky Survey blue print. ©1960, National Geographic Society-Palomar Sky Survey. Reproduced by permission of the California Institute of Technology.

The image is so faint and amorphous on the red print as to be almost undetectable.

The character of the integrated density map of the cloud is also quite amorphous. The main part of the image, and the only part which can be regarded as real with a reasonable degree of certainty, shows no clear structure, though its diameter is about 0!6, twice the beam width. About 1' northeast of this main cloud, some lower-level emission is seen. On a photographic display of the single-channel maps, this material gives an impression of extending continuously to the main cloud, but, as Figure 3-11 shows, the fainter material is only marginally above the noise. There is also a small, relatively bright knot just north of the main cloud, which has a slightly different velocity (by about 15 km/s) than any other hydrogen in this area. This knot may well be caused by a noise spike. However, we have treated all three of these components as representing real signal, for the purposes of determining the H I mass of the cloud. After primary beam correction, the total H I mass is 4.6×10^7 $M_\odot$. This assumes that the distance to the feature is 11 Mpc.

The distribution of mean velocities in the object is rather irregular. The velocities are, on average, perhaps 10 km/s lower in its southern tip than in the northern, which may constitute a marginal detection of rotation in the object. The detailed, small amplitude variations in mean velocity are, however, too chaotic to allow construction of a meaningful contour diagram.

A very crude estimate of the total mass of this object can be made from the width of its global spectrum, which is plotted as Figure 3-13. This figure was prepared from data which have not been Hanning smoothed. For the present purpose of determining the mass, Hanning smoothing has

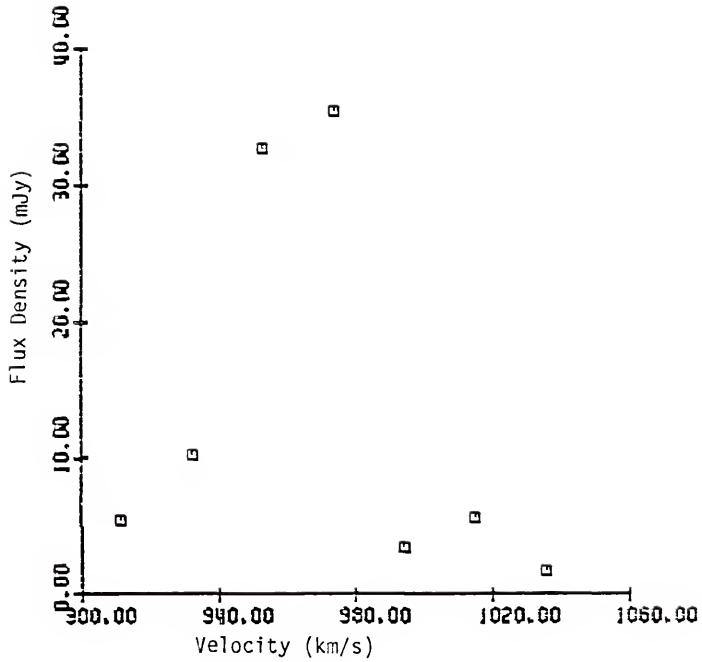


Figure 3-13. The observed global spectrum of H I in the companion galaxy to NGC 3359. No correction has been applied for the 25.3 km/s instrumental bandpass. The data from which this profile was generated have not been Hanning smoothed.

the advantage of increasing the signal-to-noise ratio, but the disadvantage of increasing the instrumental contribution to the profile width. The analysis was, therefore, carried out both ways. The procedure is to fit a Gaussian curve to the part of the global profile which is significantly above the noise. The width of this fit is then corrected for the instrumental response (which, of course, is broader in the case of the Hanning smoothed data), giving the dispersion of the intrinsic profile, which is also assumed to be Gaussian. The half-width, at the 20% level of this intrinsic Gaussian profile is taken to be the indicative velocity V_i of the system. No correction for inclination is attempted. The mass estimate is

$$M = \frac{V_i^2 R}{G} , \qquad (3\text{-}12)$$

where R is the estimated radius of the object. Clearly, this method will give only a rough idea of the mass of the system, but we lack the velocity resolution, the spatial resolution, and the sensitivity to justify a more elaborate analysis.

The results of the Gaussian fitting do not differ significantly in the smoothed and unsmoothed cases. The values of the intrinsic dispersion of the spectrum, and the mean velocity, are 11.4 km/s and 961 km/s, respectively, from the unsmoothed data, and 11.2 km/s and 963 km/s from the Hanning smoothed data. An element of caution in accepting these results is in order, however, as the widths, even of the instrumentally convolved profiles, are too narrow to be determined very accurately. Taking the mean of these two determinations, the intrinsic half-width at

20% is 20.3 km/s. The mean velocity of the cloud is 47 km/s in approach, relative to the systemic velocity of NGC 3359. Assuming the satellite galaxy to be bound to NGc 3359, the projected velocity and separation give a lower limit to the mass of the latter. Computing this limit for a circular orbit gives a mass of 2.5×10^{1}xy0 $M_{\odot}$, well below the mass inferred from the rotation curve. Therefore, it is probable that the satellite is indeed bound.

For the purpose of estimating the radius of this object, we consider only the main cloud. After deconvolving the observed map with the synthesized beam, its radius is about 0.5, or 1.6 kpc. The mass is then calculated, from equation (3-12), to be 1.5×10^{8} $M_{\odot}$. This value should be regarded as an order of magnitude estimate only.

Recalling that the mass of NGC 3359, within a radius of 21.6 kpc, is about 1.2×10^{11} $M_{\odot}$, the isolated cloud clearly cannot be expected to exert a significant influence upon the dynamical evolution of the larger galaxy. Even if our value of its mass should be an underestimate by a full order of magnitude, which is unlikely, the mass ratio is more than sufficiently large to ensure that the gravitational attraction of this outlying object is negligible at a projected distance of almost 50 kpc.

On account of the paucity of reliable data, a lengthy discussion of the physical nature of this object is not in order at the present time. However, we offer the following brief remarks. On the basis of its proximity to NGC 3359 in angular position and in velocity, we suggest that the object is a dwarf satellite galaxy to NGC 3359, probably of irregular type. The impression from the Sky Survey print is of some kind of barred structure, but the image is both so small and so faint that its

morphology is uncertain. The ratio of H I mass to total mass which we obtain is 0.31, although the estimate of the latter is admittedly crude. It seems certain, at least, that the satellite galaxy is fairly rich in H I for its mass. We have no estimate for its magnitude, as its existence has been, to the best of our knowledge, unknown until now. It is not listed in the Master List of Nonstellar Optical Astronomical Objects (Dixon and Sonneborn, 1980), for example. However, its ratio of hydrogen mass to blue luminosity would seem to be quite high, judging by the difficulty of seeing it at all on the Sky Survey. At the same time, it is a decidedly blue object, from a comparison of the blue and red images. We speculate that this may be a galaxy in which the efficiency of star formation has remained low throughout most of its history.

CHAPTER IV
THE STELLAR BAR OF NGC 3359

Surface Photometry: Introduction

In the preceding chapters, we have presented observations of neutral hydrogen in NGC 3359. From the calibrated, single-channel H I maps, we have derived maps of the integrated column density and mean velocity across the galaxy, suitable for comparison with surface densities and velocities generated by hydrodynamical models. Before proceeding to discuss those models, however, we wish to introduce one more type of observation, near infrared surface photometry of the galaxy's bar, which will be very helpful in modeling the dynamics of the gas in NGC 3359.

The neutral hydrogen data give information about the response of the gas to the imposed gravitational field of the galaxy, and the aim of our dynamical modeling is to produce a similar response of the model "gas" to the imposed potential of the model galaxy. However, a common problem in work of this kind is that the uniqueness of the results sometimes can be in question. That is, even if one is successful in reproducing what one feels to be the essential behavior of the system in a given model, it does not follow that the model is the only one that would give this outcome. In the present context, the fundamental input parameters of our model are, broadly speaking, the mass distributions of the axisymmetric and nonaxisymmetric components of the galaxy, the angular velocity of the latter, and, though this is often overlooked in

numerical modeling, the computer code itself: we are assuming that the physics of the gas in NGC 3359 is represented adequately by the hydrodynamics of the beam scheme when the gas moves in a given external gravitational field. The numerical viscosity of the beam scheme may cause some misrepresentation of the strengths of shocks in the gas, but it will certainly indicate their presence. All of the hydrodynamical codes used to date for modeling barred spirals have some difficulty in calculating the precise details of shocks, owing to the necessarily somewhat coarse spatial resolution. However, as we mentioned in Chapter I, all of these codes, including the beam scheme, give similar results when convolved to the still coarser resolutions of our H I data. The point of real interest in our investigation is the determination of those "input parameters" in the physical galaxy, and the inference we hope to draw is that they are the same as those we have used in some model galaxy which we judge to be successful. That inference must be made with considerable caution, however. Only by an exhaustive search in "parameter space," that is, by trying every conceivable combination of the input parameters, could we draw it in anything approaching complete safety. In practice, we are able to explore only a limited volume of that space. We will restrict ourselves entirely to the beam scheme program, for example, because of the development effort required in creating reliable hydrodynamical computer codes. Additionally, there are certainly gravitating configurations whose potentials we do not know about, or do not know how to write. Above all, there is no assurance, obviously, that all pertinent "input parameters" of the real galaxy are even included among those which span our parameter space.

While the uniqueness problem, therefore, can never be solved completely, there is a way to ameliorate it: add more data. The more observational constraints we can require models to satisfy, the smaller the chance that a given model happens to resemble the real system in those properties we have observed, but not in some others. It is especially helpful to add data of a radically different nature from what one has on hand already, particularly if they provide a direct measure of one of the fundamental physical parameters. Near infrared surface photometry of the stellar bar meets these criteria admirably in the present instance.

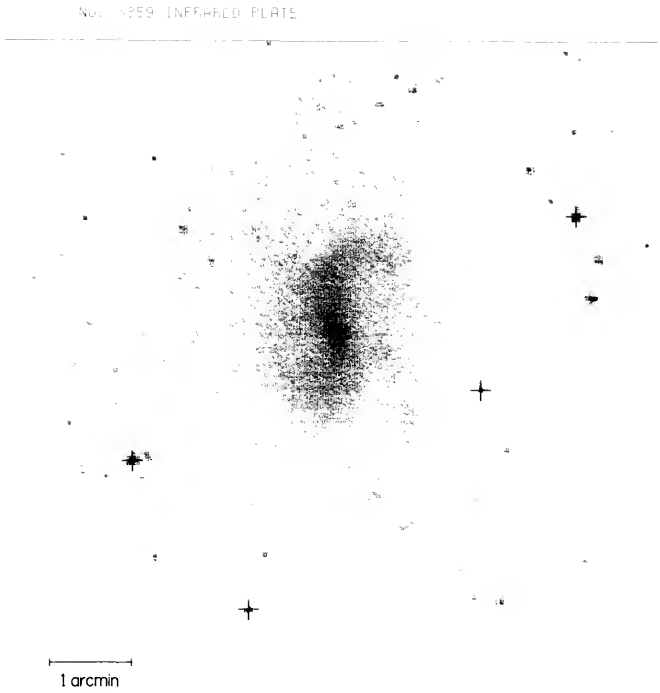
Since the models we use simply follow the response of the gas to an imposed gravitational field, which presumably is dominated by the galaxy's stellar component, information on the mass distribution of the latter is of central importance. The kinematics of the gas give indirect evidence of this distribution, through the rotation curve, but an observation of the surface brightness of the light from the stars provides direct evidence. In an extremely naive view, it might even be thought that galaxian dynamics could be predicted largely on the basis of the mass distributions, as inferred from surface photometry of the stellar components. There are, unfortunately, insuperable obstacles to such a program. The most important of these are (1) mass-to-light ratio (M/L) variations, due to changing stellar populations within a galaxy; (2) internal obscuration by dust; (3) the difficulty of trying to extract three-dimensional shapes from observations of two-dimensional images; and (4) the evident presence of large quantities of dark matter, whose relative contribution to both the mass density and the gravitational force may be very significant, especially in the outer parts of galaxies.

These complications prevent the determination of the global mass distributions of galaxies from their light distributions.

However, data of this type can be much more useful in the investigation of a smaller, more homogeneous portion of a galaxy. In the study of barred spiral galaxies, for example, photometric data which give information about the mass distribution of the stellar bar would obviously be an important tool. Dr. D. M. Elmegreen has generously provided us with a digitized version of her plate of NGC 3359 in the I passband, an analysis of which is particularly valuable for our purposes. This plate has been published previously by Elmegreen (1981). The I band has an effective wavelength of 825 nm, and a shortward cutoff of 730 nm, so that light of the hydrogen alpha line is excluded. The image of the galaxy at this wavelength should be dominated by red stars, and one might expect the bar to be especially prominent. In fact, that is exactly what is seen in Figure 4-1, a grey tone representation of the plate. Note the remarkable smoothness of the image, especially in the inner parts of the galaxy. The dominance of the bar, and of the inner galaxy in general, accords well with the results of Blackman (1983) for the barred spiral NGC 7479. He found that the inner components of that galaxy became more prominent as one moved from the U filter toward redder colors.

These data are affected less by the complicating effects noted above than most surface photometry. The stellar population of the bar is expected to be relatively well-mixed (Contopoulos, 1983). This will diminish the importance of point (1). The long wavelength makes point (2) somewhat less worrisome than in, say, a blue image. Little can be done about item (3), as we have stated, but the image structure at 825 nm does turn out to be reasonably simple, leading us to hope that we may be

Figure 4-1. The I passband image of NGC 3359, from a digitized plate made available by Dr. D. M. Elmegreen. Intensity levels of this figure have been adjusted to increase the visibility of the components outside the bar. The original plate was exposed for two hours on the Palomar Schmidt telescope. Crosses indicate the positions of the foreground stars used as fiducial marks in this dissertation (see Figure 1-1). Here north is at the top, east to the left.


NGC 4559 INFRARED PLATE
1 arcmin

able to fit it by a simple function, without doing too much violence to the actual shape of the physical bar. Lastly, the proportion of the total density which is in the form of dark matter is relatively low in the inner parts of spiral galaxies, if current ideas are correct, and so if we confine our attention to the bar itself, we should be able to neglect that influence. In practice, only relative calibration is available for the plate, owing to the lack of photoelectric work in the I band. Therefore, our "mass-to-light ratio" for the infrared isophotes contains an arbitrary scaling factor, and so the effects of any dark matter in the bar region will, to first order, appear only in this factor. We simply assume that M/L within the bar itself is constant.

The near infrared surface photometry of the bar thus provides a good measure of the unobscured surface brightness of the bar, which in turn can be used to give an estimate of the mass density within the bar. This is of great importance in helping to deal with the uniqueness problem, as it restricts rather severely the multitude of plausible bars that might be used in the hydrodynamical models of the galaxy. Before we can proceed to extract the mass distribution, however, we must make some assumptions about how the surface brightness and the volume mass density in the bar are related. Our first choice is to assume that the mass-to-light ratio within the bar is constant, as we have said. We must also find a way to pass from the two-dimensional surface brightness distribution to a volume brightness distribution. As we have emphasized above, there is no unique solution to this problem, and in order to proceed we must assume some relatively simple three-dimensional shape for the bar. The most general form of a three-dimensional, elongated object which is analytically tractable is the triaxial ellipsoid. This form is often

speculated to be the three-dimensional shape of the bars in barred galaxies, and, we might note in passing, of elliptical galaxies as well (see references in Ball and Hunter, 1984).

Projection Effects in Triaxial Figures

The isophotal projection of a triaxial figure has been considered by Stark (1977), who extended the techniques of Lindblad (1956). Specifically, Stark works out the case in which the volume brightness is constant on similar, concentric, triaxial ellipsoidal surfaces. He demonstrates that, no matter what the viewing angles, such a figure will have isophotes which consist of a set of similar, concentric ellipses, and that the converse relation also holds. We will model the bar of NGC 3359 as a triaxial figure having nested, elliptical isophotes, although there are, in practice, some departures from this ideal. This is a physically plausible figure, and it is the most complex for which the projection problem has been solved. We will comment below on the consequences of this idealization.

In Stark's formalism, the orientation of the triaxial figure being observed is given in terms of four parameters: the axial ratio of the isophotes, β, and the three Euler angles, which relate the observer's coordinates to those in which the equations of the ellipsoidal surfaces have the form

$$(tx)^2 + (uy)^2 + z^2 = a_v^2. \qquad (4\text{-}1)$$

Here a_v parameterizes the volume brightness. Two of the three angles can be specified for the case of a barred spiral galaxy, if we assume that the triaxial surfaces have one axis normal to the disk of the

galaxy. Then the Euler angles become θ, the complement of the usual inclination angle i of the galaxy; ψ, which is the angle in the sky plane between the major axis of the isophotes and the line of nodes of the disk; and Φ, which is the one remaining unknown. In physical terms, this is the angle, also in the sky plane, between the line of nodes and the projected direction of the shorter of the two axes which lie in the plane of the disk. Since the true axes of the ellipsoids are not in the same directions as the axes of the isophotes (see Figure 4-2 for an illustration of the geometry), Φ has a finite range of allowed values.

The possible solutions for the three-dimensional brightness distribution, therefore, form a one-parameter family with Φ as the controlling variable. Each member of this family is characterized by t, u, and α by t, u, and α. The latter, a dimensionless scale factor, gives for each triaxial surface the length of the semi-axis which is perpendicular to the disk, in units of the apparent semi-major axis of the corresponding isophote. From Stark's equations (9) and (12), one can solve for t, u, and α in terms of the projection parameters β, θ, ψ, and Φ. The solutions are:

$$u^2 = -(d_8 d_{12}+d_9)(d_5 d_{12})^{-1}, \tag{4-2}$$

$$t^2 = d_{12} u^2, \tag{4-3}$$

and

$$\alpha^2 = d_1 t^2 u^2 [\beta^2 (d_2 t^2 + d_3 u^2)]^{-1}, \tag{4-4}$$

where:

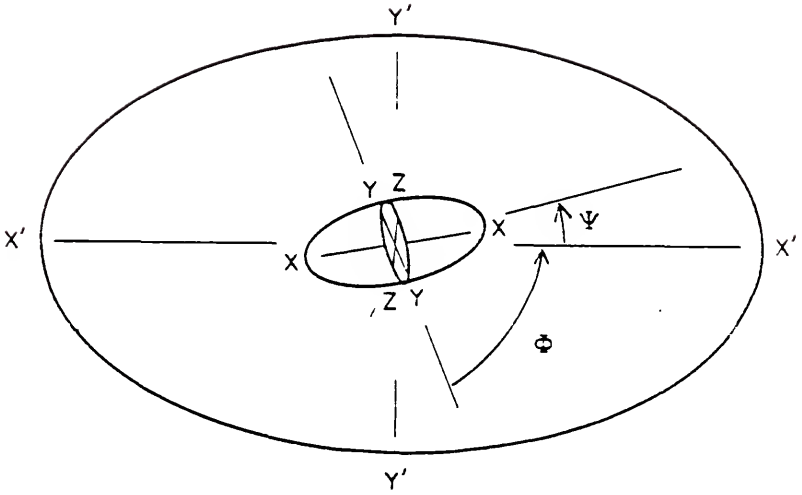


Figure 4-2. Geometry of a triaxial bar in a disk galaxy. Here, the plane of the paper represents the plane of the sky. The outer ellipse, with axes x', y', is the galaxian disk, in projection. The outermost bar isophote is shown as the ellipse of intermediate size. The three principal axes of the bar are labeled x, y, and z; the innermost ellipse is the (y,z) principal plane of the bar. By hypothesis, x and y are in the same plane as x' and y'. The two Euler angles Ψ and Φ, both measured on the plane of the sky, are indicated. Note that as the bar becomes more elongated, the x axis becomes closer to the position indicated by Ψ, and the (y,z) plane is seen more nearly edge-on.

$$d_1 = \sin^2\psi + \beta^2\cos^2\psi, \tag{4-5a}$$

$$d_2 = \sin^2\Phi, \tag{4-5b}$$

$$d_3 = \cos^2\Phi, \tag{4-5c}$$

$$d_4 = \cos^2\psi + \beta^2\sin^2\psi, \tag{4-5d}$$

$$d_5 = \sin^2\theta, \tag{4-5e}$$

$$d_5 = \cos^2\Phi\cos^2\theta, \tag{4-5f}$$

$$d_7 = \sin^2\Phi\cos^2\theta, \tag{4-5g}$$

$$d_8 = d_6 - d_2 d_4 d_1^{-1}, \tag{4-5h}$$

$$d_9 = d_7 - d_3 d_4 d_1^{-1}, \tag{4-5i}$$

$$d_{10} = \sin\psi\cos\psi(1-\beta^2), \tag{4-5j}$$

$$d_{11} = \sin\psi\cos\psi\cos\theta, \tag{4-5k}$$

and

$$d_{12} = (d_1 d_{11} - d_3 d_{10})(d_1 d_{11} + d_2 d_{10})^{-1}. \tag{4-5l}$$

The limiting values of Φ occur where u vanishes, and where t goes to infinity. We can understand these limits physically by noting that, if we call the semi-axis perpendicular to the galaxian disk c, and the longer and shorter of the other two axes a and b, respectively, then from equations (4-1) through (4-3),

$$\frac{b}{a} = \frac{u}{t} = d_{12}^{-1/2},$$

(4-6a)

and

$$\frac{c}{a} = u.$$

(4-6b)

Thus, if we assume that a is always the longest of the three axes (as is appropriate when dealing with barred spirals), the two limits correspond to the cases where the other two axes vanish. Accordingly, the allowed range in Φ encompasses all possible ratios of these two shorter axes, and the bar can be prolate, perfectly flattened either in the plane of the galaxy or at right angles to this plane, or any intermediate form. This result, of course, takes only the photometry itself into account. A stellar bar embedded in a flat disk will itself be flattened, in the same direction, by the gravitational attraction of the disk, and so we may reject models for which $c/b > 1$ on physical grounds.

The upper limit on Φ can be found by setting the denominator of d_{12} to zero. This yields a simple, closed form for the maximum value of Φ, in terms of β, θ, and ψ,

$$\tan \Phi_{max} = \frac{\cos\theta(\sin^2\psi + \beta^2\cos^2\psi)}{(\beta^2-1)\sin\psi\cos\psi}.$$

(4-7)

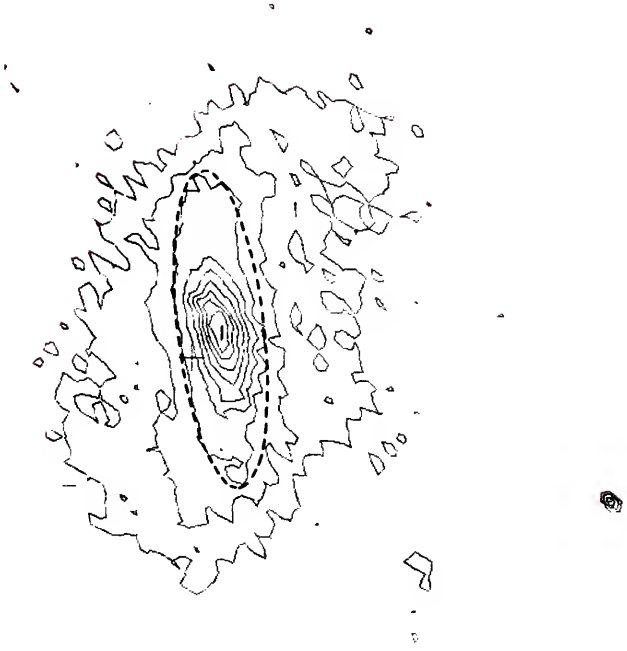
The lower limit occurs where $d_\beta d_{12} = -d_\beta$. This is most easily found iteratively. In general, the larger the value of β, the more narrowly Φ is constrained. This is suggested by Figure 4-2, and is

readily explained by a consideration of the geometry of the situation. The uncertainty in Φ arises because, for inclined triaxial figures with three finite axes, the longest axis of the ellipsoid need not point in the same direction as that of its two-dimensional projection. However, as the figure becomes more and more elongated in one direction, thus growing to resemble more and more closely a rod of infinitesimal thickness, this discrepancy becomes smaller. In the opposite limit, the sphere, the direction of the "longest axis" obviously becomes undefined. Since β exceeds unity by several times in most "strongly barred" spirals, the range of Φ is typically only a few degrees. Nevertheless, the arguments given above remain valid, and the ratio of the two shorter axes is still indeterminate.

In light of this discussion, let us now examine the bar of NGC 3359 more closely. Figure 4-3 is another representation of Elmegreen's I plate, this time displayed as a contour map. The contour levels here are evenly spaced, but are, of necessity, arbitrary, since the plate itself has only relative calibration; as there is no magnitude calibration available for the galaxy in this passband, one can only measure the intensity in arbitrary units. The bar is clearly the dominant feature in this image. Fortunately for our purposes, there is little apparent dust to interfere with an analysis of the bar brightness, as happens even at this wavelength in the barred spiral NGC 3992, for instance (Hunter et al., 1984).

Before discussing the application of Stark's method to these data, let us comment briefly on the relevance of the method itself to real galaxies. Obviously, the depiction of the isophotes as nested ellipses is a simplification. The observed isophotes will deviate from this ideal

Figure 4-3. Contour plot representation of Elmegreen's near infrared plate of NGC 3359. The contour levels are spaced at equal intervals, of one-tenth the maximum brightness. Overlaid on these contours is a member of the family of nested ellipses which is used to model the brightness distribution of the bar (dashed line). The small cross locates the center of the digitized image. One of the foreground stars used as fiducial marks may be seen to the southwest of the galaxy; see Figure 4-1 for comparison.



for three reasons: the departure of the true form of the bar from a perfect triaxial ellipsoid, confusion with other galaxian components (spiral arms, etc.), and observational noise. However, the model of the bar volume brightness which this technique allows one to derive has a crucial importance to the dynamics of the system, which is not vitiated by these objections. First of all, the model is a reasonably good representation of the data, as we shall see. Secondly, it fulfills its primary purpose well: though the model may not give a perfect fit to the photometric data, most model bars which one could use in the hydrodynamical modeling would give a far worse fit. The model bar derived from Stark's method, therefore, provides a valuable means of discriminating among mass distributions which would otherwise be only weakly constrained, and so is quite useful in dealing with the nonuniqueness problem. The final justification for this approach is the nature of the physical problem under consideration. Since the force associated with an extended mass distribution is given by an integral equation, it is not very dependent on the fine details of that distribution. Our concern is with the global behavior of the galaxy's gaseous component in the field of the bar and underlying disk, and local irregularities in that field are beyond the scope of this study. For these reasons, important information on the dynamics of the galaxy can be obtained from Stark's method, despite the minor difficulties of fitting nested ellipses to the isophotes.

The first of these difficulties is that, as one would expect with unsmoothed data, the isophotes are not perfect ellipses. Besides small, random departures which are within the noise, the ends of the isophotes tend to be somewhat broader than for true ellipses, especially in the

outer parts of the bar. Those portions of the bar probably have some contribution from the spiral arms, however, and the effect is not a major one. Secondly, the axial ratio and position angle of the best fit ellipse change slightly as one goes outward in radius. A similar effect is widely observed in the isophotes of elliptical galaxies (see Mihalas and Binney, 1981, p. 331, for a discussion, and the surveys of Strom and Strom, 1978a, b, c, and of Bertola and Galletta, 1979, for examples). The usual interpretation of the phenomenon in that context is a change in true axial ratio from the center of the system outwards. In the case of barred spirals, it may be more conventional to consider this as evidence for the presence of two separate components, one being the bar proper, and the second, inner component some sort of nuclear core. Analysis of the isophotes under that interpretation, however, is to be avoided, as it simply adds too many free parameters. As we shall see below, the mass of the bar in NGC 3359 resides largely in its outer regions, and those are also the more important dynamically, as they provide most of the nonaxisymmetric stirring force which provokes a spiral response in the gaseous disk. Therefore, it is much more important to provide a good fit to the outer than to the inner bar isophotes. Effects like the isophotal twist and axial ratio change, being confined mostly to the inner bar, will not be treated in our fitting procedure. The magnitudes of both of these effects are rather uncertain, owing to resolution effects, namely seeing and finite pixel spacing. The pixel size used in scanning the plate was approximately 3.″37. Since the seeing disk on the original plate was somewhat smaller than this, but not many times smaller, there may be some loss of resolution, especially near the center of the galaxy where the gradient is steepest. The same regions are affected, beyond question,

by the finite pixel size: on the bar minor axis, the spacing between adjacent contour levels in Figure 4-3 is nowhere more than a single pixel width. The spacings indicated between these levels cannot be regarded as real in these circumstances. Thus, the actual change in axial ratio is probably less than a naive measurement of Figure 4-3 would imply. The apparent value of this ratio varies from approximately 2.1 to about 3.7 as one goes from the center to the outer bar, but a crude, first-order deconvolution suggests that the true ratio at the center is likely to be in the range 2.6 to 3.0, depending on the size of the seeing disk on the original plate. The apparent twisting of the position angle of the major axis amounts to about 10°. This effect, considered as a qualitative description, may be real, but its true magnitude is uncertain, and in any case is rather small. In practically every case, adjustment of the contour by a single pixel would line up each isophote with the major axis of the outer bar. We neglect these indications that the bar may take on a bit more of the character of a nuclear bulge at its center, for the reasons cited above, and because the very distinct elongation of even the innermost isophotes indicates that the change is not a clear-cut one. However, it should be noted for the record that the bar may be a somewhat more complex entity than represented by our model. Fortunately, our interest is primarily in the dynamical effects of the bar, not in its internal structure. Therefore, we will simply fit ellipses as well as we can to the outermost bar isophotes and accept the imperfection of the model. This plan allows us to proceed directly from our fit to a determination of the bar forces, which are, after all, our real concern.

The dashed curve in Figure 4-3 indicates the ellipse which has been fit to the bar isophotes. Only a single contour level, near the outer

limit of the bar, is shown; the other contours would, of course, be ellipses which are similar to this one and concentric with it. The parameters that specify the fit are x_c and y_c, the Cartesian coordinates of the center of the system of ellipses: , the position angle of the major axes, in the (x,y) coordinate system; and β, the apparent axial ratio. Since precise positional data are not available for the digitized image, the location (x_c,y_c) is of no intrinsic interest. It should be noted, though, that the centers of the isophotes do vary to some extent. The adopted value of (x_c,y_c) is at the center of the outer bar isophote, and also lies almost exactly on the central brightness peak, but the centers of some of the intermediate isophotes are displaced slightly to the south of this position. The values of β and η for the adopted ellipse are 3.7 and 7°, respectively, where η is specified in the normal astronomical sense.

We can now explore the allowed range of values of the geometrical parameters, t, u, and α, of the triaxial bar corresponding to our isophotal fit. In Chapter III, we adopted 51° and -8° as the inclination angle and position angle of the line of nodes, respectively, for the disk of NGC 3359. This inclination implies that the Euler angle θ is 39°. Combining the line of nodes angle with η yields 15° for the second Euler angle, ψ. The remaining Euler angle Φ is restricted, by the adopted values of θ, ψ , and axial ratio β, to the range 66°.28 to 68°.96. Although it would be absurd to consider the limiting values of Φ to be determined to an accuracy like those quoted here, the magnitude of the range in Φ is of interest; its narrowness is on account of the large value of β, as discussed above. Despite Φ being quite well determined, the ratio of the two short axes is, of course, unknown. The case Φ = 67°.3

is prolate; lower values of Φ give bars which are flattened in the direction normal to the disk component. Table 4-1 lists the quantities of interest for several assumed values of Φ.

Brightness Distributions

Before using the geometrical fit given above in dynamical models, we must specify the three-dimensional mass density within the bar as a function of semi-major axis. We will, as we have said, assume a constant ratio of mass to luminosity in the bar. It remains to determine the run of luminosity per unit volume, in arbitrary units, from the photometry. Stark has shown that, for the triaxial geometry under discussion, the volume brightness F_V and surface brightness F_S are related by

$$F_V(a_V) = -\frac{f^{1/2}}{\pi} \int_{a_V}^{\infty} \left[\frac{d}{da_S} F_S(a_S)\right](a_S^2-a_V^2)^{-1/2} da_S, \tag{4-8}$$

where

$$f = t^2\sin^2\theta\sin^2\Phi + u^2\sin^2\theta\cos^2\Phi + \cos^2\theta \tag{4-9}$$

is independent of position for a given bar and projection. In order to make use of equation (4-8), it is helpful to have an expression for F_S which is simple enough so that the integral in (4-8) can be evaluated analytically. At the same time, F_S must be a reasonable fit to the observed surface brightness at radii from the center of the bar to its greatest extent, and must go to zero at infinity. With these requirements in mind, let us examine the observed brightness distribution.

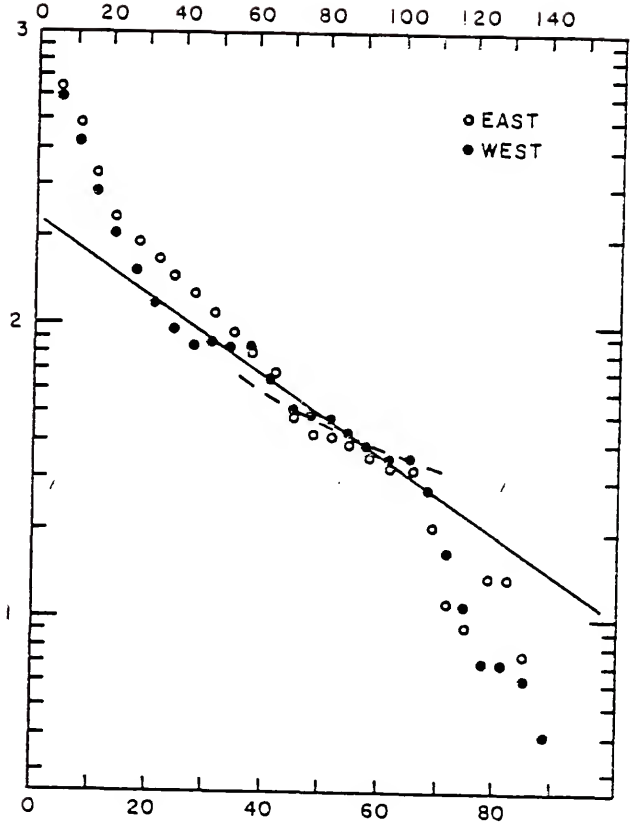
The first step in deriving F_S is to remove any contribution of disk light from the bar isophotes. To this end, we have analyzed the

TABLE 4-1
BAR PROJECTION PARAMETERS

Φ	u^2	t^2	α^2	b/a	c/a
66°5	0.0141	0.142	0.0164	0.31	0.12
67.0	0.0476	0.589	0.0554	0.28	0.22
67.5	0.0824	1.336	0.0957	0.25	0.29
68.0	0.1189	2.852	0.1374	0.20	0.34
68.5	0.1567	7.624	0.1806	0.14	0.40

intensity profile along the disk minor axis of NGC 3359. Since the bar is aligned only 15° from the disk major axis, the minor axis profile is useful for determining the brightness of the disk alone. The data were first smoothed by convolution with a Gaussian beam of dispersion σ_{bm} = 2 pixels or 6".7, in order to improve the signal-to-noise ratio in the rather faint disk. Figure 4-4 shows the profiles along the two halves of the minor axis, on a logarithmic intensity scale; the intensity units are arbitrary. There are several noteworthy features in this plot. First, there is a notable lack of symmetry, indicated by the divergence of the two curves, in the region of transition from bar to disk around 25". As one moves away from the bar in the direction of the more dominant, western arm, the contribution of light from the bar at first drops more quickly than in the opposite direction. It is not clear what significance this might have, although it is in keeping with the departures from detailed symmetry of the two halves of the spiral pattern, which are observed repeatedly in this galaxy. For the purposes of trying to estimate the contribution of a background disk, the portion of the profile which may be affected by bar light should be rejected. We therefore exclude all data interior to a distance of 35" from this analysis. At approximately twice that radius, there is another abrupt change in slope, as the surface brightness begins to fall off very rapidly. We take 65" to be the outer limit of the region used for estimation of the disk intensity. Between these two limits, the data can be represented adequately by an exponential function, although admittedly, with fewer than ten independent points on each half of the profile, the functional form of the best fitting curve is rather open to question. It could be argued that the range of data should be restricted still further, because the lower end of the selected segment of the intensity profile,

Figure 4-4. Near infrared brightness profiles, perpendicular to the major axis of the disk of NGC 3359. Values taken from the regions east and west of the galaxy center are shown as open and filled circles, respectively. The intensity scale is logarithmic, in arbitrary units. Units of the abscissa are seconds of arc. The lower axis gives the distance as measured on the sky, while the upper scale has been corrected for the inclination of the galaxy. The data lying between uncorrected radii of 35" and 65" have been fit to an exponential function by the method of least squares. This solution appears as a solid line. The dashed line is a typical fit to the data under the assumption that the surface mass density is that of a Mestel disk.


0
20
40
60
80
100
120
140
3
2
1
EAST
WEST
0
20
40
60
80

at about 60", constitutes the optical spiral arm. We retain this part of the profile, however, mainly from necessity, and proceed with fitting an exponential function to the observed profile. The result of a least squares fit to the data in the chosen region is

$$I(r) = 232.1 \exp(-r/31\overset{''}{.}7), \qquad (4\text{-}10)$$

which is shown as a solid line in Figure 4-4. When this exponential is rectified to the major axis of the disk, the exponential length scale increases from 31".7 to 51".4.

Before continuing with our determination of the bar brightness distribution, we offer a few more remarks on Figure 4-4. Though the present data do not go quite deep enough to warrant a detailed discussion of the surface photometry of components other than the bar, it is instructive to compare what we have with the invaluable surface photometry by Blackman (1983) of the barred spiral galaxy NGC 7479. One of his principal findings was that the outer isophotal profiles of that galaxy revealed the presence of an exponentially falling disk, with a very steep gradient. Furthermore, if one were to extrapolate this component inward all the way to the center of the galaxy, the resulting central brightness would greatly exceed the observed value. The behavior of the disk profile beyond 65" in Figure 4-4 seems to hint at similar behavior in NGC 3359. In NGC 7479, the abrupt change in slope occurred as the optical spiral arm was crossed. The behavior of the isophotes outside the bar of NGC 7479, in other words, can be summarized by saying that the arms enclose an area of high surface brightness, which decreases only gradually between the bar and the arms, while beyond the arms the intensity falls much more quickly. Though the data we have available for NGC 3359 must be regarded as suggestive

rather than conclusive, the overall picture suggested by Figure 4-4 accords very well with Blackman's results. We might note that Blackman also found the regions immediately outside the bar to be rather asymmetric in brightness level, as we have observed for the bar to disk transition area of NGC 3359. While on the subject of asymmetries, we point out that the western arm is seen, in Figure 4-4, to be only marginally brighter than its weaker, eastern counterpart. This tends to confirm the subjective judgment, based on inspection of Figures 1-1 and 4-1, that the apparent dominance of this arm owes nearly as much, perhaps, to the coherence of its structure, as to an actual brightness difference between the arms.

The dynamical meaning of these photometric results is problematic. We certainly do not interpret the exponential fit to the restricted segment of the intensity profile discussed above as being the true distribution of surface mass density in the galaxy's disk. It is intended solely as a means of estimating the proportion of light in the bar region which might be associated with the disk component. Even setting aside the obvious questions about the arbitrary nature of the fitting procedure, it would be very imprudent to assume that the distribution of light in this "inner disk" component is an accurate tracer of the mass density. Seiden, Schulman, and Elmegreen (1983), for example, argue from the stochastic, self-propagating star formation theory that seemingly exponential galaxian disks are a consequence of the interplay of the molecular component of the interstellar medium and the star formation rate, and that "the exponential fit has no intrinsic physical meaning." Their argument is particularly directed at galaxies in which the neutral hydrogen density and the rotation curve are both almost constant, as is roughly the case for NGC 3359 in the narrow range of deprojected radii between one and two arc

minutes, where our "exponential" disk is located. This is only one of several plausible scenarios in which the mass-to-light ratio of the disk might depend strongly on the radius. Against any such interpretation must be set the strongly suggestive results of van der Kruit and Searle (1981a, b; 1982a, b) that the mass-to-light ratios of galaxian disks do, in fact, remain relatively constant over large scales. If this evidence is accepted, the additional mass must reside in halos. In the case of NGC 3359, the spiral arms may well play an important role, for example by driving star formation in the inner galaxy, since the bright, optical arms are observed to demarcate the region of high surface brightness. If a mechanism involving the arms does help regulate the brightness in some manner, the arms must be quite long-lived, since the I plate used in our analysis records predominantly the light of old stars.

As we saw in Chapter III, the rotation curve of NGC 3359 can be fit reasonably well by a Toomre n=0 disk, also called the Generalized Mestel Disk or GMD (Hunter, Ball, and Gottesman, 1984), and we shall find in Chapter V that a disk of this type also seems to give a good representation of the underlying disk potential in hydrodynamical models of the system. This disk is considerably less centrally concentrated than the exponential disk. Its surface density is given by

$$\Sigma(r) = \Sigma_0 (r^2 + b_0^2)^{-1/2}, \qquad (4\text{-}11)$$

where Σ_0 and b_0 are constants. We consider here the infinite, untruncated GMD, as no truncation signature (Casertano, 1983; Hunter, Ball, and Gottesman, 1984) is visible in the rotation curve. The steepest possible gradient in the surface density for this disk, therefore, occurs when the

length scale b_0 is zero, corresponding to Mestel's (1963) original disk. In that case, the surface density is simply inversely proportional to the radius. An example of this law, fit to the portion of the intensity profile which we have designated as representing the disk, is shown as a dashed line in Figure 4-4. Clearly, the fit is significantly flatter than the data. The GMD which actually fits the rotation curve has a value of b_0 significantly larger than zero, and is much flatter in surface brightness than the example illustrated in Figure 4-4. We must conclude, therefore, either that the GMD does not accurately represent the surface density of the disk, or that the disk mass-to-light ratio in the I band increases significantly with radius, or, quite conceivably, both. The relatively good fit of the GMD to the rotation curve need not be physically meaningful if substantial matter is present in the form of a halo; as we remarked in Chapter I, one cannot distinguish between the presence of an axisymmetric disk component and a halo component from analysis of the rotation curve. From this point of view, the GMD may be nothing more than a convenient fiction for specifying the axisymmetric force in a galaxy whose rotation curve is asymptotically flat. This is the interpretation supported by the work of van der Kruit and Searle, for example. Alternatively, it could be the case that the GMD does prescribe the density, and that M/L increases radially on account of differences in star formation rate or some similar physical process. The lack of a truncation signature in the rotation curve tends to support this hypothesis. Ultimately, both explanations amount to the statement that the effective mass-to-light ratio of the galaxy as a whole increases from the center outwards. That is the strongest conclusion we can draw with any degree of safety from the photometric data beyond the bar. Data in additional colors would be

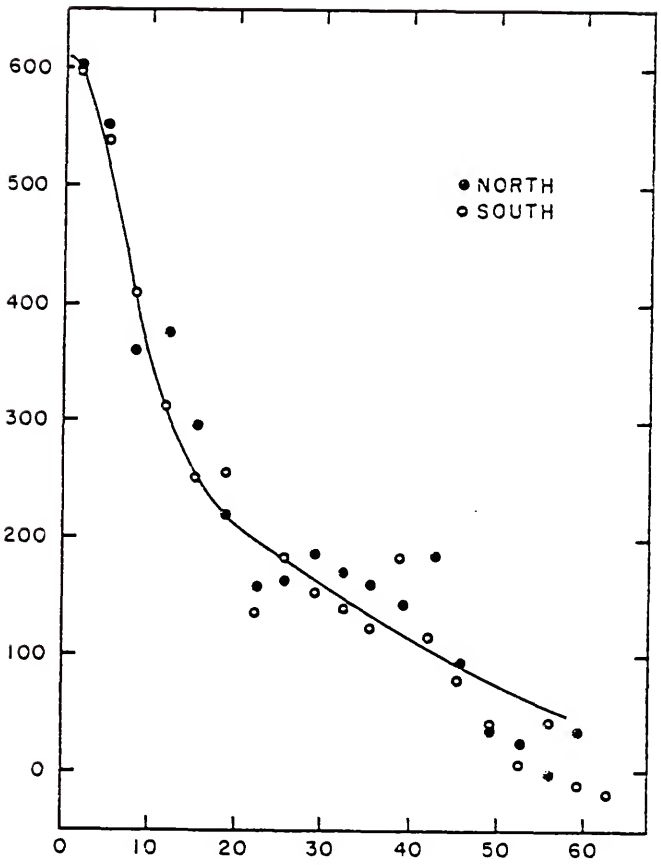
valuable in addressing the possibility that the M/L is significantly different in this near infrared image than in bolometric luminosity.

Let us now return to a discussion of the isophotes of the bar itself. We proceed by subtracting the disk brightness as estimated by the exponential fit in Figure 4-4. The radial scale of that fit was corrected for projection effects, first by rectifying it to the disk major axis, and then allowing for the orientation of the bar itself, as found in the isophotal fitting procedure described earlier. The disk major axis was taken to be that adopted in Chapter III. Figure 4-5 displays the brightness profile along the bar major axis. Unlike Figure 4-4, this plot is derived from unconvolved photometry with the full spatial resolution of the digitized plate.

The most striking feature of this plot is the abrupt flattening of the slope at a distance of about 22" from the center. This is definitely well within the apparent bar, and corresponds to the last of the crowded inner contours in Figure 4-3. Visual inspection of plates, in any of several colors, places the end of the bar at a distance of around 48" from the center, corresponding to the final plunge of the profile in Figure 4-5 into the noise. Resolution effects on a point-like nucleus are also insufficient to explain the break in slope, as the steep inner portion extends far enough that its reality is not in question. The true explanation is most probably that the character of the bar undergoes a transition to a distribution more like that of the concentrated nuclear bulges of NGC 7479 (Okamura, 1978; Blackman, 1983) and NGC 3992 (Hunter et al., 1984) near its center.

The nature of this profile makes it impossible to obtain an adequate fit to the data with any single-component model whose functional form is

Figure 4-5. Brightness profiles along the bar of NGC 3359. The linear intensity scale used here has arbitrary units. The distance scale, in arc seconds, has not been corrected for the inclination. The exponential fit to the disk brightness, indicated in Figure 4-4, has been subtracted from the raw data to give the values plotted here. Points taken from the northern and southern halves of the bar are plotted as filled and open circles, respectively. The solid curve is the nested, two-Gaussian model described in the text.


600
500
400
300
200
100
0
0
10
20
30
40
50
60
NORTH
SOUTH

simple enough to allow analytical evaluation of equation (4-8). Instead, a model was adopted which is the sum of two Gaussian profiles. In the arbitrary intensity units of Figures 4-4 and 4-5, the final model adopted has brightness I_M given by

$$I_M(r) = 365 \exp(-r^2/2\sigma_1^2) + 250 \exp(-r^2/2\sigma_2^2), \qquad (4\text{-}12)$$

with $\sigma_1 = 6''.7$ and $\sigma_2 = 32''.0$. The model is shown as the solid line in Figure 4-5. It should be understood that this model of the bar in NGC 3359 is a "two-component" model in a formal sense only; we assume that the axial ratios and mass-to-light ratios of the two components are identical, so that physically there is only one component, with an apparent surface brightness specified by equation (4-12).

The volume brightness corresponding to a Gaussian surface brightness is also Gaussian. To show this, we let F_S in equation (4-8) be Gaussian:

$$F_V(a_V) = -\frac{f^{1/2}}{\pi} \int_{a_V}^{\infty} \left[\frac{d}{da_S} I_0 \exp(-a_S^2/2\sigma^2)\right](a_S^2 - a_V^2)^{-1/2} da_S. \qquad (4\text{-}13)$$

Differentiation and change of variable to $w = a_S^2$ give

$$F_V(a_V) = \left(\frac{I_0 f^{1/2}}{2\pi\sigma^2}\right) \int_{a_V^2}^{\infty} (w - a_V^2)\exp(-w/2\sigma^2)dw, \qquad (4\text{-}14)$$

and this may be evaluated using formula 3.382.2 of Gradshteyn and Ryzhik (1980):

$$F_V(a_V) = \left(\frac{I_0 f}{2\pi\sigma^2}^{1/2}\right)[(2\sigma^2)^{1/2} \exp(-a_V^2/2\sigma^2)\,\Gamma(\tfrac{1}{2})],$$

or

$$F_V(a_V) = \left(\frac{I_0^2 f}{2\pi\sigma^2}\right)^{1/2} \exp(-a_V^2/2\sigma^2).$$

(4-15)

With the assumptions of constant M/L and similar ellipsoidal shapes, we can compare the masses of the two Gaussian distributions. The volume of a triaxial body of semi-axes a, b, c is

$$\tau = \frac{4}{3}\pi a\, b\, c.$$

(4-16)

If b and c are considered to be $b = k_b a$ and $c = k_c a$, with k_b and k_c constant, the incremental volume of a thin, triaxial shell of thickness da becomes

$$d\tau = \frac{\partial \tau}{\partial a}\, da = 4\pi a^2 k_b k_c\, da.$$

(4-17)

Hence, the mass of the entire ellipsoid whose density is $\rho = \rho(a)$ is

$$M = 4\pi k_b k_c \int_0^{\infty} a^2 \rho(a)\, da.$$

(4-18)

Let us write $\rho(a) = \rho_0 \exp(-a^2/2\sigma^2)$ for the Gaussian mass density. Evaluation by formula 3.461.2 of Gradshteyn and Ryzhik leads to

$$M = (8\pi^3)^{1/2} \rho_0 k_b k_c \sigma^3. \qquad (4\text{-}19)$$

Equation (4-19) allows us to compare the masses of the two components of equation (4-12). If the mass of the inner component is defined as unity, that of the outer component is then 73.4. Evidently, the outer Gaussian is far more important to dynamics of the galaxy. This conclusion is strengthened further by the distribution of the two: to material in the outer regions of the galaxy, the inner component will look much more like a point mass at the galaxy center. Thus, the driving perturbation responsible for the spiral structure, if that structure is in fact a response to the bar, is due almost entirely to the outer of these two components. This result is the principal reason why we have not gone to great lengths, in the ellipse fitting procedure, to represent the inner Gaussian component with a high degree of accuracy.

Having arrived at an estimate of the mass density in the bar of NGC 3359, under the assumption that the bar is a triaxial figure, we are now in a position to integrate this mass distribution into hydrodynamical models of the galaxy. This approach should be quite valuable in reducing the uncertainty associated with the nonuniqueness of successful models. The importance of the use of the photometric data is that they give an independent, direct measurement of one of the three fundamental input parameters of our conceptual model of a barred spiral galaxy. To reiterate a

point made early in this Chapter, those parameters are the mass distributions of the axisymmetric and nonaxisymmetric components of the galaxy, and the angular velocity of the latter. The rotation curve gives observational evidence on the first of these quantities, and the photometry on the second, leaving the pattern speed of the bar as the only major physical parameter for which there is no direct observational estimate.

In order to use the surface photometry results in the hydrodynamical models, we must have a way of calculating the force which arises from the observed mass distribution. This could be done simply by direct numerical integration, but at considerable computational expense. We have discovered simple expressions for the potentials and forces of triaxial, ellipsoidal shells, or homoeoids. The expressions for the forces, in particular, can be evaluated exactly, without resort to elliptic integrals or numerical differentiation. Therefore, these results are very well suited to the problem of calculating forces due to a body whose equidensity surfaces are triaxial. Besides the application to barred spiral galaxies, the new expressions may be of interest in studies of elliptical galaxies as well (Ball and Hunter, 1984). The remaining section of this chapter develops these results in some detail. In Chapter V, we will use them to incorporate the triaxial bar we have derived from the photometry into hydrodynamical models of NGC 3359.

The Attraction of a Triaxial Homoeoid

We wish to find the gravitational force of a shell of material bounded by two concentric, similar, triaxial ellipsoids. Let the semi-axes of the inner ellipsoid be (a,b,c), where $a > b > c$, in the

directions of the Cartesian coordinates (x,y,z), respectively. The outer ellipsoid is considered to have greatest semi-axis (a+da), and axial ratios equal to those of the inner figure. The foundation of our approach is laid out by MacMillan (1958). In his Section 11, he demonstrates that the force on a particle in the interior of this homoeoid is zero. We shall consider, therefore, the solution at a point (x,y,z) exterior to the homoeoid. We use the following two additional results from MacMillan, which he derives using Ivory's method. First, the potential of the solid, homogeneous body which is bounded by our inner ellipsoid is

$$V_0 = \pi G\rho abc \int_{\kappa}^{\infty} \left(1 - \frac{x^2}{a^2+s} - \frac{y^2}{b^2+s} - \frac{z^2}{c^2+s}\right) x$$

$$x \left[(a^2+s)(b^2+s)(c^2+s)\right]^{-1/2} ds, \qquad (4\text{-}20)$$

where ρ is the mass density, and κ is the algebraically largest root of the equation

$$\frac{x^2}{a^2+\kappa} + \frac{y^2}{b^2+\kappa} + \frac{z^2}{c^2+\kappa} = 1. \qquad (4\text{-}21)$$

The largest root of this equation is real and positive (MacMillan, p. 52). Second, in taking partial derivatives of V_0, the terms containing partial derivatives of κ vanish.

The incremental potential dV associated with our homoeoid is

$$dV = \frac{\partial V_\Omega}{\partial a}\, da. \qquad (4\text{-}22)$$

Noting that $b = k_b a$ and $c = k_c a$, where k_b, k_c are constants, the differentiation gives

$$\begin{aligned} dV = \frac{3V_\Omega}{a}\, da + \pi G\rho a^3 k_b k_c \{ & -a\int_\kappa^\infty (a^2+s)^{-3/2}(b^2+s)^{-1/2}(c^2+s)^{-1/2}ds - \\ & -k_b^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-3/2}(c^2+s)^{-1/2}ds - \\ & -k_c^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-1/2}(c^2+s)^{-3/2}ds - \\ & -x^2[-3a\int_\kappa^\infty (a^2+s)^{-5/2}(b^2+s)^{-1/2}(c^2+s)^{-1/2}ds - \\ & -k_b^2 a \int_\kappa^\infty (a^2+s)^{-3/2}(b^2+s)^{-3/2}(c^2+s)^{-1/2}ds - \\ & -k_c^2 a \int_\kappa^\infty (a^2+s)^{-3/2}(b^2+s)^{-1/2}(c^2+s)^{-3/2}ds] - \\ & -y^2[-a\int_\kappa^\infty (a^2+s)^{-3/2}(b^2+s)^{-3/2}(c^2+s)^{-1/2}ds - \\ & -3k_b^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-5/2}(c^2+s)^{-1/2}ds - \\ & -k_c^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-3/2}(c^2+s)^{-3/2}ds] - \\ & -z^2[-a\int_\kappa^\infty (a^2+s)^{-3/2}(b^2+s)^{-1/2}(c^2+s)^{-3/2}ds - \\ & -k_b^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-3/2}(c^2+s)^{-3/2}ds - \\ & -3k_c^2 a \int_\kappa^\infty (a^2+s)^{-1/2}(b^2+s)^{-1/2}(c^2+s)^{-5/2}ds]\}da. \end{aligned} \qquad (4\text{-}23)$$

After some simplification, this becomes

$$\begin{aligned} dV = \pi G\rho bc\, da\{ & 3I_1(a,b,c) - (3x^2+a^2)I_2(a;b,c) - \\ & -(3y^2+b^2)I_2(b;a,c) - (3z^2+c^2)I_2(c;a,b) + \\ & +(x^2b^2+y^2a^2)I_3(a,b;c) + (x^2c^2+z^2a^2)I_3(a,c;b) + \\ & +(y^2c^2+z^2b^2)I_3(b,c;a) + 3x^2a^2I_4(a;b,c) + \\ & +3y^2b^2I_4(b;a,c) + 3z^2c^2I_4(c;a,b)\}, \end{aligned} \qquad (4\text{-}24)$$

where:

$$I_1(l,m,n) = \int_\kappa^\infty (l^2+s)^{-1/2}(m^2+s)^{-1/2}(n^2+s)^{-1/2}ds, \tag{4-25a}$$

$$I_2(l;m,n) = \int_\kappa^\infty (l^2+s)^{-3/2}(m^2+s)^{-1/2}(n^2+s)^{-1/2}ds, \tag{4-25b}$$

$$I_3(l,m;n) = \int_\kappa^\infty (l^2+s)^{-3/2}(m^2+s)^{-3/2}(n^2+s)^{-1/2}ds, \tag{4-25c}$$

$$I_4(l;m,n) = \int_\kappa^\infty (l^2+s)^{-5/2}(m^2+s)^{-1/2}(n^2+s)^{-1/2}ds. \tag{4-25d}$$

The ten distinct integrals appearing in equation (4-24) may be evaluated with the aid of formulae in Section 3.13 of Gradshteyn and Ryzhik (1980, hereinafter GR). However, their formula for $I_3(b,c;a)$, equation 3.135.1 in their enumeration, contains a typographical error. This integral can be evaluated by appropriate substitutions, although a moderate amount of manipulation is required. The solution is given in Appendix A. Here we will simply quote the result.

With this correction, the integrals of equation (4-24) are specified as follows. We introduce the notations:

$$p = (a^2-b^2)^{1/2}(a^2-c^2)^{-1/2}, \tag{4-26a}$$

$$\alpha = \sin^{-1}[(a^2-c^2)^{1/2}(a^2+\kappa)^{-1/2}], \tag{4-26b}$$

$$g_1 = (a^2-b^2)^{-1}, \tag{4-26c}$$

$$g_2 = (b^2-c^2)^{-1}, \tag{4-26d}$$

$$g_3 = (a^2-c^2)^{-1/2},$$

(4-26e)

$$h_1 = (a^2+\kappa)^{1/2},$$

(4-26f)

$$h_2 = (b^2+\kappa)^{1/2},$$

(4-26g)

$$h_3 = (c^2+\kappa)^{1/2}\ .$$

(4-26h)

The solutions for the integrals depend on the assumption $a > b > c>$, and on the fact, mentioned above, that $\kappa > 0$. Then

$$I_1(a,b,c) = 2g_3F(\alpha,p),$$

(4-27a)

$$I_2(a;b,c) = 2g_1g_3[F(\alpha,p)-E(\alpha,p)],$$

(4-27b)

$$I_2(b;a,c) = 2g_3^{-1}g_1g_2E(\alpha,p)-2g_1g_3F(\alpha,p)-2g_2h_3h_1^{-1}h_2^{-1},$$

(4-27c)

$$I_2(c;a,b) = 2(-g_2)g_3E(\alpha,p)+2g_2h_2h_1^{-1}h_3^{-1},$$

(4-27d)

$$I_3(a,b;c) = 2g_2g_1^{2}g_3[(a^2+b^2-2c^2)E(\alpha,p)-2g_2^{-1}F(\alpha,p)]-$$
$$-2g_1g_2h_3h_1^{-1}h_2^{-1},$$

(4-27e)

$$I_3(a,c;b) = 2g_1g_2g_3^{3}[(2b^2-a^2-c^2)E(\alpha,p)-g_2^{-1}F(\alpha,p)]+$$
$$+2g_2g_3^{2}h_2h_1^{-1}h_3^{-1}\ ,$$

(4-27f)

$$I_3(b,c;a) = 2g_1g_2^2g_3[g_2^{-1}F(\alpha,p)-(2a^2-b^2-c^2)E(\alpha,p)] +$$
$$+ 2g_2^2h_1^{-1}h_2^{-1}h_3^{-1}(b^2+c^2+2\kappa),$$

(4-27g)

$$I_4(a;b,c) = \frac{2}{3}g_1^2g_3^3[(3a^2-b^2-2c^2)F(\alpha,p)-2(2a^2-b^2-c^2)E(\alpha,p)] +$$
$$+ \frac{2}{3}g_3^2g_1h_3h_2h_1^{-3},$$

(4-27h)

$$I_4(b;a,c) = \frac{2}{3}g_1^2g_2^2g_3[2g_3^{-2}(a^2+c^2-2b^2)E(\alpha,p) +$$
$$+ g_2^{-1}(3b^2-a^2-2c^2)F(\alpha,p)]- \frac{2}{3}[3a^2b^2-a^2c^2 +$$
$$+ 2b^2c^2-4b^4+ (2a^2-3b^2+c^2)]g_1g_2^2h_3h_1^{-1}h_2^{-3},$$

(4-27i)

$$I_4(c;a,b) = \frac{2}{3}g_2^2g_3^3[2(a^2+b^2-2c^2)E(\alpha,p)-(b^2-c^2)F(\alpha,p)] +$$
$$+ \frac{2}{3}[a^2b^2-3a^2c^2-2b^2c^2+4c^4- (2a^2+b^2-3c^2)] \times$$
$$\times g_3^2g_2^2h_2h_1^{-1}h_3^{-3},$$

(4-27j)

where $F(\alpha,p)$ and $E(\alpha,p)$ are the incomplete elliptic integrals of the first and second kinds, respectively. Table 4-2 gives the specific formulae from GR used in evaluating these integrals, except for $I_3(b,c;a)$, which, as described above, had to be evaluated through the somewhat more laborious procedure detailed in Appendix A.

TABLE 4-2
EVALUATION OF INTEGRALS FROM GRADSHTEYN AND RYZHIK (1980)

Integral	GR Formula
$I_1(a,b,c)$	3.131.1
$I_2(a;b,c)$	3.133.1
$I_2(b;a,c)$	3.133.7
$I_2(c;a,b)$	3.133.13
$I_3(a,b;c)$	3.135.9
$I_3(a,c;b)$	3.135.5
$I_4(a;b,c)$	3.134.1
$I_4(b;a,c)$	3.134.7
$I_4(c;a,b)$	3.134.13

It is clear that we may rewrite the potential dV as

$$dV = \pi G\rho bc\, da\{ PF(\alpha,p) + QE(\alpha,p) + R \}.$$

(4-28)

The functions P, Q, and R can be specified by combining equations (4-24), (4-27), and (4-28).

$$\begin{aligned} P = {} & g_1^2 g_2 g_3^3 \{[6g_1^{-2}g_2^{-1}g_3^{-2} - 2a^2 g_1^{-1}g_2^{-1}g_3^{-2} + \\ & + 2b^2 g_1^{-1}g_2^{-1}g_3^{-2}] + x^2[-6g_1^{-1}g_2^{-1}g_3^{-2} - 4b^2 g_2^{-1}g_3^{-2} - \\ & - 2c^2 g_1^{-1}g_2^{-1} + 2a^2 g_2^{-1}(3a^2 - b^2 - 2c^2)] + \\ & + y^2[6g_1^{-1}g_2^{-1}g_3^{-2} - 4a^2 g_2^{-1}g_3^{-2} + 2c^2 g_1^{-1}g_3^{-2} + \\ & + 2b^2 g_3^{-2}(3b^2 - a^2 - 2c^2)] + z^2[-2a^2 g_1^{-1}g_2^{-1} + \\ & + 2b^2 g_1^{-1}g_3^{-2} - 2c^2 g_1^{-2}] \}, \end{aligned}$$

(4-29)

$$\begin{aligned} Q = {} & 2g_1^2 g_2^2 g_3^3 \{[a^2 g_1^{-1}g_2^{-2}g_3^{-2} - b^2 g_1^{-1}g_2^{-1}g_3^{-4} + \\ & + c^2 g_1^{-2}g_2^{-1}g_3^{-2}] + x^2[3g_1^{-1}g_2^{-2}g_3^{-2} + \\ & + b^2 g_2^{-1}g_3^{-2}(a^2 + b^2 - 2c^2) + \\ & + c^2 g_1^{-1}g_2^{-1}(2b^2 - a^2 - c^2) - 2a^2 g_2^{-2}(2a^2 - b^2 - c^2)] + \\ & + y^2[-3g_1^{-1}g_2^{-1}g_3^{-4} + a^2 g_2^{-1}g_3^{-2}(a^2 + b^2 - 2c^2) - \\ & - c^2 g_1^{-1}g_3^{-2}(2a^2 - b^2 - c^2) + 2b^2 g_3^{-4}(a^2 + c^2 - 2b^2)] + \\ & + z^2[3g_1^{-2}g_2^{-1}g_3^{-2} + a^2 g_1^{-1}g_2^{-1}(2b^2 - a^2 - c^2) - \\ & - b^2 g_1^{-1}g_3^{-2}(2a^2 - b^2 - c^2) + 2c^2 g_1^{-2}(a^2 + b^2 - 2c^2)] \}, \end{aligned}$$

(4-30)

$$R = 2g_1g_2^2g_3^2h_1^{-3}h_2^{-3}h_3^{-3}\{[b^2g_1^{-1}g_2^{-1}g_3^{-2}h_1^2h_2^2h_3^4 -$$
$$- c^2g_1^{-1}g_2^{-1}g_3^{-2}h_1^2h_2^4h_3^2]+x^2[-b^2g_2^{-1}g_3^{-2}h_1^2h_2^2h_3^4 +$$
$$+ c^2g_1^{-1}g_2^{-1}h_1^2h_2^4h_3^2+a^2g_2^{-2}h_2^4h_3^4] +$$
$$+ y^2[3g_1^{-1}g_2^{-1}g_3^{-2}h_1^2h_2^2h_3^4-a^2g_2^{-1}g_3^{-2}h_1^2h_2^2h_3^4 +$$
$$+ c^2g_1^{-1}g_3^{-2}h_1^2h_2^2h_3^2(b^2+c^2+2\kappa)-b^2g_3^{-2}h_1^2h_3^4 \times$$
$$\times (3a^2b^2-a^2c^2+2b^2c^2-4b^4+\kappa(2a^2-3b^2+c^2))] +$$
$$+ z^2[-3g_1^{-1}g_2^{-1}g_3^{-2}h_1^2h_2^4h_3^2+a^2g_1^{-1}g_2^{-1}h_1^2h_2^4h_3^2 +$$
$$+ b^2g_1^{-1}g_3^{-2}h_1^2h_2^2h_3^2(b^2+c^2+2\kappa)+c^2g_1^{-1}h_1^2h_2^4 \times$$
$$\times (a^2b^2-3a^2c^2-2b^2c^2+4c^4-\kappa(2a^2+b^2-3c^2))]\}.$$

(4-31)

Upon replacing the g's and h's by their definitions in terms of a,b,c, and κ, and manipulating the resulting expressions, equations (4-29) through (4-31) yield remarkable simplifications:

$$P = 4g_3 ,$$

(4-32)

$$Q = 0,$$

(4-33)

$$R = 0 .$$

(4-34)

The result for R relies upon the definition of κ. Combining these results with equations (4-26e) and (4-28), we obtain the very simple form for the potential:

$$dV = 4\pi G\rho bc\, da(a^2-c^2)^{-1/2}F(\alpha,p).$$

(4-34)

In the oblate limit (a = b), we have p = 0, and the elliptic integral F (α,p) becomes simply α. If we further restrict ourselves to points in the equatorial plane of the spheroid, z = 0, then $\kappa = r^2-a^2$,

where $r^2 = x^2+y^2$. The potential becomes, from (4-26b) and (4-35),

$$dV = 4\pi G\rho ae^{-1}(1-e^2)^{1/2}da\ \sin^{-1}(\frac{ae}{r}),$$

(4-36)

where we have defined $e = (a^2-c^2)^{1/2}a^{-1}$. Equation (4-36) is the familiar expression for an oblate shell (Mestel, 1963). The prolate form, which is very similar to this, can be recovered by interchanging c and a, and using the identities $(a^2-c^2)^{1/2} = i(c^2-a^2)^{1/2}$ and $-i\ \sin^{-1}i\theta = \sinh^{-1}\theta$ (MacMillan, 1958, p. 63).

An even further simplification ensues when the forces are computed. It arises because the position (x,y,z) enters into (4-35) only through the argument α of the elliptic integral. Therefore, by the definition of $F(\alpha,p)$ (e.g., GR, p. 904), the x component of the force, for example, becomes:

$$dF_x = 4\pi G\rho bc(a^2-c^2)^{-1/2}da\ x$$
$$x\ \frac{\partial}{\partial x}\int_0^{\sin\alpha}(1-t^2)^{-1/2}(1-p^2t^2)^{-1/2}dt.$$

(4-37)

Recalling the definition of α in equation (4-26b), we obtain

$$dF_x = 4\pi G\rho bc(a^2-c^2)^{-1/2}\ da(1-\sin^2\alpha)^{-1/2}\ x$$
$$x\ (1-p^2\sin^2\alpha)^{-1/2}(-1/2)(a^2-c^2)^{1/2}(a^2+\kappa)^{-3/2}\ \frac{\partial\kappa}{\partial x}\ ,$$

(4-38)

or

$$dF_x = -2\pi\rho bc\ da[(a^2+\kappa)(b^2+\kappa)(c^2+\kappa)]^{-1/2}\frac{\partial\kappa}{\partial x},$$

(4-39a)

and, similarly,

$$dF_y = -2\pi G\rho bc\ da[(a^2+\kappa)(b^2+\kappa)(c^2+\kappa)]^{-1/2}\frac{\partial\kappa}{\partial y},$$

(4-39b)

$$dF_z = -2\pi G\rho bc\ da[(a^2+\kappa)(b^2+\kappa)(c^2+\kappa)]^{-1/2}\frac{\partial\kappa}{\partial z}.$$

(4-39c)

The only difficulty in computing the force, then, is the evaluation of κ and of $\frac{\partial\kappa}{\partial x}$, etc. In the general case, κ is the greatest root of a cubic equation whose coefficients are functions of (a,b,c) and (x,y,z), and can be found by application of standard techniques (e.g., Abramowitz and Stegun, p. 17) at each point (x,y,z). Implicit differentiation of the cubic equation readily yields analytical expressions for the partial derivatives of κ, in terms of (a,b,c), (x,y,z), and κ. Therefore, after one has solved for κ, evaluation of the three force components is utterly straightforward. Here we illustrate the procedure in the simpler case where the point of evaluation is in one of the planes of symmetry of the homoeoid, say z = 0. In this case, κ is the larger root of the quadratic equation implied by

$$\frac{x^2}{a^2+} + \frac{y^2}{b^2+} = 1.$$

(4-40)

Therefore,

$$\kappa = \frac{1}{2}[(x^2+y^2)-(a^2+b^2)] + \frac{1}{2}\{[(x^2+y^2)-(a^2+b^2)]^2 - 4(a^2b^2-x^2b^2-y^2a^2)^{1/2}\}. \tag{4-41}$$

Defining $r^2 = x^2 + y^2$ and $l^2 = a^2-b^2$, this can be rearranged to:

$$\kappa = \frac{1}{2}[r^2-(a^2+b^2] + \frac{1}{2}[r^4+l^4(r^2-2y^2)]^{1/2} \tag{4-42}$$

Differentiation with respect to x, for example gives

$$\frac{\partial\kappa}{\partial x} = x\{1+(r^2-l^2)[r^4+l^4-2l^2(r^2-2y^2)]^{-1/2}\}, \tag{4-43}$$

and this simplifies to

$$\frac{\partial\kappa}{\partial x} = x\left(\frac{2\kappa+2b^2}{2\kappa+a^2+b^2-r^2}\right). \tag{4-44}$$

Similarly,

$$\frac{\partial\kappa}{\partial y} = y\left(\frac{2\kappa+2a^2}{2\kappa+a^2+b^2-r^2}\right). \tag{4-45}$$

The system of equations consisting of (4-39a), (4-39b), (4-42), (4-44), and (4-45) make the exact computation of the force due to a triaxial homoeoid very easy, therefore, in one of its principal planes.

Equations (4-39) are equally valid outside of the planes of symmetry. In that case, κ, being the root of a cubic equation, is more cumbersome to write explicitly. Once it has been found, however, equations for $\frac{\partial \kappa}{\partial x}$, $\frac{\partial \kappa}{\partial y}$, and $\frac{\partial \kappa}{\partial z}$ can be written which are similar in form to (4-44) and (4-45). Then one can calculate the force due to a mass density that varies in an arbitrary manner, provided only that it is constant on triaxial surfaces, via numerical integration of equations (4-39), to any desired degree of accuracy.

CHAPTER V
HYDRODYNAMICAL MODELS

The Beam Scheme

The computer code used to follow the dynamics of the gas in our hydrodynamical models of NGC 3359 is the two-dimensional "beam scheme" program, introduced by Sanders and Prendergast (1974, hereafter SP). The version of the program used here was developed and made available to us by Dr. J. M. Huntley. As mentioned in Chapter I, Huntley and others have published a number of models of barred spiral galaxies which make use of this code (Sanders and Huntley, 1976; Huntley, 1977; Huntley, Sanders, and Roberts, 1978; Sanders and Tubbs, 1980; Huntley, 1980; Schempp, 1982). The hydrodynamical properties of the beam scheme are reviewed by SP, and the references listed above have discussed at some length its application to the modeling of barred spirals. We will give only a brief summary of its numerical characteristics, along with an outline of its use in the present context.

The beam scheme takes its name from the representation of the gas in each cell of the numerical grid by a small number of discrete beams. A two-dimensional, Cartesian, Eulerian grid is used for the calculations reported in this chapter, but the salient characteristics of the code are more easily illustrated by considering the one-dimensional example. The quantitative discussion below follows that of the Appendix of SP.

The fundamental assumption of the beam scheme is that the gas is characterized, at all times and in each cell of the grid, by a Maxwellian velocity distribution. In the one-dimensional case, this means that the

velocity distribution in each cell is given by a function of the form

$$F_j(u,v,w) = A_j \exp\ -\frac{1}{2\sigma_j^2}[(u-\bar{u}_j)^2+v^2+w^2] \ .$$

(5-1)

Here the subscript j refers to the jth cell of the grid; A_j is proportional to the mass density in the cell; σ_j is the velocity dispersion; u, v, and w are the x, y, z velocities of the matter in the cell. Assuming that the x coordinate specifies the one degree of freedom in the problem, only this cordinate has an allowed bulk velocity, which is $\bar{u}_j$. The problem of the beam scheme program is to approximate the distribution of equation (5-1) by a set of discrete beams under appropriate constraints. These are chosen by requiring the first few moments of the distribution to be the same as in the continuous case. In particular, the beam scheme works with the first three moments, which are, respectively, the mass density ρ_j, momentum density p_j, and kinetic energy density E_j associated with cell j:

$$\rho_j = \int_{-\infty}^{\infty} F_j(u,v,w)\ du\ dv\ dw,$$

(5-2)

$$p_j = \int_{-\infty}^{\infty} F_j(u,v,w)\ u\ du\ dv\ dw,$$

(5-3)

$$E_j = \int_{-\infty}^{\infty} \frac{1}{2} F_j(u,v,w)(u^2+v^2+w^2)\ du\ dv\ dw.$$

(5-4)

In the computer program, F_j (u,v,w) is approximated by three distinct beams of gas, in each of which all the matter moves at a single velocity. Each beam contains a particular fraction of the mass in its cell, and it

is required that the distribution of the mass and momentum over the available beams preserve the moments ρ_j, p_j, and E_j of the Maxwellian distribution $F_j(u,v,w)$. The "central beam," with most of the mass, moves at the velocity u_{1j}. Additionally, there are two "side beams," offset in velocity from the central beam by equal but opposite increments Δu. Individually, these are designated the forward and back beams; their associated mass densities are equal. For problems of higher dimensionality, two additional side beams are required for each added dimension. The complexity of the bookkeeping needed goes up very quickly, but there are no fundamental differences from the one-dimensional case.

Let us denote the mass densities in the central beam, and in a single side beam, by a_j and b_j, respectively. These weights may be chosen arbitrarily. However, SP show that the choice

$$a_j = 4b_j \tag{5-5}$$

allows the fourth-order moment of the continuous distribution $F_j(u,v,w)$ to be preserved as well. This choice determines the distribution of beam velocities:

$$u_{1j} = \bar{u}_j, \tag{5-6a}$$

$$(\Delta u_j)^2 = 3\sigma_j^2 . \tag{5-6b}$$

Table 5-1 gives the densities of mass, momentum, and energy carried by each of the three beams.

TABLE 5-1
BEAM PROPERTIES, ONE-DIMENSIONAL CASE

Central beam

$$\rho_{1j} = 4b_j$$
$$p_{1j} = 4b_j\bar{u}_j$$
$$E_{1j} = 2b_j\bar{u}_j^2+4b_j\sigma_j^2$$

Back beam

$$\rho_{2j} = b_j$$
$$p_{2j} = b_j(\bar{u}_j - \sqrt{3}\ \sigma_j)$$
$$E_{2j} = b_j(\frac{1}{2}\ \bar{u}_j^2 - \sqrt{3}\ \bar{u}_j\sigma_j + \frac{5}{2}\sigma_j^2)$$

Forward beam

$$\rho_{3j} = b_j$$
$$p_{3j} = b_j(\bar{u}_j+\sqrt{3}\ \sigma_j)$$
$$E_{3j} = b_j(\frac{1}{2}\ \bar{u}_j^2 + \sqrt{3}\ \bar{u}_j\sigma_j + \frac{5}{2}\sigma_j^2)$$

The evolution of the system is followed by means of a series of discrete time steps. The size Δt of each time step is computed by the Courant condition,

$$\Delta t < \frac{L}{u_{max}}, \qquad (5\text{-}7)$$

where L is the cell size and u_{max} is the largest velocity present on the grid. Equation (5-6) is a necessary condition for numerical stability. Since u_{max} varies with time, Δt differs from one time step to the next.

In each time step, every beam transports a proportion of its mass, momentum, and energy into neighboring cells, the proportion being equal to the fraction of a cell length traversed during the time step. We may write this fraction as g_{ijk}, where the first subscript identifies the beam (using the convention i = 1 for the central beam, i = 2 for the back beam, and i = 3 for the forward beam), the second subscript the cell of origin, and the third the destination cell. Thus,

$$g_{ijk} = \frac{u_{ij}\Delta t}{L} . \qquad (5\text{-}8)$$

Of course, in practice one must be careful that these fractions are computed only for the appropriate destination cell. Cell j retains the fraction $g_{ijj} = (1-g_{ijk})$ of the same quantities during the time step.

The following simple example serves to illustrate the action of a time step in the beam scheme. Suppose that gas is initially flowing in

the direction of the higher numbered cells. Specifically, consider the transport of matter from cell 1 to cell 2. We will further assume that the flow velocities are large enough so that all beams in each cell move in the same direction. The density in cell 2, after time step Δt, is

$$\rho_2' = g_{112}a_1+g_{212}b_1+g_{312}b_1 \; + \; + \; g_{122}a_2+g_{222}b_2+g_{322}b_2 \; . \tag{5-9}$$

Substitution from Table 5-1 for the quantities on the right-hand side of this equation gives, simply,

$$\rho_2' = \rho_2 + \frac{\Delta t}{L} (\rho_1\bar{u}_1-\rho_2\bar{u}_2), \tag{5-10}$$

or, in general,

$$\frac{\Delta\rho_j}{\Delta t} = \frac{1}{L} (\rho_{j-1}\bar{u}_{j-1} - \rho_j\bar{u}_j). \tag{5-11}$$

Equation (5-11) is a finite-difference version of equation (1-1), the hydrodynamical continuity equation. However, it differs in one important respect from the most straightforward finite-difference representation of that equation. Equation (5-11) is not centered; that is, the right-hand side, which is equivalent to a spatial differentiation of the momentum, is computed by reference to cell j and one of its neighbors, while the other neighbor is ignored. In effect, the center of the

differencing is moved half a cell in the direction from which matter is flowing. This is known as "upwind differencing," and hydrodynamical codes which employ it are well known to have a high degree of numerical stability (Roache, 1976, pp. 64-67). The equations of motion and of energy are approximated, in the beam scheme, by finite-difference equations in which the spatial differentiations are of precisely the same form as in equation (5-11).

When the densities, momenta, and energies of all cells in the grid have been updated in this way, the time step is complete. Before proceeding with the next, these quantities are redistributed over the three beams of each cell. In other words, a Maxwellian velocity distribution is re-established. The physical interpretation of this redistribution is that the velocities in the gas are randomized, by collisions, over a distance scale equal to the cell size L, or, equivalently, that shear viscosity acts, with a mean free path of about L. The effects of this viscosity have been discussed by SP, and, at more length, by Huntley (1977) and by Huntley, Sanders, and Roberts (1978, hereafter HSR). The effective shear viscosity at cell j is approximately

$$\eta = \rho_j \Delta u_j L. \qquad (5\text{-}12)$$

Obviously, this viscosity is artificial, but the justification for it is that a real, physical viscosity must be present in the gaseous components of disk galaxies, in order to prevent gas streamlines from crossing. Indeed, HSR conclude, based on their results and on the Lagrangian calculations of Sanders (1977), that this viscosity is essential to the development of trailing, two-armed spirals in barred galaxies. The details of the viscosity were found to be unimportant, to the limits of resolution of the respective calculations. This circumstance is

fortunate, given our ignorance of the real viscosity in the interstellar medium.

Having briefly described the hydrodynamical properties of the beam scheme, we will now comment with equal brevity on the implementation of the program for the models to be presented in this chapter. The Cartesian grid employed in these calculations is two-dimensional, of extent 32 x 64 cells. This arrangement was developed specifically for the computation of barred spiral models, under the assumption that the force, and the gas response, are bisymmetric. Thus, the symmetry inherent in the problem is used to allow the representation of a square, 64 x 64 grid with only half the storage space. One of the two longer boundaries of the 32 x 64 grid is conceived as defining a meridian plane of the galaxy, and matter crossing this boundary re-enters the grid at a point which is symmetric, relative to the galaxy center, with that which it has just left. Each cell in the computed grid, therefore, represents two cells in the full grid, the second of which would be located diametrically to the computed cell.

The gas is only allowed to inhabit cells whose centers are within 32 units of the galaxy center, defining a circular disk; in other words, the corners of the square grid are empty. Material which passes into cells beyond this radius simply leaves the calculation permanently. Consequently, the total mass of gas in the system decreases continually, though slowly. This leakage across the outer boundary is the only sink of gas in the program, and there are no sources. Also, the gas is modeled as a single-phase, continuous medium. These conditions amount to the assumption that the large-scale dynamics of the system do not depend on the complex, local interactions of the neutral gas with other

components of the galaxy. The gaseous disk in these calculations is maintained in a strictly isothermal state, with a constant velocity dispersion of 17 km/s. HSR found in test cases that no significant differences arose among models having velocity dispersions of 10, 17, and 25 km/s.

The mass density is initially constant over the grid. However, the initial velocity conditions of the gas must be fixed with some care. If the velocity field at any given time step is not rather close to the equilibrium field for the forces then extant, severe disruption of the flow can result, quite possibly vitiating the subsequent evolution of the model entirely. We have followed the practice of Huntley by starting each calculation with the gas in pure, solid-body rotation and in equilibrium. The angular velocity of this rotation is equal to ($\sqrt{2}/2$ times that of the galaxian bar which eventually will be established. Over the course of the first 100 time steps of the run, the force changes linearly to that of the imposed gravitational field chosen for that model. That is, in each time step 1% of the initial uniform-rotation force is "turned off," and 1% of the desired force "turned on." After 100 time steps, the force no longer evolves. This procedure allows the gas to adjust quasi-statically to the noncircular forcing employed in these barred spiral models.

The models to be presented here have been followed for 400 time steps, long enough for several rotations of the bar to have taken place, and for a stable response to have emerged in the gas. This expectation has been confirmed repeatedly by running selected models for 600 time steps, or occasionally longer. After a time, usually amounting to about 300 time steps, the models show almost steady-state behavior, with the

only significant evolution being a slow, secular increase in the contrast between the highest and lowest densities present in the disk, the former occurring at the center of the computational grid. This reflects the pattern of inflow and outflow associated with numerical effects in the program. At the very center of the grid, there is a steady accumulation of matter caused by numerical diffusion. This phenomenon is primarily caused by the spatial discretization of the grid, which is of relatively greater importance at its center (Huntley, 1977). At the same time, the outer regions of the galaxy constantly lose gas by thermal diffusion and noncircular flow across the outer boundary.

As we have implied above, the gravitational forces of the disk and bar of a particular model do not evolve with time, in accordance with our assumption that the dynamics of the gas are essentially a response to the gravitational fields of the more massive components of the system. Obviously, the attraction of one element of the modeled gas for another--the "self-gravity" of the gas--cannot be included in such an approach. As the H I mass is typically only a few percent of the total mass of a galaxy, there is some justification for neglecting this force. On the other hand, because of its ability to dissipate energy in shocks, gas can exhibit considerably greater enhancements at the locations of shallow potential minima than can relatively collisionless objects like stars. Consequently, at such locations the local density of gas can become a substantial fraction of the total density. It is then not so clear that its self-gravity is unimportant in influencing its large-scale behavior.

There are two decisive factors in our decision to proceed with models that fail to include self-gravity. The first is that a considerable savings in computational time and effort is made possible by this

choice. The more important reason, however, is a negative one. Huntley (1980) investigated models based on the numerical bar found by Miller and Smith (1979) in their n-body calculations; this bar is known from their experiments to be an equilibrium figure. Huntley's experience was that deviations of more than about 5% from the equilibrium values, either in bar rotation rate or in mass distribution, produced models in which no stable structure emerged, but only a succession of transient figures. Assuming this result to be general, the difficulty of finding self-consistent disk-bar systems for computing models of an observed galaxy seems daunting indeed. We will return to a discussion of this point later in this chapter.

The neglect of self-gravity, besides yielding somewhat inaccurate estimates of the forces acting on the gas, is part of a larger inconsistency in our models. The response which is evoked by the existing potentials does not cause any rearrangement of the mass distributions which give rise to those potentials. Thus, at the deepest level, there is a total lack of dynamical feedback in these models; this is rather worrisome, since we by no means know that our chosen potentials and pattern speeds correspond to equilibrium figures. If a perfect galaxy model is ever computed, it will have to include all known components of the system and demonstrate that their dynamics are mutually consistent. For the present, we only hope to make progress in setting limits on the plausible.

Finally, we must mention the problem of the proper computational frame. This question has been discussed at some length by Huntley (1977, 1980). The two obvious reference frames in which to compute models of a barred spiral galaxy are the inertial frame and the rotating frame whose

angular velocity is equal to the pattern speed of the bar. Unfortunately, there are numerical disadvantages inherent in each of these frames.

An elementary stability analysis of the beam scheme, summarized in Appendix B, suggests that the code may be unstable when velocities fall into the subsonic regime. In an inertial frame, this presents no problem, since the rotational velocities in disk galaxies easily exceed thermal speeds. For any rotating frame, however, there will be a "corotation zone" in which the rotational velocity of the material in the disk matches the angular velocity of the frame, so that the computed velocities are virtually zero. We will call this condition "corotation-with-the-frame," to distinguish it from the physically significant corotation-with-the-bar, or simply corotation. Unfortunately, when computations are performed in the rotating frame of the perturbation, there is no distinction between the two. In models computed in this frame by Huntley (1977), no actual instabilities arose, but a nonphysical enhancement of the gas density, tending to form a ring, was observed. This ring originates simply because, since the velocity goes to zero as one approaches this ring from either side, the gas finds it much easier to enter this zone than to leave it.

The drawback of the inertial frame occurs because in this case the pattern speed of the bar relative to the computational grid is maximized. Consequently, the bar may sweep past the grid at a rate greater than one cell per time step; in this event, the gas is unable to respond quickly enough to the perturbation, which is, in effect, smeared out over a broad area. This is equivalent to a greatly increased effective viscosity in the affected parts of the grid; manifestly, the effect increases linearly with radius.

Both of these problems may be avoided by choosing a computational frame which rotates at an intermediate speed. We have adopted the practice of selecting a frame speed which places the estimated position of corotation-with-the-frame five cells beyond the maximum radius used in the calculation. This ensures that velocities at all points on the computed grid are high enough to prevent numerical artifacts, while minimizing the problems of added viscosity and lower effective resolution in the outer part of the disk.

Models Using Triaxial Bars

We now pass from a review of the general properties of the beam scheme, and its use in barred spiral galaxy models, to a consideration of the specific problem at hand. So far, we have not addressed the question of the fixed, external forces which may be included in the hydrodynamical computations. In the models to be discussed in this chapter, these explicit forces are simply the gravitational forces associated with the axisymmetric and nonaxisymmetric components of NGC 3359. As we have said earlier, these forces, along with the angular velocity or pattern speed of the bar, are the physically meaningful input parameters which distinguish various beam scheme models of the galaxy from one another.

For the sake of simplicity, the only axisymmetric components we will consider in this section are disks, as there is no difference in the gas response between this case and multi-component models, as long as the observed rotation curve is reproduced adequately. In Chapter III, we saw that the Generalized Mestel Disk (GMD) is a reasonable first-order fit to the rotation curve of NGC 3359. Initially, we attempted to determine the mass distribution of the bar from

observations as well, using the approach detailed in Chapter IV. The outcome of the modeling based on this approach, however, was less than satisfactory, as we shall soon explain. We were thus led to investigate radically different bar potentials from those derived using the surface-photometry method. The remainder of this chapter describes our results for two types of bar potential: triaxial, or "massive" bars (in keeping with the analysis of Chapter IV), and oval distortions, which are sometimes referred to as massless bars. We will discuss the characteristics of the models resulting from these two cases in turn, beginning with those employing triaxial bars.

In Chapter IV, we applied the method of Stark (1977) to the determination of a family of model bars whose mass density distributions fit that of the bar in NGC 3359, within the observational uncertainties. The bars allowed by this procedure are inhomogeneous, triaxial ellipsoids, whose shortest axes coincide with the rotation axis of the galaxy. From the observed run of surface brightness along the bar major axis, the density distribution in the bar of NGC 3359 was inferred to consist of two nested, Gaussian functions, whose amplitudes and length scales are specified by the model surface brightness of equation (4-12). The geometrical parameters of the bars are typified by the members of the family listed in Table 4-1. The technique for incorporating these bars into the hydrodynamical models is quite simple. The only free parameters are the pattern speed of the bar, its mass, relative to that of the disk, and its degree of flattening in the direction perpendicular to the galaxy plane. By varying the values of these three quantities for different runs of the program, one can explore empirically the response of the gaseous disk to the full range of triaxial bars.

The results of these models are both clear-cut and unexpected. The model galaxies which use triaxial bars derived from the photometry are emphatic failures in reproducing the observed distribution and kinematics of the gas in NGC 3359.

The nature of this failure is as follows. In the density response, there is a fair degree of variation among particular runs, but several characteristics recur in each case. Near the center of the galaxy, the density constrast between the bar major and minor axes tends to be quite large, with a well-defined enhancement in the gas density along the bar itself. This gas bar is a common feature of barred spiral models in the beam scheme, but is not necessarily real; to a large extent it owes its existence to the inevitable infall of material to the center of the grid. As discussed in the previous section, this is a numerical effect. The more important phenomenon in the triaxial models under discussion is the substantial void which commonly appears to surround the bar in these models. Clearly, the triaxial bars are efficient at sweeping up the gas in their immediate vicinities (Hunter et al., 1984). As we have seen, the H I observations of NGC 3359 imply that a rather mild sweeping of the bar zone may have taken place in that galaxy, assuming that the apparent depletion of gas is not compensated by the presence of significant quantities of molecular hydrogen.

However, the density distributions in the outer regions of the models are in serious disagreement with the observations. While many of the runs show slight, trailing spiral enhancements in the outer disk, the contrast between these features and the ambient densities at the same radius is invariably quite low. The response, in other words, is primarily axisymmetric, and is aptly characterized as a broad ring of moderately high density surrounding the aforementioned inner zone of partial depletion. Furthermore, the

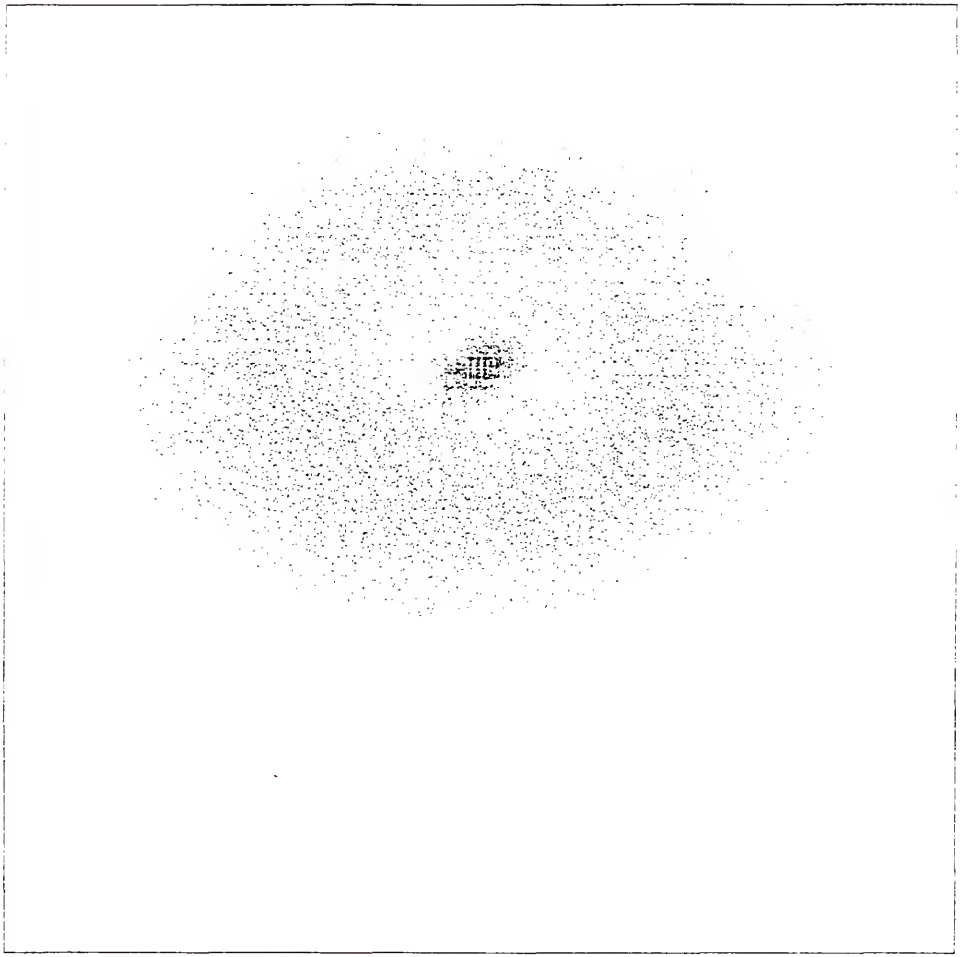
vestigial arms that do appear are not smoothly connected to the bar region, and do not, in general, extend as far inward as the inner portions of the observed arms in NGC 3359.

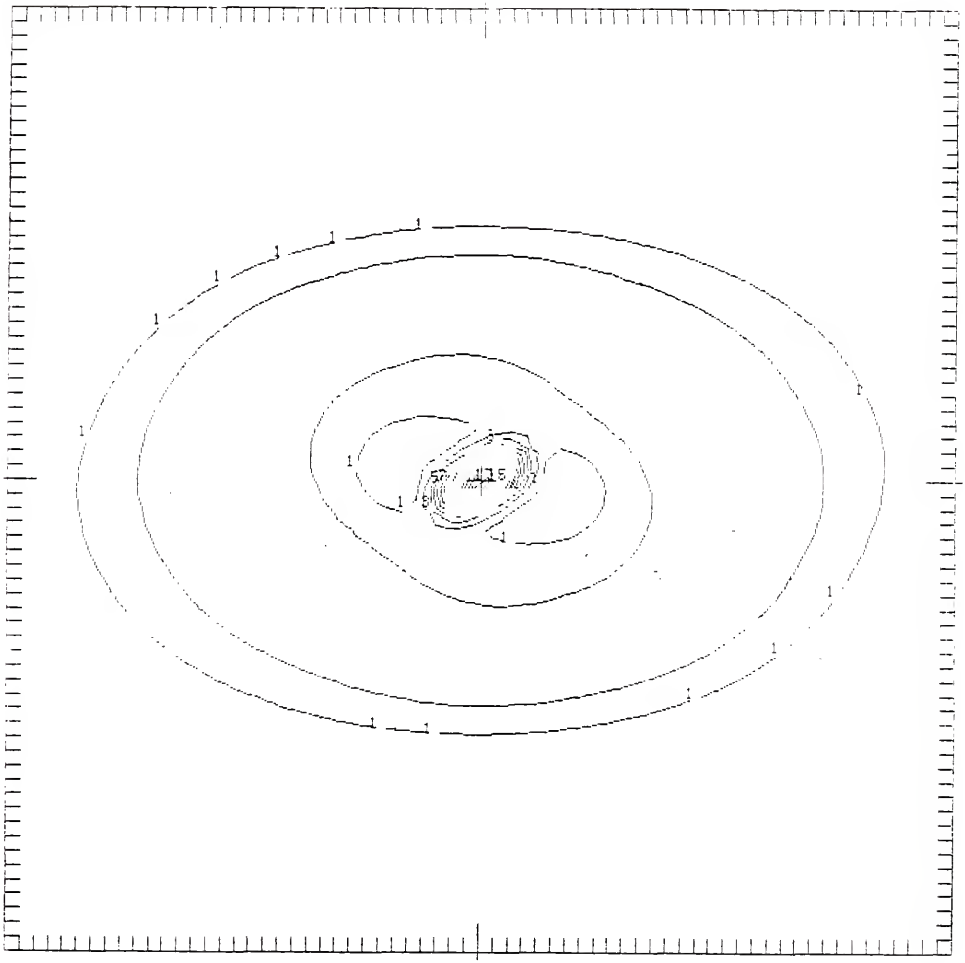
The shortcomings of the triaxial models are exposed most clearly in their velocity fields, which bear almost no resemblance to that of NGC 3359, beyond their gross rotational properties. The presence of the triaxial bar causes a pronounced disturbance at the center of each model galaxy, skewing the velocity contours across the minor axis sharply in the immediate neighborhood of the bar, while having relatively little effect at greater radii. The kinematical perturbations associated with the spiral arms, seen so clearly in the observed velocity field of Figure 3-6, are practically nonexistent. This behavior is a striking illustration of the fundamental defect of the triaxial models, namely, that the nonaxisymmetric forcing is much too weak in the outer galaxy relative to its value in the bar region. We will have a good deal more to say on this subject shortly.

The characteristics of models based on triaxial bars are well illustrated by the pair of examples whose density and velocity responses are displayed in Figures 5-1 and 5-2. In both of these runs, as with all of the models we have computed, the area to be modeled is restricted to a galactocentric radius of 3!75, or 12.0 kpc at the assumed distance of 11 Mpc. This restriction was necessitated by the absence of well-defined velocities around a complete circumference of the observed galaxy, at larger radii. These models employ a GMD and triaxial bar, adjusted to give a composite rotation curve which matches that of NGC 3359. In each case, the mass of the bar is 11% that of the disk, and the pattern speed of the bar is 91.5 km/s/arc minute, placing corotation a short distance beyond the end of the bar. In these displays, and

Figure 5-1. Gas response in the first of the purely triaxial models of NGC 3359 discussed in this chapter. The results of the model have been projected in the same manner as the observed galaxy, and convolved with a Gaussian beam of full width at half power 18". The maximum radius of the model galaxy is 3.'75, and the spacing of the tick marks is 0.'117. The model is displayed with its line of nodes horizontal. (a) Gas surface density, grey tone display. (b) Gas surface density, contour display. (c) Line of sight, density-weighted velocities in the gas. The contour interval is 20 km/s, and the contours are labeled with the velocities in units of this interval.

NGC 3359 MODEL -- CONVOLVED





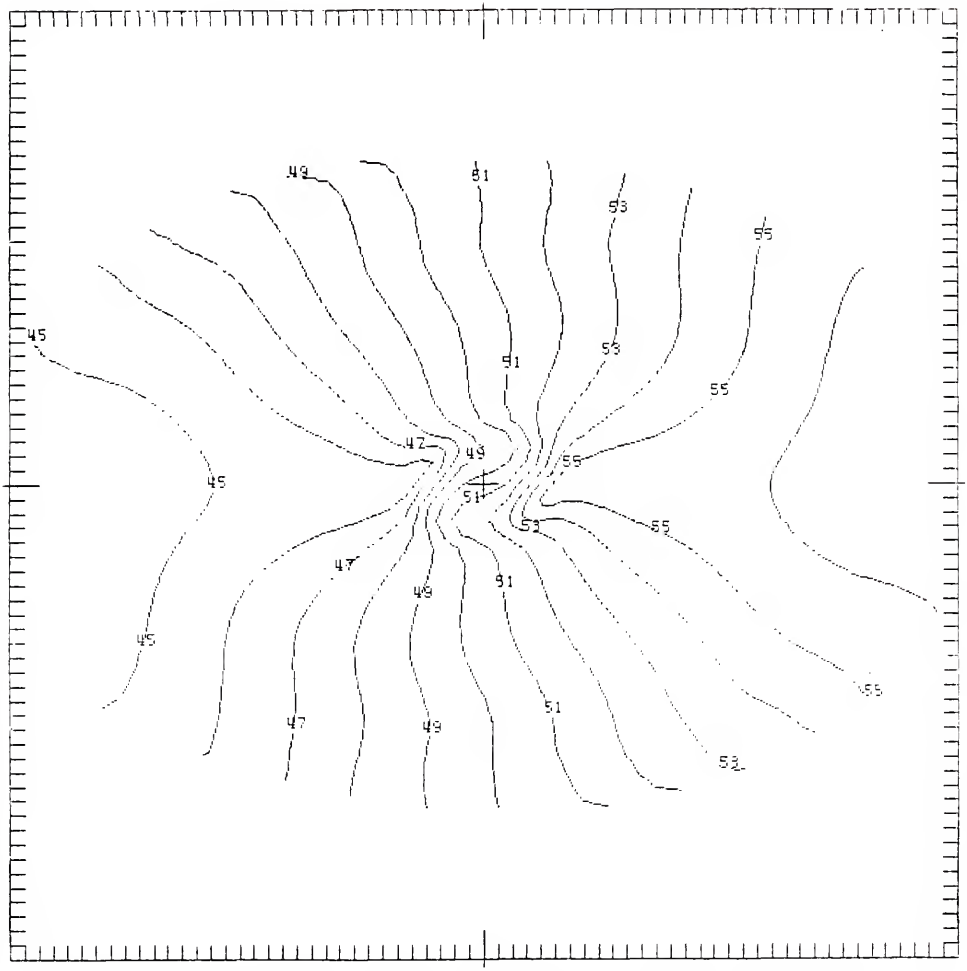
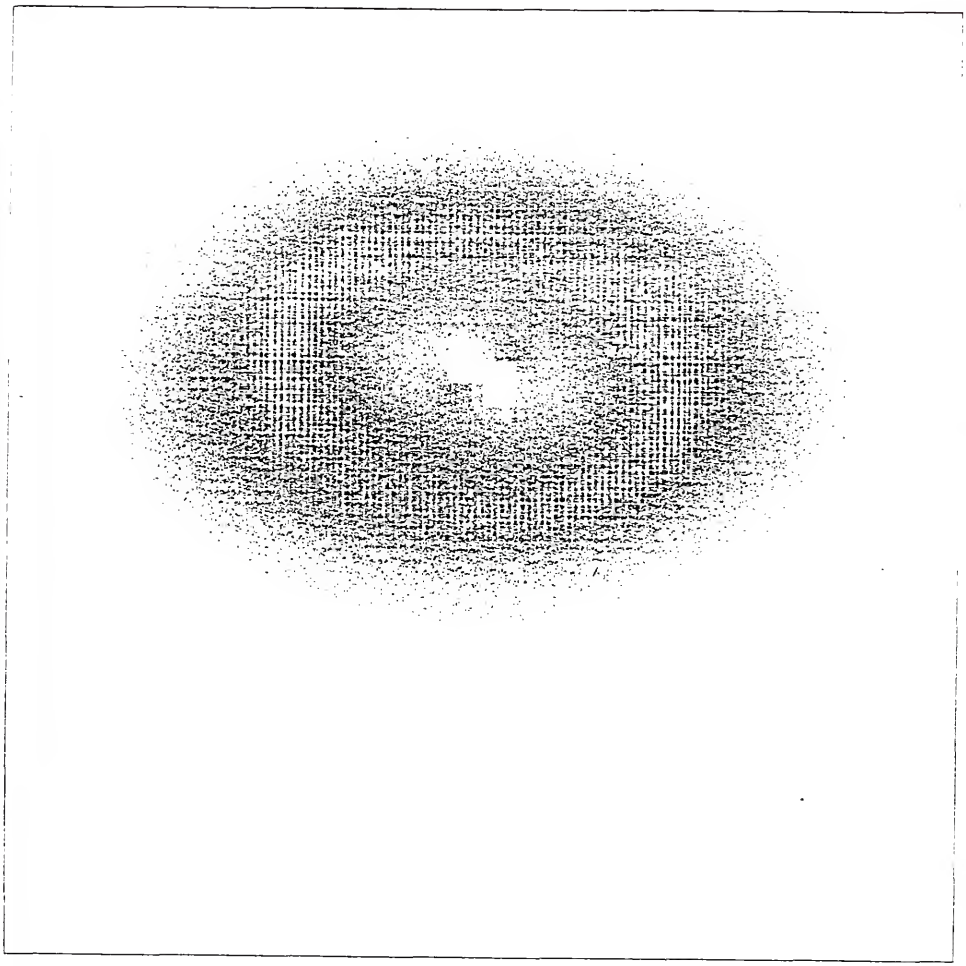
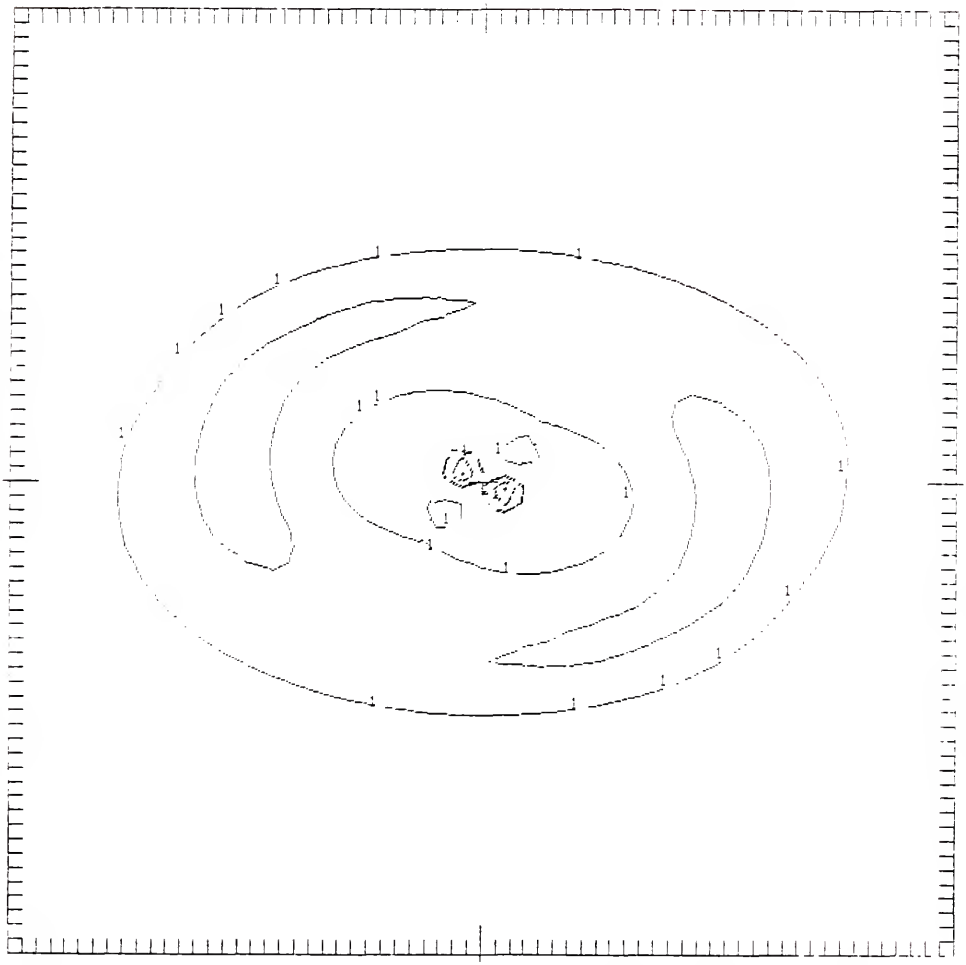
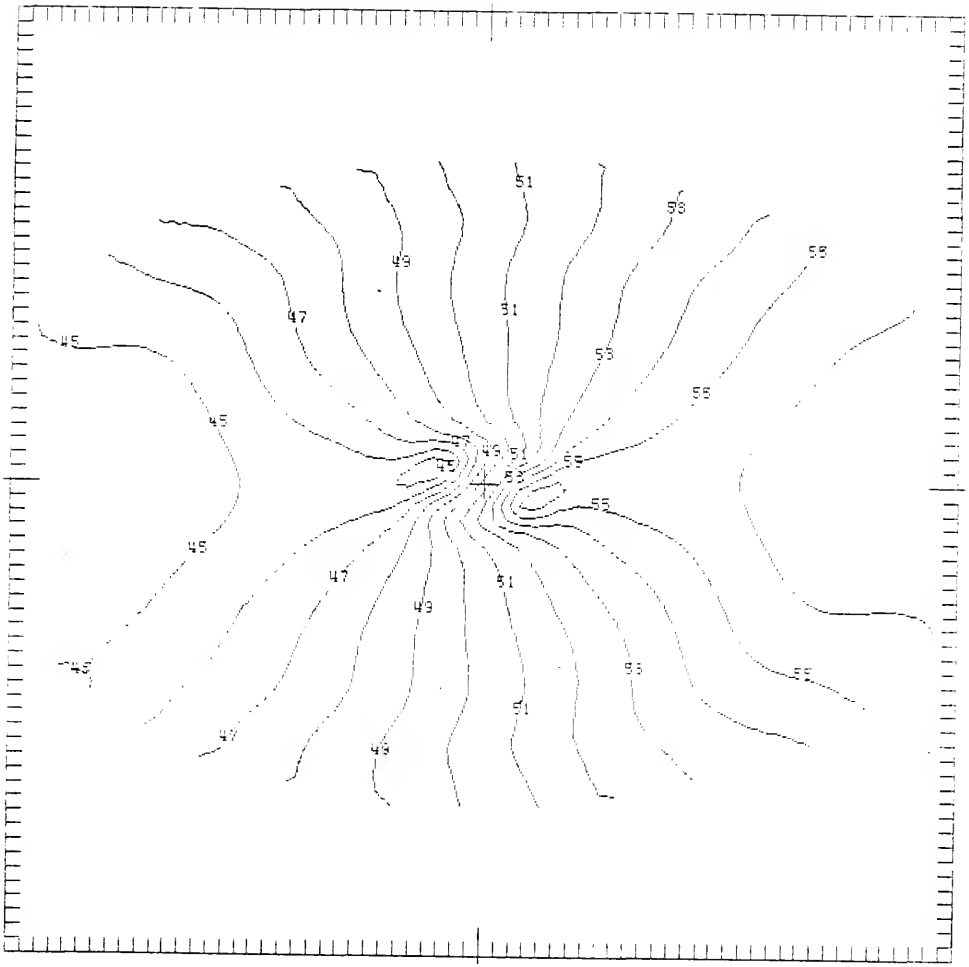

49
51
53
55
45
51
53
55
47
49
55
45
45
47
49
51
53

Figure 5-2. Gas response in the second of the purely triaxial models of NGC 3359 discussed in this chapter. This model differs from that shown in Figure 5-1 by having a more flattened bar. The identification of parts (a) through (c) are as in Figure 5-1.

NGC 3559 MODEL -- CONVOLVED






45
47
49
51
53
55

throughout this chapter, the models have been projected in space to match the orientation of NGC 3359, and then convolved with a Gaussian beam of the same width as that obtained at the VLA. The velocities were weighted by the densities in performing the convolution, with the central spike in the model density field truncated to avoid distorting the velocities in its vicinity. The model galaxies are displayed with the line of nodes horizontal. The full extent of the modeled area, inside the tick marks on the contour plots, is 7.5. The contour plots of the surface densities do not use a constant set of contour levels; the numerical values of the density are not necessarily comparable between runs, as different models are affected by the numerical density artifacts to differing degrees. Rather, the contour interval has been selected, in each case, to help bring out the relevant structure.

The difference between these two models is in the degree of flattening. The model of Figure 5-1 has a prolate bar, whose axial ratios k_b and k_c are both 0.262, in accordance with the findings of the surface photometry in Chapter IV. The bar employed in calculating Figure 5-2, by contrast, is extremely flattened, with k_b = 0.327 and k_c = 0.010. That is, the height of the bar perpendicular to the disk is only 1% of its length. Clearly, this model bar should indicate the limiting behavior as the triaxial bar approaches a two-dimensional form.

The density distributions of these models are markedly different in their centers. The flat bar is extremely effective in removing gas from the zone through which it rotates, so that the usual enhancement along the bar is not seen. In that respect, this is an atypical model among those using triaxial bars. The central hole is, indeed, so abrupt that the interpolatory program, which must be used to compute the

projected appearance of the galaxy, has obtained slightly negative "densities" in this region. The model with the prolate bar is a more representative example of the class.

In the region beyond the bar, the two models are not greatly different, despite the appearances of the grey tone displays. (The grey tone plot of the prolate model is saturated for the central enhancement.) As suggested by the density contour plots, only very mild asymmetries are present in the outer parts of the model galaxies. The flattened bar is slightly more successful at generating armlike features, although the appearance of this is enhanced by the particular values of the contour levels in the two figures.

The similarities and differences between the two models are more easily visible in their respective velocity contour plots. The greater disturbance of the flow in the center, for the case of the flattened bar, is obvious here. A telling similarity, however, occurs in the outer portion of the disk. By comparison with Figure 3-6, it is clear that neither of these models begins to approach the requisite degree of spiral structure in its kinematics. They share this key point with all other models we have computed which use a triaxial bar alone to give the nonaxisymmetric forcing in NGC 3359.

The two examples we have shown are illustrative of this class of models. We have produced a large number of such models with the free parameters (pattern speed, bar mass, and flattening) varying throughout, and even beyond, the range allowed by plausible arguments regarding the nature of the bar. Considerable effort was expended to correct the problems of excess disturbance at the center of the velocity field, and too little disturbance in the spiral arms. We have made substantial headway in neither. More importantly, any improvement in one symptom

comes at the price of aggravating the other. Thus, one way of adding more spiral structure in the outer disk is to increase the mass of the bar, relative to the disk. We have pursued this tactic to extremes approaching the absurd, with the bar having up to 70% the mass of the disk. Even in this case, the arm response, though stronger, was considerably too weak. At the same time, the disruption of the gas in the central region was violent. If the mass of the bar is decreased, the departures from axisymmetry become imperceptible, even in the unconvolved response. We have also tried decreasing the rate at which the density in the bar falls off, with much the same effect.

The situation is no better with respect to the other free parameters. If the pattern speed is increased, the well-known tendency of rapidly-rotating bars to evoke rings soon destroys the nonaxisymmetry of the response. More slowly-rotating bars eventually give rise to long, barlike forms with little winding. As for the flattening parameter, we have seen that making the bar flatter than the prolate case can increase the severity of the kinematical distortion in the bar zone, but does little to change the picture further out.

The findings outlined in this section may be summarized as follows. Based on an extensive survey of the available parameter space, we find no combination of input values which reproduces, even qualitatively, the gas response in NGC 3359, if the assumption is made that the nonaxisymmetric force is produced entirely by the observed optical bar. This statement is especially true when considering the velocity behavior of the galaxy. Models using a bar derived from the near infrared surface photometry exhibit noncircular velocities which are far too strong in the neighborhood of the bar, and too weak in the outer disk. The

cause of this behavior will be explored later in this chapter. Before passing on to discuss more successful models, however, we wish to emphasize that this negative result may be the most important finding of this chapter.

Models Using Oval Distortions

We have seen that hydrodynamical models of NGC 3359 which employ a triaxial bar derived from the observed surface photometry are unable to reproduce the essential kinematical behavior of the gas flow. In this section, we discuss details of models whose azimuthal forces result from oval distortions of the background disk. Both the driving potential and the gaseous response in this type of model are radically different from the triaxial case.

Oval distortion potentials are sometimes described as "massless bars." The meaning of this term is that a disk galaxy which has been modified by the addition of an oval distortion does not increase in total mass. Rather, the surface density present in each infinitesimal annulus of the galaxy has been redistributed in azimuth, so that both the surface density and the potential become doubly periodic functions of the azimuthal angle θ. The resulting disk-bar system has equipotential surfaces which are qualitatively similar to those of galaxies containing more conventional, "massive" bars, in that the potential minimum is bisymmetric, and its position angle is constant with radius in the disk. However, there are differences of detail which have important dynamical consequences. The most obvious is that an oval distortion, being a two-dimensional redistribution of the matter in a flattened disk, is itself a perfectly flat distribution. As mentioned in our discussion of models with triaxial bars, though, the degree of flattening of a bar does not appear to be of primary importance in its dynamical consequences for the gaseous disk.

To illustrate the more subtle features of the density distributions associated with oval distortions, let us consider the specific example used in the models of NGC 3359. Huntley (1977) demonstrated the following result for members of Toomre's (1963) family of galaxian disks, of which the infinite GMD (i.e., one not truncated after the manner of Hunter, Ball, and Gottesman, 1984) is an example. Let the surface density $\sigma_0(r)$ of the unperturbed disk be redistributed according to the equation

$$\sigma(r) = \sigma_0(r)[1+\varepsilon r^2(1+(\frac{r^2}{b_n^2}))^{1/2}\ (\beta^2+r^2)^{-k-3/2}\beta^{2k-2}\cos 2\theta]. \tag{5-13}$$

Here ε is an amplitude parameter, b_n the length scale associated with the Toomre disk of index n, β the length scale of the distortion, and k an arbitrary exponent. The total mass of the disk within a given radius is unchanged. Furthermore, the distorted disk remains an exact solution of Poisson's equation. Examples of hydrodynamical models of barred spirals which use potentials of this type may be found in Huntley (1977) and in Huntley, Sanders, and Roberts (1978).

The basic character of the potential arising from this mass distribution is a consequence of the dependence of the distortion on r and θ. At very large r, the density involved in the distortion falls off as r^{-2k}, but its behavior is much more complicated in the regime where r is of the order of b_n and β. In general, the decline in density of the nonaxisymmetric component will be considerably more gradual in this regime, which usually coincides with the portion of the galaxy where spiral structure is observed. This is a fundamental contrast between the

oval distortion and the triaxial figure, where the density residing in the nonaxisymmetric mode is identically zero outside the physical radius of the bar. It should also be noted that, since the azimuthal behavior of the distortion is uncoupled from the radial dependence, the degree of elongation of the figure is a function of radius.

We have investigated distortions of the form given by (5-13) with $k = 1$ and with $k = 2$. The functional forms for the potential and forces are much simpler for the $k = 2$ case. Unfortunately, this distortion also decays much more quickly with radius, and for the purposes of modeling barred spirals we have found the $k = 1$ form much more useful. All the oval distortions discussed in this chapter are of this type. The potentials and forces of both distortions are derived by Huntley (1977). For convenience, we reprint here the radial and azimuthal components of the perturbing force, f_r and f_a respectively, for the GMD with $k = 1$. The same forms, except for a constant scaling factor, apply to the Toomre $n = 1$ disk. Recalling that the mass of this disk is parameterized by a constant, C_0 (cf. equation 3-5), the expressions are

$$f_r = \frac{r\, C_0^2 \cos 2\theta}{3b_n}\left[\frac{3\beta(2\beta^2-r^2)(\beta^2+r^2)^{1/2} - (r^2\beta^2+4r^4)+6\beta^4}{[\beta+(\beta^2+r^2)^{1/2}]^3(\beta^2+r^2)^{5/2}}\right], \tag{5-14}$$

$$f_a = \frac{-2\, rC_0^2 \sin 2\theta}{3b_n}\left[\frac{\beta+2(\beta^2+r^2)^{1/2}}{[\beta+(\beta^2+r^2)^{1/2}]^2(\beta^2+r^2)^{3/2}}\right]. \tag{5-15}$$

In using this oval distortion to model an observed galaxy, one begins by selecting a disk whose rotation curve matches the data as closely as possible. There is less uncertainty in this step than for

models with triaxial bars, since the angle-averaged rotation curve is not affected by the inclusion of the oval distortion. As before, the model parameters which are not well constrained by the observations are the bar pattern speed and mass distribution, the latter being specified by ε and β. Once again, one proceeds by a series of numerical experiments, comparing the density and kinematical responses of the models to those of the observed galaxy.

As with the triaxial bars, we have computed a number of models of NGC 3359 based on oval distortions, varying the values of the input parameters. Although we have not been able to find a single model which agrees with the observations in every respect, the oval distortion models bear much more resemblance to the galaxy than those with triaxial bars alone, and the essential dynamics of the system are represented quite well by this class of model. We have, therefore, been able to place limits on the parameters which control the dynamics of NGC 3359. We will now proceed to discuss these results, together with some necessary caveats, in more detail.

The density response of the gas to oval-distortion bars is typically dominated by an open, trailing spiral which shows no sharp differentiation between bar and arms, but a gradual bending of the gaseous bar enhancement into a trailing spiral pattern. The spiral arms continue smoothly outward from the bar, with a decreasing pitch angle (i.e., they become more tightly wound at large radius), the whole resembling an open letter "S." Figure 1c of Huntley, Sanders, and Roberts (1978) is a good illustration of such a response. The strength of this spiral structure increases with the amplitude of the oval distortion, and also with the pattern speed, until the latter becomes so high as to produce a ring of material in place of

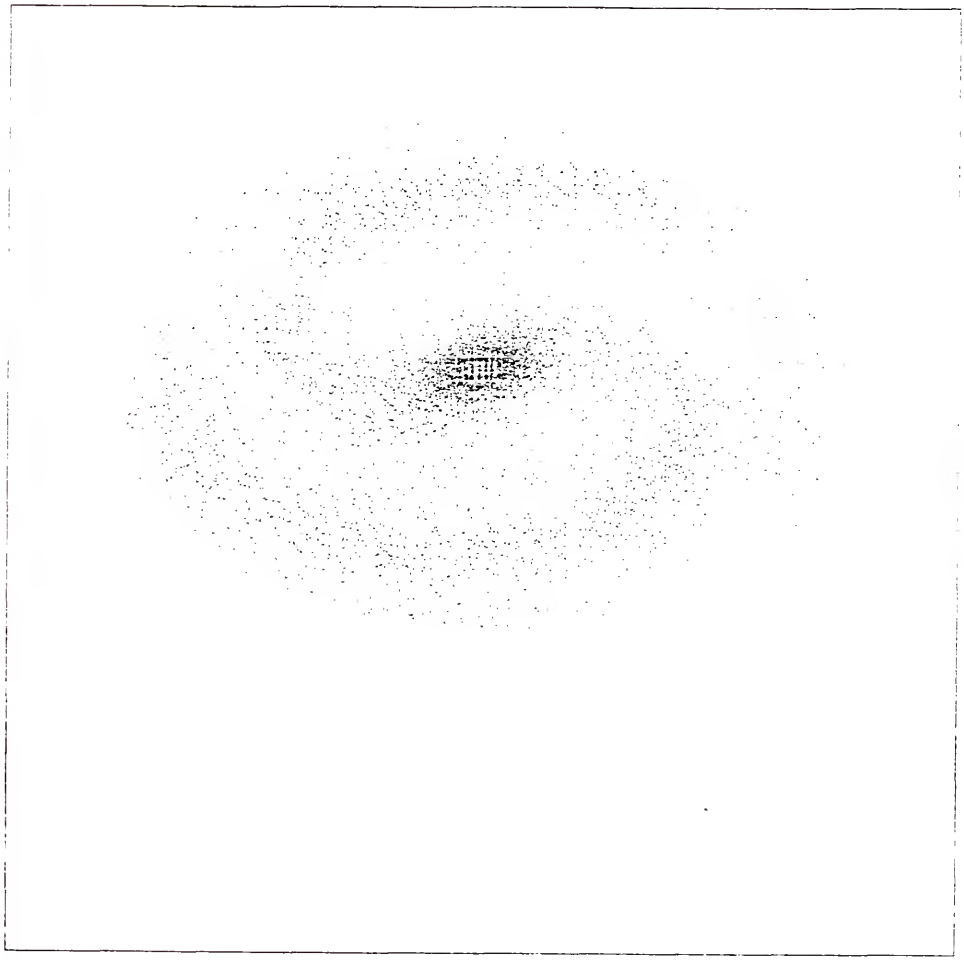
arms. Associated with the spiral pattern are large kinematical deviations from circular flow, which are especially apparent in the outer parts of the model galaxy. As there is no great physical distinction between the bar zone and the rest of the disk, the velocity disturbances in this zone are simply a smooth continuation of the overall flow pattern.

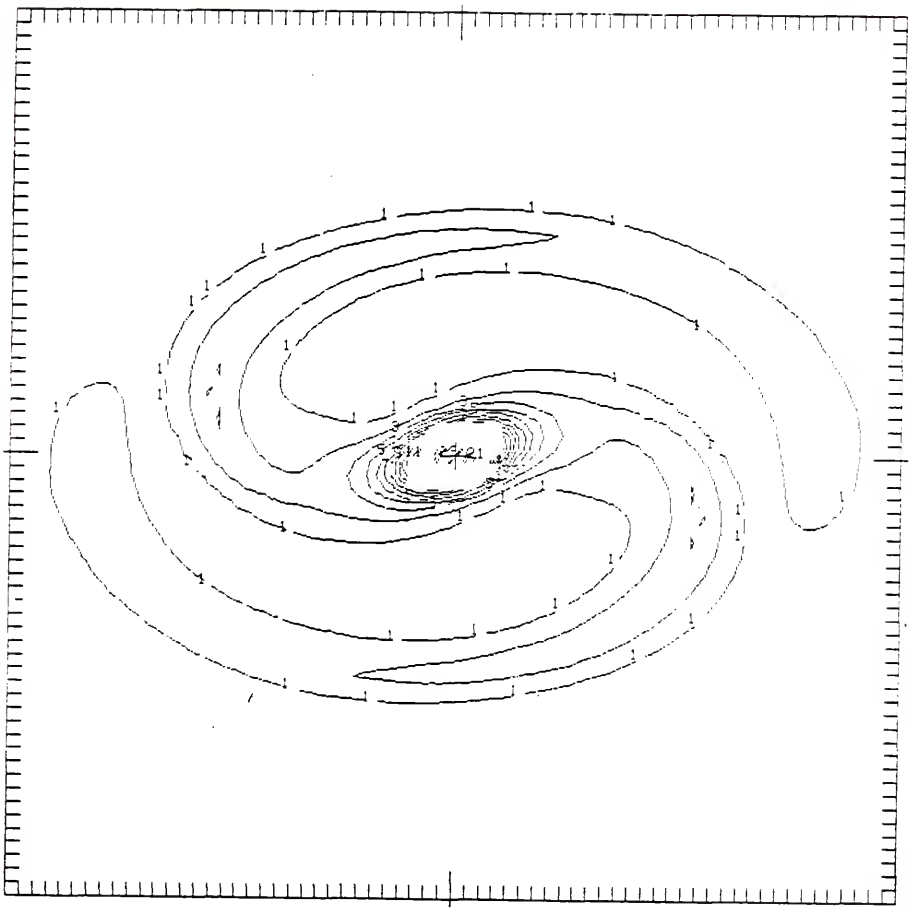
As an example of the behavior of oval distortion models, we show in Figure 5-3 the density and velocity fields computed in a typical run. The model illustrated has a GMD identical to the mass model of NGC 3359 which is described in Chapter III, and whose rotation curve is plotted in Figure 3-9: its length scale b_0 is 0.5, and its amplitude constant C_0 is 155 km/s. As in the triaxial models, the computational grid has been scaled to a maximum radius of 3.75. Superposed on this disk is an oval distortion of the type described above, with length scale $\beta = 0.82$ and amplitude $\varepsilon = 2.5$. The pattern speed at which this bar rotates, when transformed to an inertial reference frame, is 100 km/s/arc minute. Note that this pattern speed is about 10% faster than those of the triaxial models discussed in the previous section. In general, the oval distortions require slightly higher angular velocities to excite spiral structure.

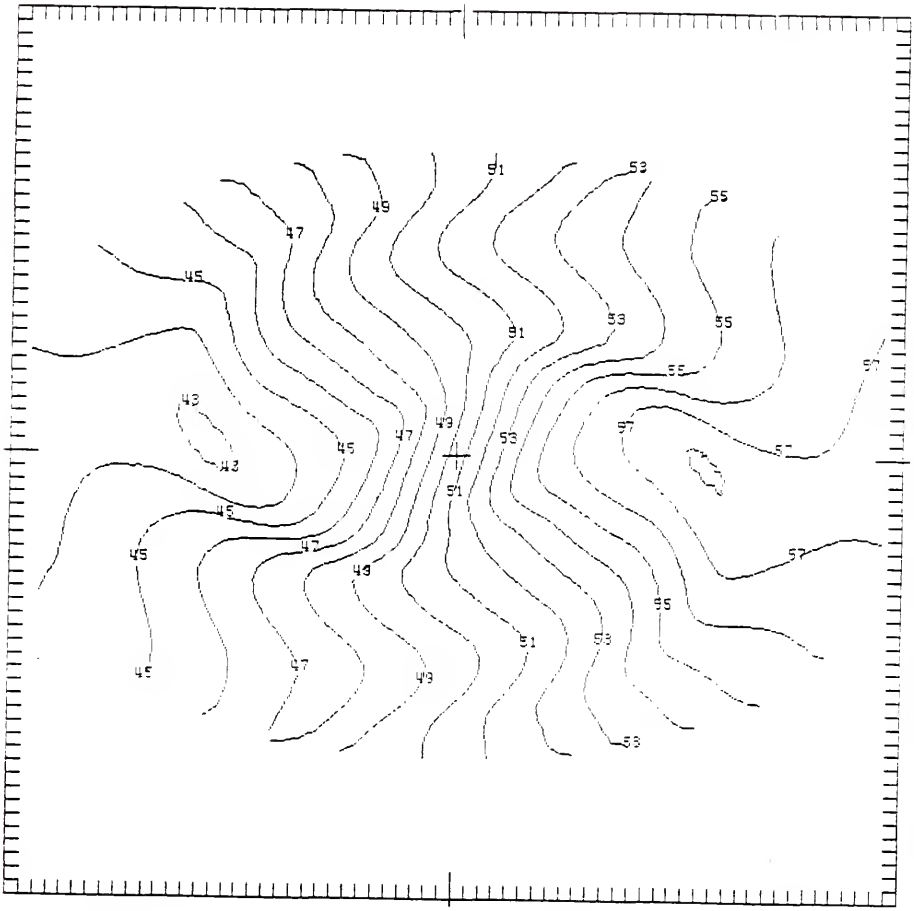
The characteristic S-shaped wave is readily evident in the density response of this model. Like the triaxial models, it exhibits a significant buildup of matter in the center of the galaxy, for the usual numerical reasons. The response in the outer galaxy, however, is quite different from that observed with triaxial bars. The spiral structure is, if anything, rather too marked, with an arm-interarm contrast of about 4.0, compared with a value around 2.5 in the observed galaxy. The reader will recall that the triaxial models, on the other hand, could produce only insignificant spiral enhancements in the region well beyond the bar.

Figure 5-3. Gas response in the pure oval distortion model of NGC 3359 discussed in this chapter. The identification of parts (a) through (c) are as in Figure 5-1.

NGC 3359 MODEL -- CONVOLVED






45
47
49
51
53
55
57
43

The spiral structure is also apparent in the model velocity field. Comparison of this velocity contour plot with Figure 3-6 shows that all of the principal features of the observed velocities are present, with perhaps slightly greater amplitude, in this model. In the region dominated by the bar, the contours are skewed somewhat away from the minor axis, but this effect is much more moderate than in the triaxial models, resembling the observed field in this respect. Additionally, the strong velocity shocks in this zone, which are so prominent in the velocity plots of the triaxial models, are absent here. This is again more consistent with the observations. Moving outward in radius, one encounters, in the velocity field of Figure 5-3, strong perturbations associated with the spiral arms. These kinematic distortions are of the same sense, and roughly comparable magnitude, as those observed in NGC 3359. The main discrepancy between the two has to do with the positions of the perturbations. The spiral arms in the model begin to be distinguishable from the bar at a significantly greater radius than in the observed galaxy, and become progressively more tightly wound as the radius increases. This behavior, as mentioned above, is common in oval-distortion models, but stands in contrast to the observed spiral pattern in NGC 3359. We found in Chapter III that, especially for the more well-defined, western arm, the pitch angle of this pattern is roughly constant with radius. Consequently, the most highly distorted contours along the major axis, the last which remain open, occur further out in the model than in NGC 3359; in both cases, they overlie the innermost portions of the spiral arms. At the same time, owing to its tight winding the spiral structure near the minor axis is found too close to the center in the model. However, these are differences of detail. It is immediately apparent that the kinematical behavior of this model is a

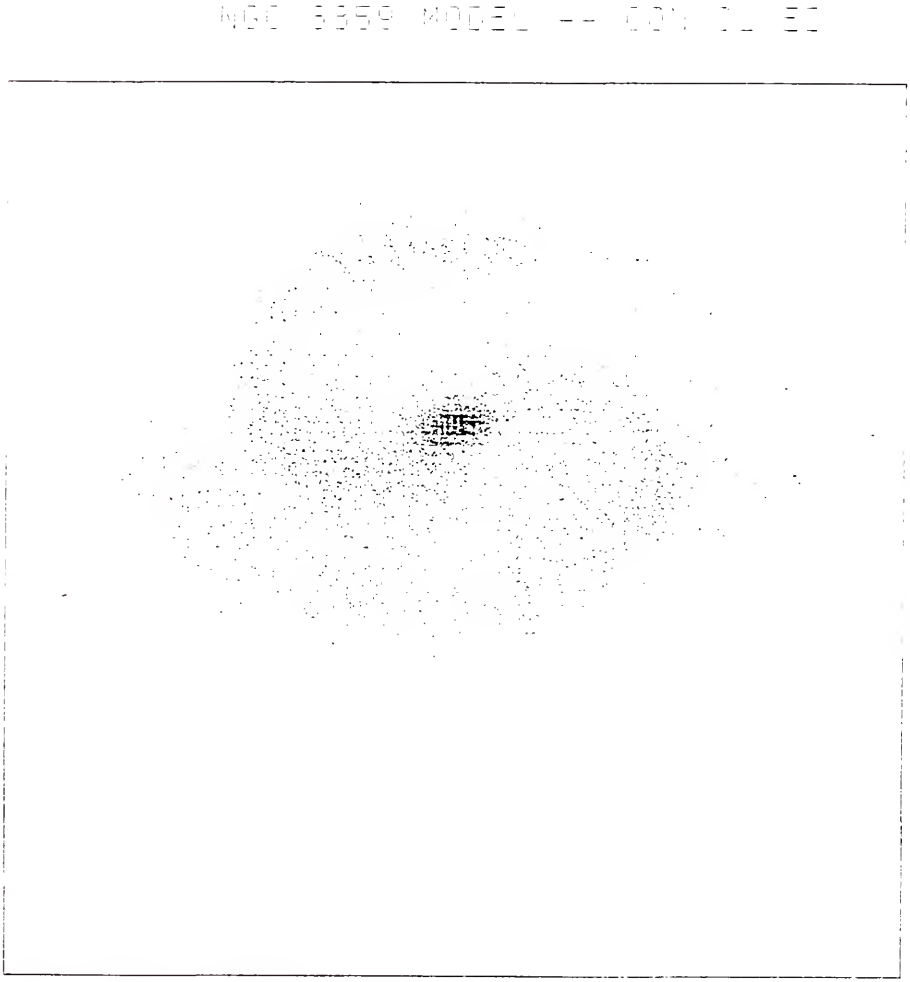
much more satisfactory approximation to that of NGC 3359 than we obtained with the models based on triaxial bars.

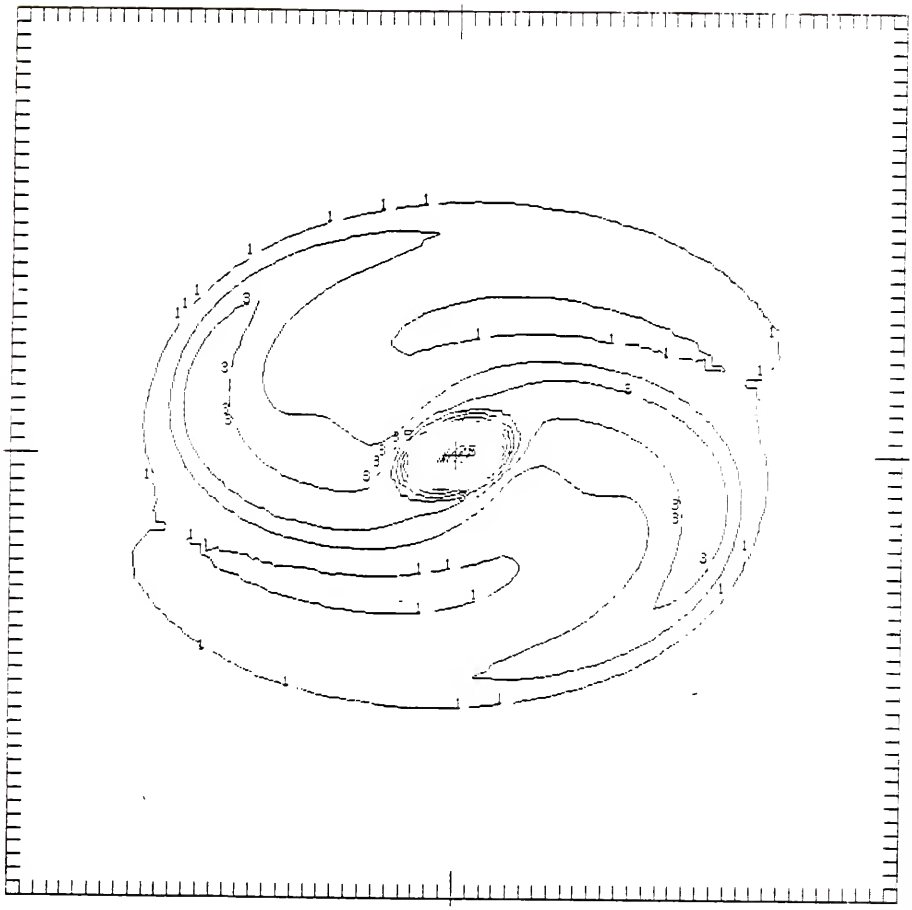
We have also computed a number of cases in which a triaxial component is placed at the center of a disk which is subject to an oval distortion. The principal reason for considering models of this kind is to reconcile, in some way, the oval distortion models with the surface photometry. Even if the stellar bar is not the driving agent for the nonaxisymmetric structure in the remainder of the galaxy, its existence must still be allowed for. One conceivable approach would be to postulate that the oval distortion produces a surface density in the galaxy's center which is related to the observed surface brightness of the bar. Even leaving aside the difficulties associated with mass-to-luminosity ratios, however, the shape of the isophotes rules out such an interpretation. Oval distortions which are sufficiently elongated to have surface-density contours with the proper axial ratio evoke much too strong a response in the gas: in every case, the resulting figure has an arm-interarm contrast of at least 20. Furthermore, their shapes remain very elongated, and in many cases even dumbbell-shaped, to radii of order two arc minutes (6.4 kpc) or more; such a system would hardly merit the designation "disk galaxy" at all. Therefore, we favor the view that the optical bar is indeed a "massive" object, which is in all probability a dynamically separate component of the system, with a configuration not unlike that derived by Stark's method. However, it is no longer seen as the principal mechanism for providing the nonaxisymmetric forces in the galaxy. Additionally, its mass relative to that of the disk must be less than assumed in the models which employed a triaxial bar only, in order to suppress the associated disruption of the central isovels.

The models displayed in Figures 5-4 through 5-7 are all of this kind. The first three of these are related models, chosen to illustrate the effects of changing certain dynamical parameters. All three use precisely the same disk and oval distortion as employed in the pure oval distortion model just discussed. Each of them also contains, at its center, a triaxial bar whose density is just as prescribed by equation (4-12). In other words, the form of this triaxial bar (which is identical in the three cases) is the same as in the purely triaxial runs, but its total mass has been scaled down, and is only 3.9% of the disk mass interior to the maximum modeled radius of 3.75. This latter mass is $5.9 \times 10^{10}\ M_{\odot}$ for the GMD being used here, so that the mass in the triaxial bar is $2.3 \times 10^{9}\ M_{\odot}$.

The model shown in Figure 5-4 also has the same pattern speed as in the oval distortion model; thus, differences between these plots and those of Figure 5-3 are due entirely to the addition of the triaxial component. Several subtle effects are in evidence. The density pattern in the bar area is altered slightly; the bar is more readily identifiable as a distinct entity, with a perceptible change in direction of the enhancement at its end, especially on the trailing edge. This effect is considerably stronger in the unconvolved response. The straight, barlike response in the gas is also somewhat shorter. The triaxial bar is thus beginning to show signs of affecting the gas density as it did in the purely triaxial runs, which would improve the oval distortion model significantly. The physical galaxy, as we have seen in Chapter III, does have some depletion of H I in the bar zone, and the arms certainly do not continue smoothly from the bar. Unfortunately, triaxial bars which are sufficiently massive to begin sweeping gas from the center give rise to significant distortions

Figure 5-4. Gas response in the first of the hybrid oval distortion and triaxial models of NGC 3359 discussed in this chapter. Except for the addition of the triaxial component, this model is identical to that shown in Figure 5-3. The identification of parts (a) through (c) are as in Figure 5-1.





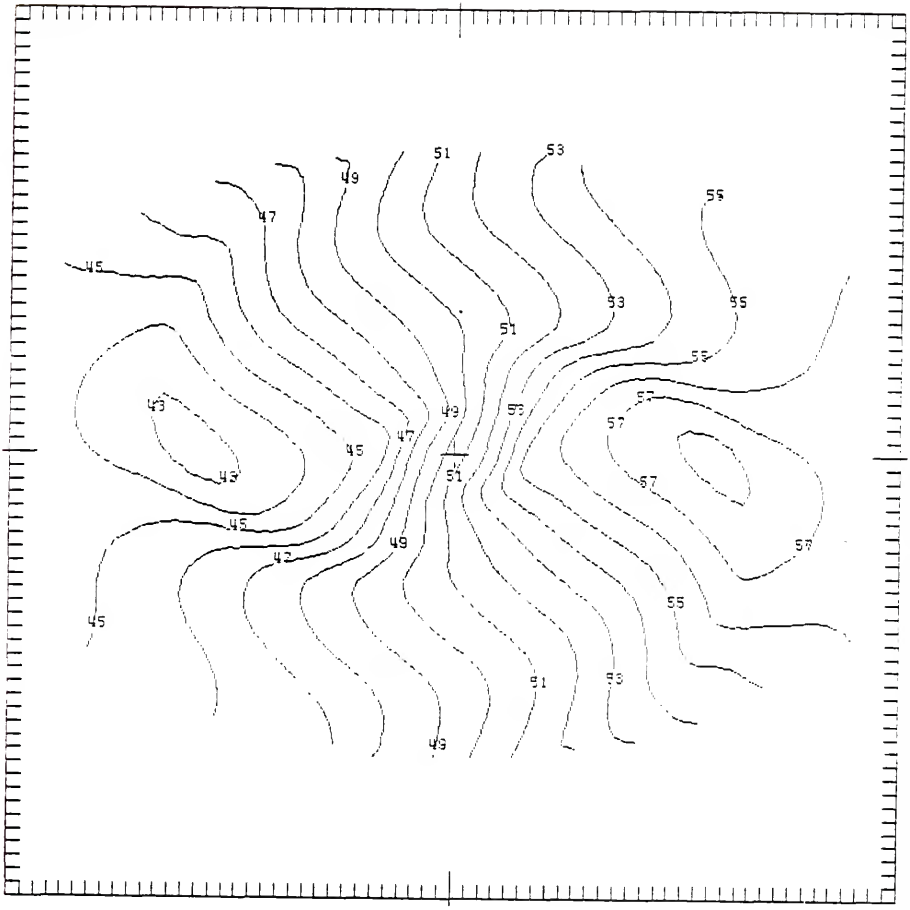
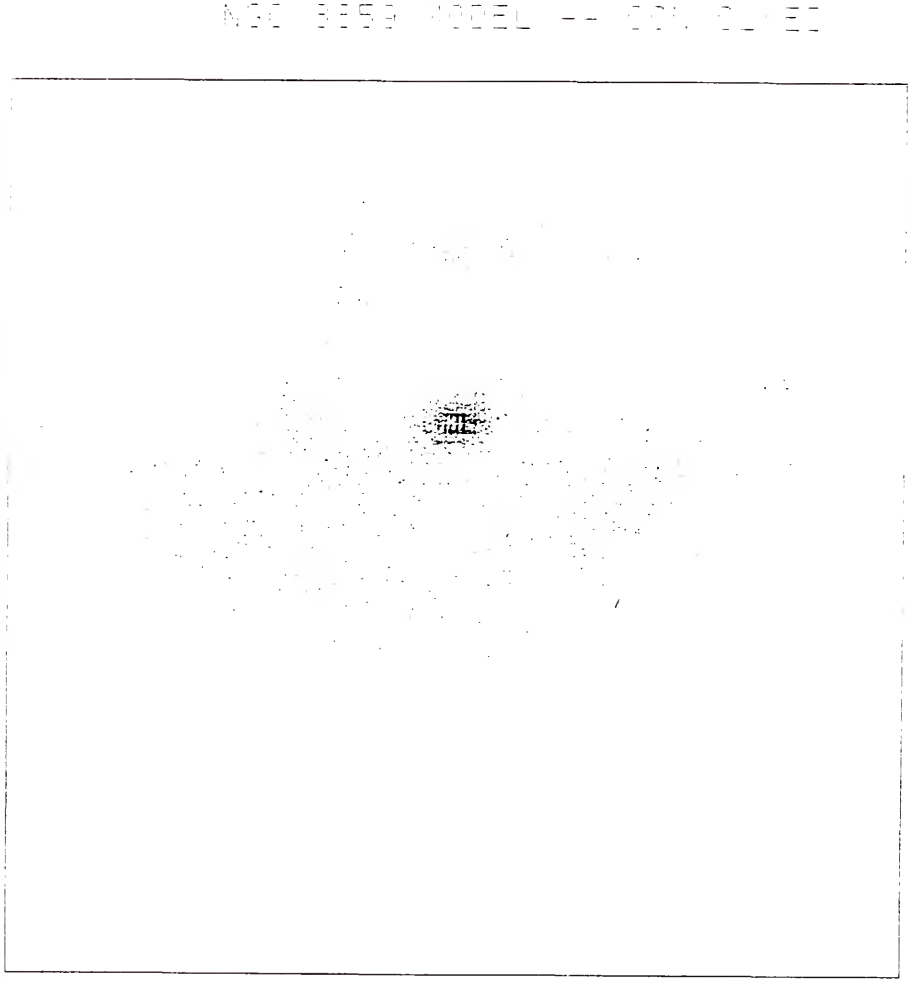
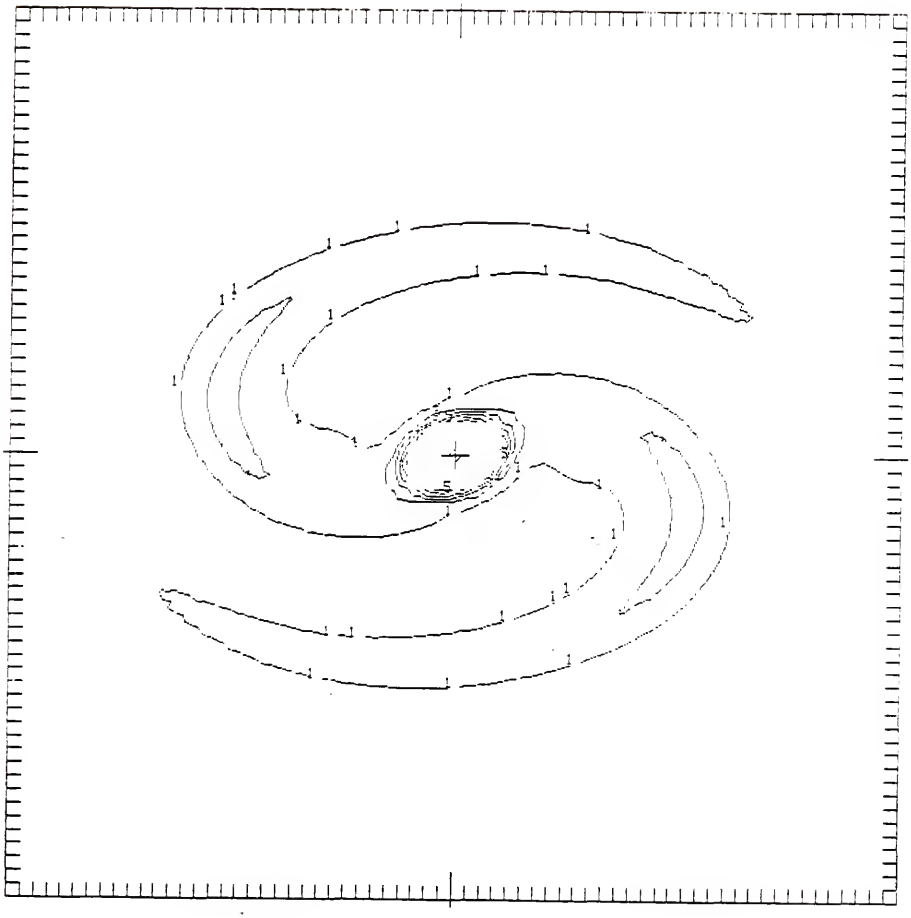

45
47
49
51
53
55
43
45
47
49
51
53
55
57
42
45
49
51
53
55
57
59
49

Figure 5-5. Gas response in the second of the hybrid oval distortion and triaxial models of NGC 3359 discussed in this chapter. This model is identical to that illustrated in Figure 5-4, except that the pattern speed is higher by 8.5%. The identification of parts (a) through (c) are as in Figure 5-1.





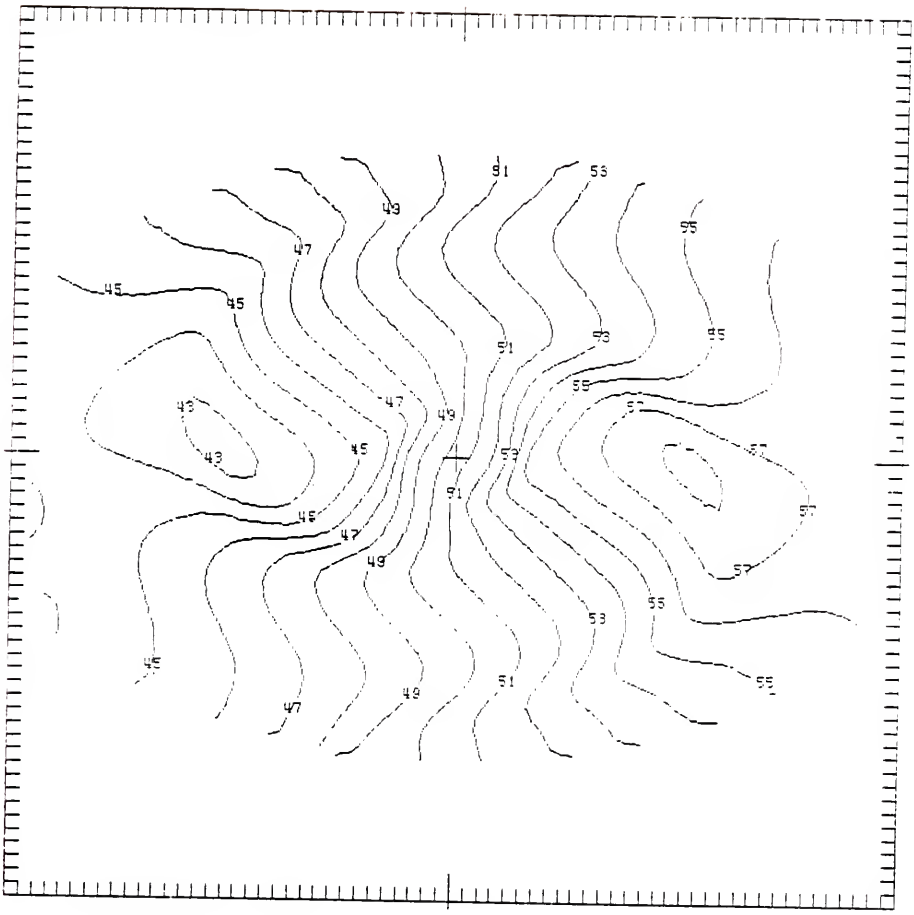
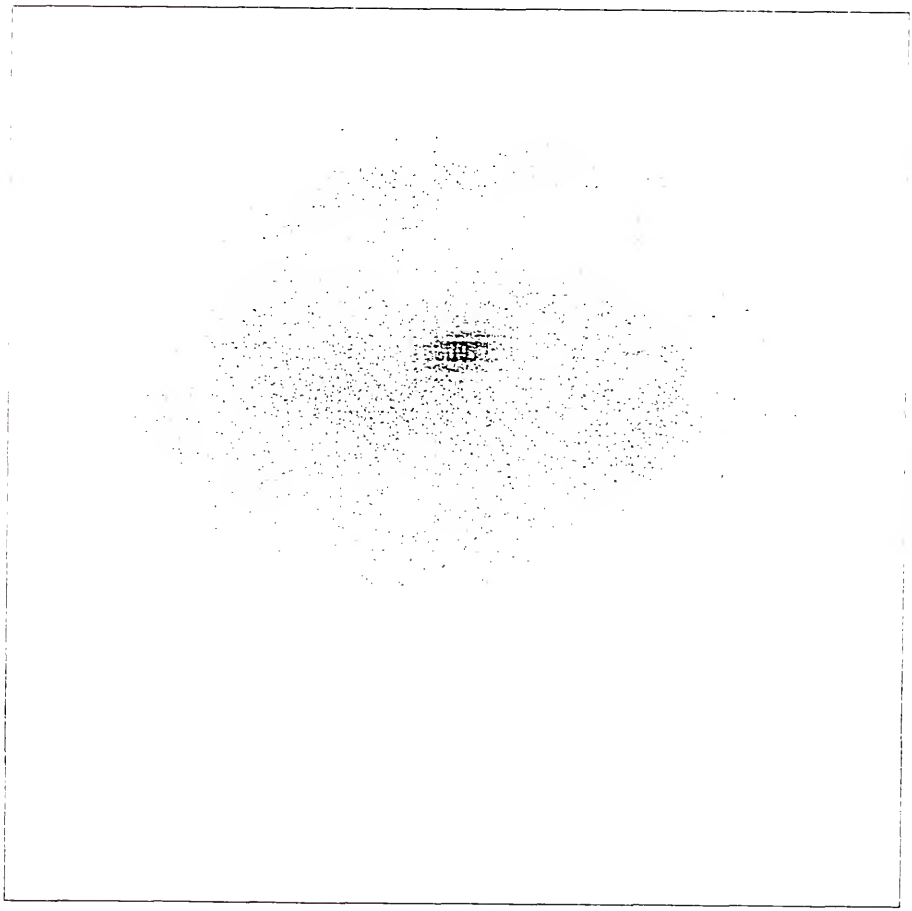
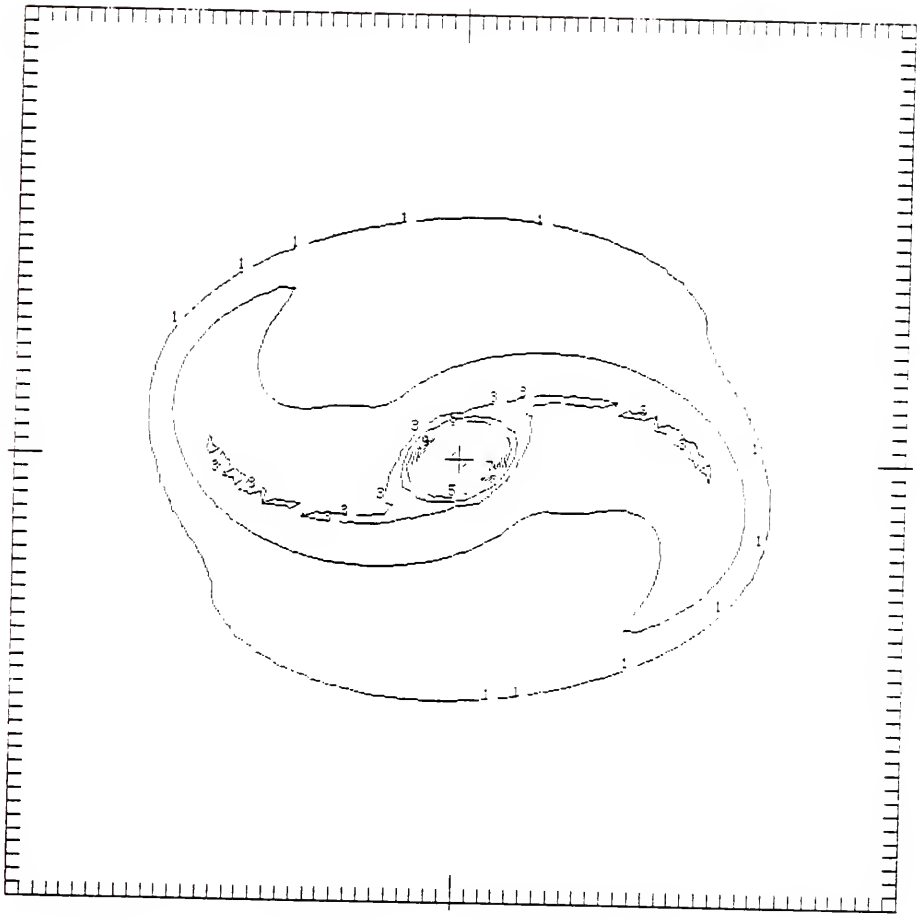

45
45
47
49
51
53
55
43
43
45
47
49
51
53
55
57
45
47
49
51
53
55
57

Figure 5-6. Gas response in the third of the hybrid oval distortion and triaxial models of NGC 3359 discussed in this chapter. This model differs from that shown in Figure 5-4 only by having the amplitude of the distortion reduced by 40%. The identification of parts (a) through (c) are as in Figure 5-1.





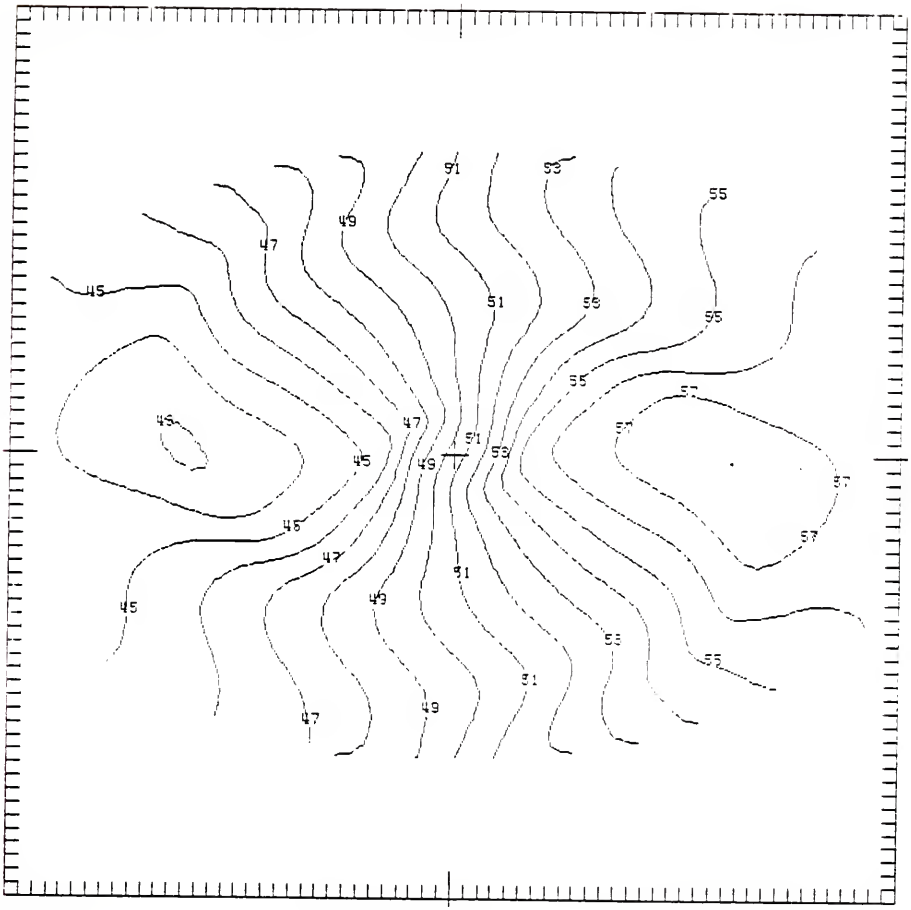
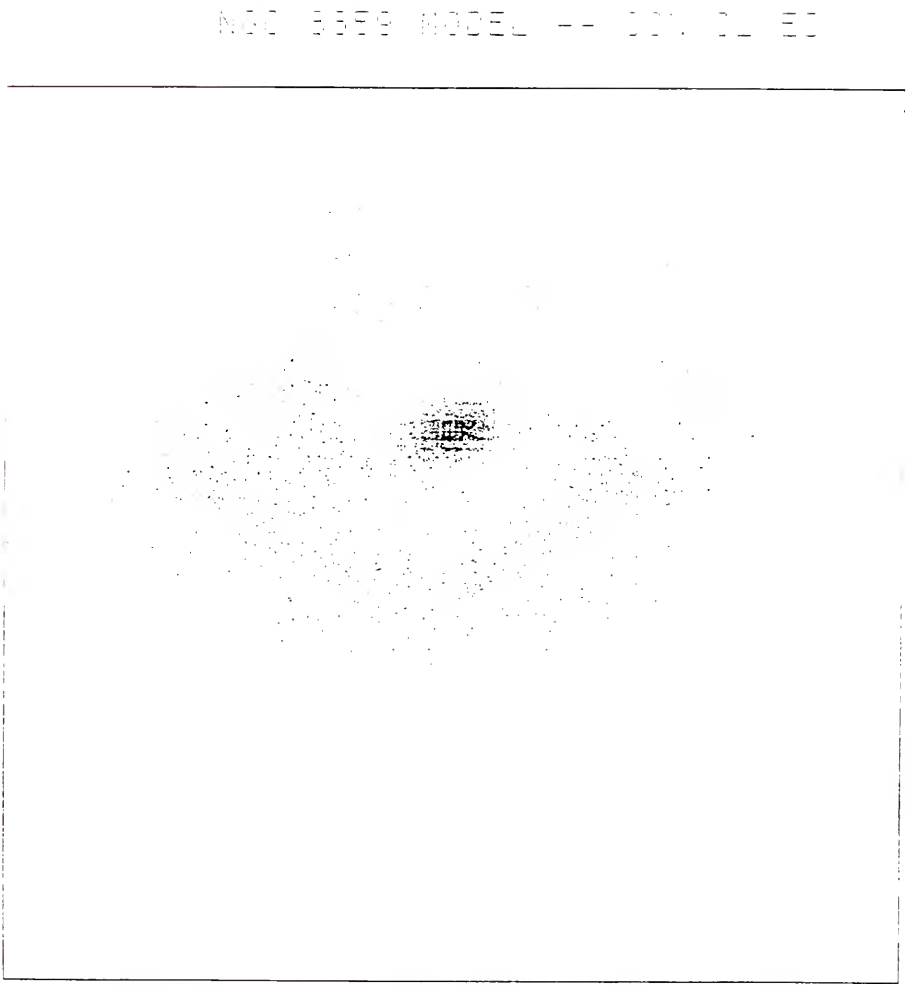
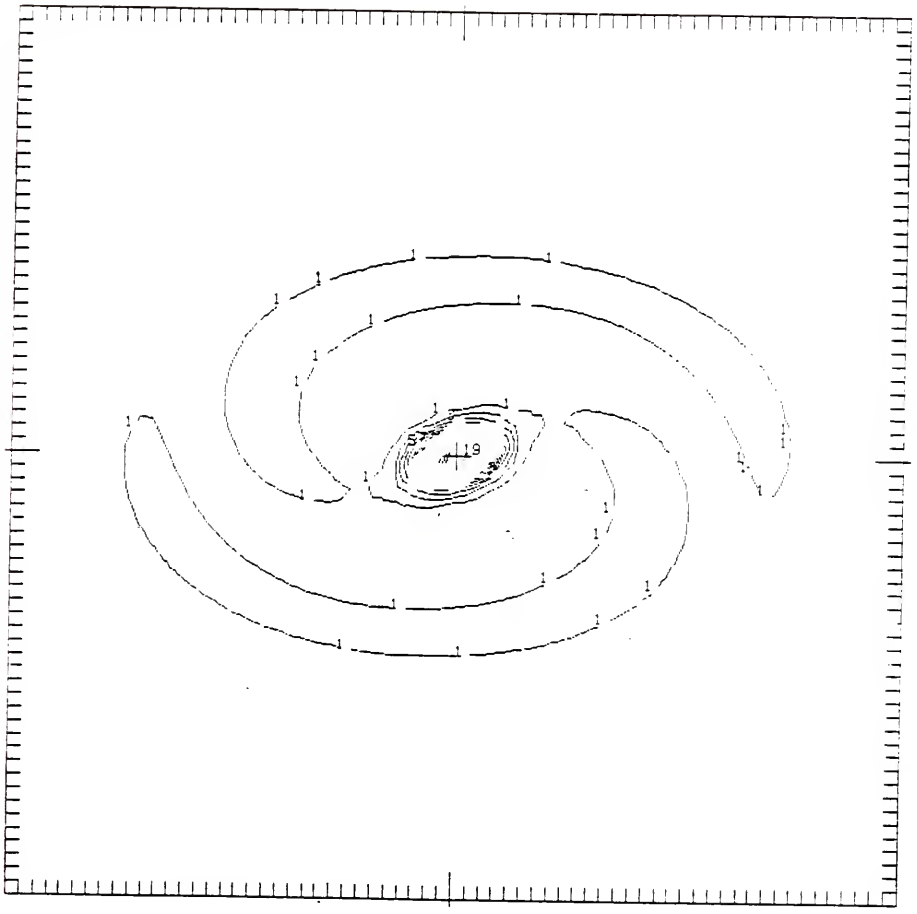
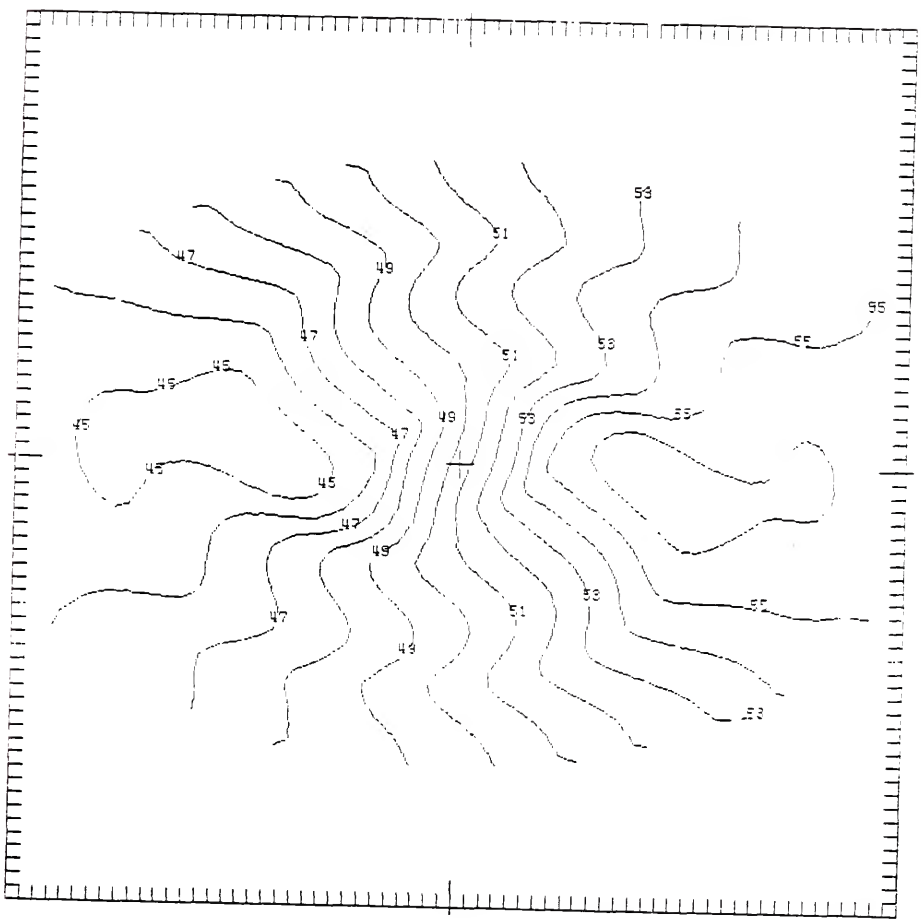

45
47
49
51
53
55
57

Figure 5-7. Gas response in the final hybrid oval distortion and triaxial model of NGC 3359 discussed in this chapter. This model is identical to that of Figure 5-5, except that the disk mass has been lowered slightly to match the observed rotation curve. The identification of parts (a) through (c) are as in Figure 5-1.






45
45
45
45
45
47
47
47
47
47
49
49
49
49
51
51
51
53
53
53
53
55
55
55
55

in the central velocity field. Bars as little as 50% more massive than this one are unacceptable on that account. As one would expect, there is little change in the velocity field of Figure 5-4, as compared to that of the oval distortion alone.

Figure 5-5 presents a model which differs from the previous one only in pattern speed. It is axiomatic, in barred spiral models, that the response one obtains is more critically dependent on the pattern speed than on any other parameter (Huntley, Sanders, and Roberts, 1978; Schempp, 1982). The type of model discussed here is no exception. The run shown in Figure 5-5 was calculated using a pattern speed of 108.5 km/s/arc minute, just 8.5% faster than that of Figure 5-4. The spiral response is significantly enhanced. The density distribution shows longer, less open, and more well-defined arms. To put this statement on a more quantitative footing, we note that the arm-interarm contrast in the gas density response, which was less than 3.0 in the previous model, has increased to about 5.0 in the model depicted here. The change also is clearly seen along the minor axis in the velocity plot, where the strength of noncircular velocities has been increased.

Figure 5-6 shows a model in which we have varied the oval distortion amplitude, which is second in importance only to the pattern speed in determining the response in models of this kind. Here the amplitude has been reduced by 40% from the run illustrated by Figure 5-4, to which this model is otherwise identical. The effect upon the density distribution is obvious; the amplitude and, especially, the length of the spiral arms in the gas are drastically reduced. At radii where the previous two models showed definite spiral structure, the densities in this case remain close to axisymmetric. Note, however, that the kinematic disturbances

associated with the spiral arms are still seen in the velocity plot. The contrast between these velocity perturbations for a relatively weak oval distortion, and those shown earlier in the chapter for the triaxial bars, is instructive.

We have now shown explicitly, if rather cursorily, the effects of changing two of the three unknown input parameters in models whose nonaxisymmetric forces arise primarily from oval distortions: namely, the angular velocity and amplitude of the distortion. A fairly wide sampling of the parameter space of those two quantities confirms that the dependence of the gas response on these two parameters is as one would expect from previous studies of barred spiral models. The pattern speed remains the single most dominant input parameter; a change of more than about 20% from its optimum value leads to a model galaxy which is unrecognizable as being related to the observed object. Because of this sensitivity of the models to pattern speed, it can be determined with more confidence than the other dynamical parameters which emerge from the modeling process. As with the triaxial models, we find that pattern speeds which are too low evoke no winding in the gas response. On the other hand, too fast a rotation of the bar smears the response into a ring. The role of the distortion amplitude is simple: it is related directly (but nonlinearly) to the amplitude of the nonaxisymmetric structure produced in the gas.

The remaining unknown parameter is the length scale of the oval distortion. It is difficult to attach a physical interpretation to this scale, which has no direct relation to the optical bar length. The scale of the distortion is simply an indication of how the amplitude of the nonaxisymmetric forcing falls off with radius. For models of NGC 3359 whose

oval distortion is specified by equation (5-13), the length scale cannot be varied by more than about 25% in either direction from the adopted value of 0!82. Longer and shorter scales produce model galaxies with too much and too little spiral structure in the velocity field, respectively. As we have said, these models employ the k = 1 distortion, which drops off less rapidly than the k = 2 version. The effects of choosing the latter are qualitatively similar to those when a shorter length scale is used in the k = 1 case.

The last model to be presented in this section is shown in Figure 5-7. This model represents the best compromise among those we have computed, as a fit to the density and velocity distributions of NGC 3359. It differs from the model shown in Figure 5-5 only in having a slightly reduced disk mass, to compensate for the addition of the triaxial bar to the rotation curve. The tendency of the earlier model toward a more well-defined break between the bar and the arms has clearly been accentuated. The spiral arms are also longer, winding through more than 180°, as in NGC 3359. The arm-interarm contrast is somewhat higher than observed, approaching a value of 4.0; indeed, it should be noted that all of the oval distortion models that we have calculated show at least a small excess in this quantity. It is likely that a very small decrease in the pattern speed (108.5 km/s/arc minute in this model) would help to alleviate this problem, and make the arms less tightly wound, as well, which would accord better with the observations. However, such a change would also reduce their extent.

The most important differences between the detailed kinematics of this model and the observed galaxy come about because the positions of the spiral features are not precisely matched. Like every oval distortion

model calculated, this one gives arms which begin too far from the center of the galaxy and then become too tightly wound in its outer parts. As a result, the relative positions of the associated velocity features do not agree with those observed. Also, the model shows overly strong perturbations at the very edge of the grid. This may also be an indication that the pattern speed of this model, which places corotation at a radius of approximately 1:2, is slightly higher than optimal.

In Figure 5-8, we plot the rotation curve of this model. Its behavior is gratifyingly similar to that obtained for NGC 3359 from the neutral hydrogen data (cf. Figure 3-7). The inner region shows essentially solid-body rotation, rising to a value slightly greater than 120 km/s at about 1:4. The curve then flattens for a short distance, before rising again as the interarm region is traversed. Allowing for the differences in the positions of the spiral features, this behavior is very like that of the observed curve. Beyond a radius of 3', the model curve decreases with considerably more rapidity than the observed curve ever does. This phenomenon has been observed in previous models, and is an artifact of the approach to the edge of the grid. When substantial noncircular velocities are present near this edge, the angle-averaged velocity is lowered by the loss of higher-velocity material across the transparent outer boundary. The magnitude of the effect in this case is another confirmation that the noncircular velocities are slightly too large in the very outer parts of the model galaxy.

Taking into account these sundry pieces of evidence, as well as the results of models at higher pattern speeds, we regard 108.5 km/s/arc minute as near the upper limit that can reasonably be allowed for the bar pattern speed in NGC 3359. The lower limit is found to lie around 95

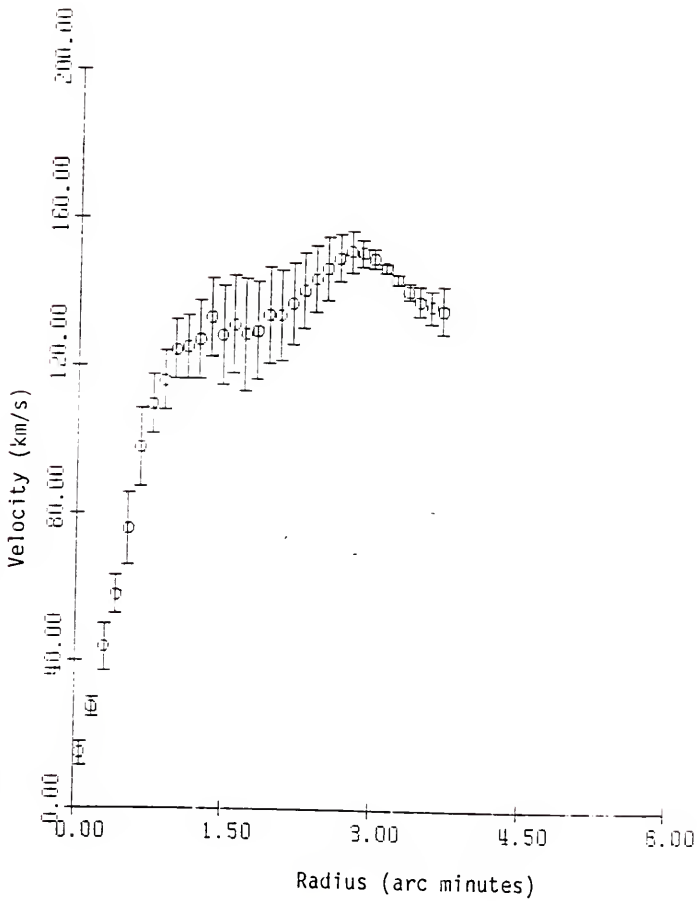


Figure 5-8. The rotation curve derived from the model illustrated in Figure 5-7. This curve has been produced by the same angle-averaging algorithm used to generate the observed rotation curve of Figure 3-7. The error bars show plus and minus one standard deviation; these formal errors are due entirely to the presence of noncircular velocities in the model galaxy.

km/s/arc minute. These results depend to some degree on the nature of the bar used: the most probable value could vary by a few percent for bars whose density distributions are significantly different than the oval distortions we have considered in this chapter.

We would like to close this chapter with a short discussion of the implications of our modeling results. The oval distortion models are somewhat unsatisfactory, not only because there remain differences in detail between their predictions and the observed densities and kinematics, but also because the approach itself is rather ad hoc when compared to the use of traixial bars derived from the photometry. Direct observational support for the potential used in an oval distortion model is nil, while the logic of the triaxial approach seems almost inescapable, building steadily from the observed surface brightness to the form of the triaxial potential. Why, then, do the oval distortions succeed where the triaxial bars fail?

The answer lies in the radial dependence of the forces in the two cases. The crucial factor for evoking a trailing spiral response in a gaseous disk is the azimuthal component of the force (Huntley, 1977). Let us consider the behavior of this component in the oval distortions and in the massive bars. In the oval distortion model of Figure 5-7, the maximum value of the tangential force, relative to the local force associated with the disk, is 35%, at the galaxy center. If one evaluates the equivalent ratio at a radius of 2.8, three-quarters of the way to the edge of the computed grid, one obtains a value of 7%. For a typical triaxial case such as illustrated early in this chapter, these numbers are, respectively, 50% and 0.2%. Simply put, the material orbiting well away from the bar zone sees the bar as being simply a point mass at the galaxy center.

Given the small radial extent of the optical bar in NGC 3359, this result is inevitable for any remotely plausible distribution of mass in the bar.

The failure of the triaxial models may well be the most important conclusion of this study of the gas dynamics in NGC 3359, because it is the most definite. Despite being an appealingly direct and elegant approach to the problem, the use of a triaxial bar based on the surface photometry does not work in the hydrodynamical models. The azimuthal force contributed by this well-defined bar is simply insufficient to produce the observed spiral response in the outer parts of the disk. Some additional source of nonaxisymmetric forcing must be present at large radii. In this study, we have used an oval distortion to supply the needed force.

The question arises as to whether there might be some other source of maintenance of the nonaxisymmetric force in the outer parts of the galaxy, whose kinematical effects are mimicked by the oval distortion. The obvious candidate is, of course, the gravitational attraction of the gas itself. If the relative density of the gas is high enough, this could well be an important factor in the gas dynamics. In a scenario of this kind, the role of the optical bar might be to provide the weak initial disturbance which is amplified by the gas response, as well as to maintain the spiral structure by continually providing rotational energy to the dissipative gaseous system. The alternative to the self-gravity hypothesis is that the galaxy as a whole is indeed slightly nonaxisymmetric to quite large radii, with the optical bar perhaps serving as a tracer of this distribution.

It would be very desirable to test this hypothesis by numerical simulations which include the self-gravity of the gas. There are two

possible approaches to such an investigation. The first, three-dimensional N-body calculations, would require large amounts of computer time on advanced parallel processors. The less costly approach would be to perform calculations like those reported here, with the self-attraction of the gas included. Unfortunately, as mentioned earlier in this chapter, such computations do not give rise to stable configurations unless the input potential and pattern speed together form a truly self-consistent equilibrium figure. Aside from the numerical model of Miller and Smith (1979), which was investigated in this manner by Huntley (1980), no such figures are known for barred spiral galaxies. Discovery of additional equilibrium figures of barred spiral galaxies would be very useful in future studies of their dynamics.

CHAPTER VI
SUMMARY OF CONCLUSIONS

Observational Results

The principal conclusions of this study are of two kinds: those derived directly from the 21 cm observations of NGC 3359, and those which depend on modeling and theoretical interpretation. The former are presented in detail in Chapter III, the latter primarily in Chapter V. In this final chapter, we shall give a brief summary of these conclusions, beginning with the observational results.

NGC 3359 has been observed at 21 cm, using the Very Large Array, with a total integration time of more than twelve hours on-source. After calibration and reduction, we have obtained a set of single-channel maps, at a velocity resolution of 25 km/s, and a spatial resolution of 18", or 0.95 kpc at the assumed distance of 11 Mpc (de Vaucouleurs, 1979). The rms noise in these single-channel maps is 2.13 K. Integrated surface density and mean velocity maps were generated by integrating the channel maps.

The total neutral hydrogen mass detected in the galaxy is 4.9×10^9 $M_\odot$. This mass is about 20% lower than detected by the single-dish studies of Rots (1980) and of Fisher and Tully (1981), despite considerable effort to detect emission of low surface brightness. While this could be, in part, an effect of the absence of spacings less than about 200 wavelengths in our data, the more likely explanation is that there is still a broad, low surface brightness component below our detection threshold.

The distribution of neutral hydrogen across the disk of NGC 3359 is complex, with local maxima and minima alternating in a manner which initially seems almost chaotic. However, when one examines the distribution of H I at column densities greater than about half the peak value (3.51×10^{21} cm^{-2}), a clear-cut, grand design spiral pattern emerges The two principal arms of this pattern coincide with the optical arms, except that the H I arms extend somewhat beyond the latter. This correlation of positions is clearly shown by superposing the H II regions catalogued by Hodge (1969) on the neutral hydrogen density map. The most significant difference between the H I spiral structure and the optical appearance is that, in the former, the two principal arms are of nearly equal strength and extent. Thus, the striking asymmetry of the galaxy in blue photographs (cf. Figure 1-1) does not reflect the distribution of atomic gas.

The most prominent feature at 21 cm, apart from the two principal arms, is an amorphous enhancement in the galaxy's northwest quadrant. This feature has the appearance of a series of irregular, quasi-spiral spurs on optical photographs, and these spurs contain several of Hodge's H II regions. In contrast to the spiral arms, the optical bar does not have an associated neutral hydrogen feature. Instead, the zone dominated by the bar is somewhat depleted in H I, suggesting that the gas in this region may have been partially swept out by the bar's rotation. This result is similar to that observed in the barred spiral NGC 3992 by Gottesman et al. (1984).

The bulk of the 21 cm line emission in NGC 3359 arises from a disk of radius approximately 4!1, or 13.1 kpc. The disk edge is rather abrupt around most of its circumference, but, at a few positions,

outlying material of slightly lower surface brightness is joined to the H I disk. This outer material is very irregular and knotty in distribution. On the largest scale, it seems to be distributed in two or three ragged, armlike features, although we cannot rule out the possibility that a broken, irregular ring of low density is present. Based on a consideration of its kinematics, we conclude that this filamentary outer material probably occupies the same plane as the well-defined neutral hydrogen disk.

The 21 cm velocity field of the galaxy is dominated by rotation, but modified by strong disturbances whose positions are well correlated with the spiral structure. The velocity contours in the neighborhood of the bar are skewed across the galaxy minor axis, toward the direction parallel to the bar. Additionally, distortions of large amplitude are located at the positions of the principal arms. These features are taken to indicate the presence of substantial streaming velocities associated with the spiral pattern. They lend support to the hypothesis that the dynamics of the galaxy are fundamentally bisymmetric, since the amplitudes of the velocity distortions of the two arms are approximately equal.

The rotation curve of NGC 3359 behaves very smoothly at radii near and beyond the edges of the photometric and H I disks, implying that there is no sharp truncation of the total disk surface density (Casertano, 1983; Hunter, Ball, and Gottesman, 1984; Gottesman et al., 1984). At large galactocentric distances, the rotation curve is virtually flat, with perhaps a very gradual decrease from a maximum deprojected velocity of about 150 km/s. This behavior continues from a radius of about 3!5 to our last measured point, at 6!75 or 21.6 kpc. The near flatness of the rotation curve implies that a considerable mass

must be present beyond the optically detectable features, and even beyond the radius of the well-defined H I disk. The total mass derived from the rotation curve, interior to 21.6 kpc, is 1.2×10^{11} $M_{\odot}$.

Comparison of the H I and total masses shows that some 4% of the mass of NGC 3359 is in the form of atomic hydrogen. This value is fairly typical for an isolated galaxy of its type, as is the value of 11 for the total mass to blue luminosity ratio in solar units.

In the course of our 21 cm observations, we have detected a small, previously unknown companion galaxy to NGC 3359. An exceedingly faint, blue image is visible on the Palomar Sky Survey prints at the location of this object. The projected separation from NGC 3359 of the presumed satellite is 47 kpc, while its relative velocity is about 45 km/s in approach. The total neutral hydrogen mass in the satellite is measured to be 4.6×10^{7} $M_{\odot}$. Its total mass is quite difficult to estimate, since there is only very marginal evidence for rotation in its velocity field. An estimate based on its global spectrum, after correction for the instrumental profile, gives 1.5×10^{8} $M_{\odot}$ for the mass, very roughly. At any rate, the mass seems definitely too low for the satellite to have an important gravitational or tidal influence on NGC 3359 itself.

Hydrodynamical Models

We have modeled the observed gas response in NGC 3359 with a well-known hydrodynamical code, the "beam scheme," a two-dimensional version of which has been made available to us by Dr. J. M. Huntley. Our immediate object in these models was to attempt to reproduce the detailed density and kinematics observed for the galaxy. The latter are especially important in evaluating models of this kind, as their density fields are subject to certain artifacts which have to do with the nature of the

two-dimensional grid. The more fundamental purpose of this program is to investigate the validity of the assumptions commonly made in published numerical models of barred spiral galaxies.

The results of the hydrodynamical models employed in this study are, broadly speaking, determined by the following input parameters: the masses and mass distributions of the bar and of the axisymmetric components, and the angular velocity at which the bar stirs the gaseous disk. The mass distribution of the axisymmetric components of the galaxy (a disk and possibly a halo) is fairly well determined from the rotation curve. The precise nature of this distribution is, in fact, unimportant for our purposes, as long as the observed rotation curve is well fit. The most important variables among the model parameters are, therefore, the characteristics of the bar. We have used two very different forms for the bar potential.

In the first approach, the bar forces were derived from near infrared surface photometry of the stellar bar of the galaxy, following the technique developed by Stark (1977). This procedure, presented in detail in Chapter IV, fits the isophotes with an inherently triaxial bar. The orientation and relative mass density of this bar can be determined directly from the observations, except for one unknown projection angle, which then becomes the controlling variable in a one-parameter family of models of the observed bar. The primary physical difference between members of this family is the ratio of the shorter two axes of the triaxial figure. Experimentation has revealed that the results of hydrodynamical models incorporating different bars from a single family show little difference in the gas response, provided that the bar mass remains constant. Therefore, with a potential of this kind, the free

parameters for hydrodynamical modeling, after fitting the surface photometry, are the bar pattern speed and its total mass, relative to that of the underlying disk. The gravitational force arising from the triaxial bar is computed by means of newly derived expressions, given in Chapter IV.

The other type of bar potential employed was an oval distortion of the underlying disk. Unlike the triaxial bar, the oval distortion does not add mass to the system. Instead, the mass already present in the disk is simply redistributed azimuthally, to give expressions for the surface density and potential which are doubly periodic as a function of azimuth angle. The resulting total surface density can be adjusted to have approximately the elongation of the bar isophotes, but this procedure is rather ad hoc by comparison to the triaxial fitting method.

Surprisingly, the results of the modeling are quite emphatic in favoring the oval distortion as the more correct representation of the bar potential in NGC 3359. The essential distinction between the two types of potential is that the triaxial bar is a massive body of limited extent, while the oval distortion is a perturbation in the disk surface density, whose amplitude only decays gradually with radius. Because of this property of the triaxial bar, test masses in the outer regions of a galaxy see it is as a point mass, to lowest approximation. At the same time, a bar of this kind tends to dominate the disk mass in the zone which it occupies, and, because of its fairly extreme elongation, is very disruptive of the gas flow in this zone. Therefore, the kinematical perturbations are much stronger in the inner part of the galaxy than in the outer regions where the spiral arms are found. A thorough exploration of the allowed parameters of the triaxial bar model in NGC 3359

has demonstrated conclusively that no combination of bar mass and pattern speed is capable of producing a velocity response throughout the gas which is similar to that observed.

The oval distortion models, on the other hand, work very well in matching the locations and the amplitudes of the velocity distortions in the galaxy. The inescapable conclusion is that the spiral structure of NGC 3359 is not driven directly by the observed, optical bar, as defined by a straightforward interpretation of the near infrared surface photometry. It remains very possible that the bar is instrumental in exciting the departures from axisymmetry in the galaxy, but some complementary process is needed to prevent the nonaxisymmetric force from falling off too rapidly in the outer portions of the gas disk. The models presented here accomplish this by the use of oval distortions of the disk potential, which extend beyond the radius of the optically defined bar. In a real galaxy, the force of the self-gravity of the gas may also be of importance. Unfortunately, one cannot include this force, and obtain a numerically stable response in the gas, unless the applied potential is truly an equilibrium figure, to a high degree of accuracy. Little is known about the variety of equilibrium configurations available to a rotating disk-bar system under typical galaxian conditions; the establishment of a diversity of such configurations is an important topic for future investigation. The major dynamical finding to arise from the present study is that the common assumption, that the visible bar directly contributes the nonaxisymmetric force in barred spiral galaxies, is not true in the case of NGC 3359.

APPENDIX A
EVALUATION OF THE INTEGRAL $I_3(b,c;a)$

In connection with the problem of finding the forces due to a triaxial homoeoid, it is desired to evaluate

$$I_3(b,c;a) = \int_{\kappa}^{\infty}(b^2+s)^{-3/2}(c^2+s)^{-3/2}(a^2+s)^{-1/2}ds, \tag{A-1}$$

where $a > b > c$, and with a,b,c, and κ all being positive, real numbers. We have evaluated several integrals similar to this in Chapter III, using the results of Section 3.13 of Gradshteyn and Ryzhik (1980, hereafter GR). However, the formula given in that Section by GR for $I_3(b,c;a)$ contains a typographical error. Therefore, it is necessary to find an alternate way of writing equation (A-1) which will allow us to evaluate $I_3(b,c;a)$.

Our approach is to re-express the integral by means of a pair of substitutions. First, let us define variables t and p by

$$t^2 = (a^2-c^2)(a^2+s)^{-1}, \tag{A-2}$$

$$p^2 = (a^2-b^2)(a^2-c^2)^{-1}. \tag{A-3}$$

Since $a > b > c$, and since the integration extends only over positive values of s, the variable of integration, both t and p are real numbers of absolute value less than one. We observe that

$$\frac{ds}{dt} = -2t^{-3}(a^2-c^2), \tag{A-4}$$

and denote the value of t when s = κ by T. Then equation (A-1) can be rewritten as

$$I_3(b,c;a) = 2(a^2-c^2)^{-5/2} \int_0^T (1-p^2t^2)^{-3/2} \times$$
$$\times (1-t^2)^{-3/2} t^4 \, dt. \qquad (A-5)$$

Because of the restrictions on the magnitudes of p and t, we can change variables again, defining z by t = sin z, α by T = sin α, and Δ by

$$\Delta = (1-p^2t^2)^{1/2}. \qquad (A-6)$$

Since (dt/dz) = cos z, the integral is now

$$I_3(b,c;a) = 2(a^2-c^2)^{-5/2} \int_0^\alpha (\sin^4 z)(\Delta^3 \cos^2 z)^{-1} dz. \qquad (A-7)$$

Equation (A-7) is in a form which is tabulated in Section 2.58 of GR. By their formula 2.581.2, we write

$$I_3(b,c;a) = 2(a^2-c^2)^{-5/2} [(\sin^3 z)(p^2 \Delta \cos^3 z)^{-1} -$$
$$- 3p^{-2} \int (\sin^2 z)(\Delta \cos^2 z)^{-1} dz -$$
$$- 3p^{-2} \int (\sin^4 z)(\Delta \cos^4 z) dz]_0^\alpha , \qquad (A-8)$$

or

$$I_3(b,c;a) = 2(a^2-c^2)^{-5/2} [p^{-2} \Delta^{-1} \tan^3 \alpha -$$
$$- 3p^{-2}(G_1+G_2)], \qquad (A-9)$$

where G_1 and G_2 are the two integrals in (A-8). By repeated substitution of the identity $\sin^2 z = (1-\cos^2 z)$ in the numerators,

it is easily shown that

$$G_2 = \int_0^{\alpha}(\Delta\cos^4 z)^{-1}dz -$$
$$- \int_0^{\alpha}(\Delta\cos^2 z)^{-1}dz - G_1$$
$$= G_3 - G_4 - G_1.$$

(A-10)

Integrals G_3 and G_4 are listed in GR as formulae 2.584.87 and 2.584.72, respectively:

$$G_3 = \frac{1}{3}(1-p^2)^{-1}[\Delta_\alpha \tan^3\alpha -$$
$$- (1-p^2)^{-1}(5p^2-3)\Delta_\alpha \tan\alpha - (3p^2-2)F(\alpha,p) +$$
$$+ 2(1-p^2)^{-1}(2p^2-1)E(\alpha,p)],$$

(A-11)

$$G_4 = F(\alpha,p) - (1-p^2)^{-1}E(\alpha,p) + (1-p^2)^{-1}\Delta_\alpha \tan\alpha.$$

(A-12)

In writing (A-11) and (A-12), we have used the elliptic integrals of the first and second kinds, $F(\alpha,p)$ and $E(\alpha,p)$, respectively. We have also denoted the value of Δ at $z=\alpha$ by Δ_α, and, finally, we have made use of the identities $F(0,p) = E(0,p) = \tan 0 = 0$.

When equations (A-11) and (A-12) are once again expressed in terms of a,b,c, and κ, the result is

$$G_3 - G_4 = \frac{1}{3}(b^2-c^2)^{-2}(a^2-c^2)[(2a^2-b^2-c)E(\alpha,p) -$$
$$- (b^2-c^2)F(\alpha,p)] + \frac{1}{3}\Delta_\alpha \tan\alpha (a^2-c^2)(b^2-c^2)^{-2} \times$$
$$\times (c^2+\kappa)^{-1}[a^2b^2 + b^2c^2 - 3a^2c^2 + c^4 +$$
$$+ \kappa(2b^2-2a^2)].$$

(A-13)

Upon combining equations (A-9), (A-10), and (A-13), the desired expression for I_3 (b,c;a) is obtained. After algebraic simplification, it may be written as

$$I_3 (b,c;a) = 2(a^2-b^2)^{-1}(b^2-c^2)^{-2}(a^2-c^2)^{-1/2} \times$$
$$\times [(b^2-c^2)F(\alpha,p) - (2a^2-b^2-c^2)E(\alpha,p)] +$$
$$+ 2(b^2-c^2)^{-2}(a^2+\kappa)^{-1/2}(b^2+\kappa)^{-1/2} \times$$
$$\times (c^2+\kappa)^{-1/2}(b^2+2\kappa+c^2). \qquad \text{(A-14)}$$

APPENDIX B
STABILITY OF THE BEAM SCHEME

In this appendix, we offer a rudimentary analysis of the stability of the beam scheme, in the one-dimensional case. This discussion is not meant to be a rigorous stability analysis, but rather to aid the reader in understanding the fundamental cause of the observed stability of this code in the supersonic velocity regime.

Our analysis begins with equation (5-11), the equation of continuity in the beam scheme. We will consider the behavior as the finite-difference equation approaches the continuous limit, that is, as the cell size L and the time step Δt go to zero. Writing equation (5-11) with the right-hand side expressed in terms of the momenta p, we have

$$\frac{\Delta \rho_j}{\Delta t} = \frac{1}{L} (p_{j-1} - p_j). \qquad \text{(B-1)}$$

The reader will recall that this equation was derived under the assumption that matter is flowing from cell (j-1) to cell j; the asymmetry of equation (B-1) reflects the upwind differencing which effectively is performed in the beam scheme.

Now suppose that we take the continuous limit of equation (B-1), as mentioned above, and expand p_{j-1} in a Taylor series about cell j. Dropping high-order terms, we obtain

$$\frac{\partial}{\partial t}(\rho) = D(\rho u),$$

(B-2)

where D is the spatial differential operator,

$$D = \frac{L}{2}\frac{\partial^2}{\partial x^2} - \frac{\partial}{\partial x} \quad .$$

(B-3)

Similarly, the equations of motion and of energy reduce, in this limit, to

$$\frac{\partial}{\partial t}(\rho u) = D(\rho u^2 + P),$$

(B-4)

$$\frac{\partial}{\partial t}(E) = D[u(E+P)],$$

(B-5)

where P and E are the pressure and the kinetic energy density, respectively. In writing these simple forms, we have assumed that no external forces are acting, and have ignored dissipative terms in the energy equation. Note the similarities of equations (B-2), (B-4), and (B-5) to equations (1-1) through (1-3).

To illustrate the stability behavior of the beam scheme, let us consider a perturbation of a system which is initially "flat," with $\rho=\rho_0$, $u=u_0$, and $P=P_0$ everywhere. This is the degenerate solution to the system of equations (B-2) and (B-4). Let the perturbation be $\rho=\rho_0+\rho_1$, $u=u_0+u_1$, $P=P_0+c^2\rho$, with c constant. Then, dropping

terms of second order in the perturbation, the equation of continuity becomes

$$\frac{\partial}{\partial t}\ (\rho_1) = u_0 D\rho_1 + \rho_0 Du_1 \ ,$$

(B-6)

while the equation of motion is

$$u_0 \frac{\partial}{\partial t}\ (\rho_1) + \rho_0 \frac{\partial}{\partial t}\ (u_1) = 2\rho_0 u_0 Du_1 + (u_0^2 + c^2) D\rho_1 \ .$$

(B-7)

These can be combined to give

$$\rho_0 \frac{\partial}{\partial t}\ (u_1) - u_0 \frac{\partial}{\partial t}\ (\rho_1) = (c^2 - u_0^2)\ D\rho_1 .$$

(B-8)

Now, if we operate on equation (B-6) with $\frac{\partial}{\partial t}$, and on (B-8) with D, and combine the results again, we obtain the equation

$$\frac{\partial^2}{\partial t^2}(\rho_1) - 2u_0 D \frac{\partial}{\partial t}(\rho_1) + (u_0^2 - c^2) D^2\rho_1 = 0.$$

(B-9)

Suppose $\rho_1 = \rho_1' \exp(ikx+\omega t)$, where k is real. We evaluate $D\rho_1$ and $D^2\rho_1$, from equation (B-3). In the latter, we drop terms of order L^2, as before. Then equation (B-9) may be written as

$$[\omega^2+(u_0Lk^2+2iu_0k)\omega+k^2(u_0^2-c^2)(iLk-1)]\rho_1 = 0 .$$

(B-10)

To solve this for ω, introduce

$$y = \omega+iu_0k .$$

(B-11)

In terms of this variable, equation (B-10) becomes

$$y^2+u_0k^2Ly+k^2c^2(1-iLk) = 0.$$

(B-12)

We can then find y by means of the quadratic formula. If we subsequently drop terms in L^2, we find that

$$y = -\frac{1}{2} u_0Lk^2 \pm kc(\frac{1}{2} Lk+i),$$

(B-13)

and so the two solutions for ω are

$$\omega = -\frac{1}{2} Lk^2(c+u_0) -ik(c+u_0),$$

(B-14a)

and

$$\omega = \frac{1}{2} Lk^2(c-u_0) +ik(c-u_0) .$$

(B-14b)

The first of these is a stable solution, but the second need not be. The condition for stability is

$$u_0 > c, \text{ stable.}$$

(B-15)

An analysis which is very much in this spirit can easily be performed for the individual beams, starting from the relations listed in Table 5-1. We will summarize the results of such an investigation very briefly. For the central beam, whose associated mass density is a, the differential equations approximated by the beam scheme are

$$\frac{\partial}{\partial t}(a) = D(au),$$

(B-16a)

$$\frac{\partial}{\partial t}(au) = D(au^2) .$$

(B-16b)

Note that this pair of equations does not contain the pressure; in the beam scheme, the pressure is carried solely by the side beams. The solution of equations (B-16) turns out to be unconditionally stable for small perturbations, of the kind we have considered in this appendix.

Performing the same kind of analysis on the two side beams considered jointly, we find the relevant differential equations to be

$$\frac{\partial}{\partial t}(b) = D(bu),$$

(B-17a)

$$\frac{\partial}{\partial t}(bu) = D(bu^2+3b\sigma^2),$$

(B-17b)

where b is the mass density associated with each side beam. These equations are of exactly the same form as equations (B-2) and (B-4),

with $(3\sigma^2)$ playing the role of c^2 in the earlier analysis. We remind the reader that, in setting up the beam scheme, we found that the velocity offset Δu of the side beams should obey $(\Delta u)^2 = 3\sigma^2$, for our choice of beam weights. We see, then, that the velocity offset Δu can be interpreted as a sound speed. The stability criterion for the side beams follows directly from equation (B-15):

$u_0 > \Delta u$, stable. (B-18)

REFERENCES

Abramowitz, M., and Stegun, I.A., eds. 1965, Handbook of Mathematical Functions (New York: Dover).

Allen, C. W. 1973, Astrophysical Quantities (Third edition, London: Athlone Press).

Allen, R. J., and Goss, W. M. 1979, Astr. Ap. Suppl., 36, 135.

Baars, J. W. M., Genzel, R., Pauliny-Toth, I. I. K., and Witzel, A. 1977, Astr. Ap., 61, 99.

Baldwin, J. E., Field, C., Warner, P. J., and Wright, M. C. H. 1971, M. N. R. A. S., 154, 445.

Ball, R., and Hunter, J. H., Jr. 1984, Ap. J., submitted.

Berman, R. H., Pollard, D. J., and Hockney, R. W. 1979, Astr. Ap., 78, 133.

Bertola, F., and Galletta, G. 1979, Astr. Ap., 77, 363.

Blackman, C. P. 1983, M. N. R. A. S., 202, 379.

Bosma, A. 1978, Ph. D. thesis, University of Groningen.

Bosma, A. 1981a, A. J., 86, 1791.

Bosma, A. 1981b, A. J., 86, 1825.

Bosma, A. 1983, In IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), p. 11.

Bosma, A., Ekers, R. D., and Lequeux, J. 1977. Astr. Ap., 57, 97.

Bracewell, R. 1965. The Fourier Transform and Its Applications (New York: McGraw-Hill).

Brinks, E., and Shane, W. W. 1984, Astr. Ap. Suppl., 55, 179.

Casertano, S. 1983, M. N. R. A. S., 203, 735.

Cheriguene, M. F. 1975, In La Dynamique des Galaxies Spirales, ed. L. Weliachew (Paris: Editions due Centre National de la Recherche), p. 439.

Christiansen, W. N., and Hindman, J. V. 1952, Austral. J. Sci. Res., A5, 437.

Clark, B. G. 1980, Astr. Ap., 89, 377.

Clark, B. G. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, ed. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 10.

Contopoulos, G. 1983, Astr. Ap., 117, 89.

Cornwell, T. J. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, eds. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 9.

D'Addario, L. R. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, eds. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Stronomy Observatory), Chapter 7.

de Vaucouleurs, G. 1979, Ap. J., 227, 729.

de Vaucouleurs, G., de Vaucouleurs, A., and Corwin, H. G., Jr. 1976, Second Reference Catalogue of Bright Galaxies (Austin: University of Texas Press).

Dixon, R. S., and Sonneborn, G. 1980, A Master List of Nonstellar Optical Astronomical Objects (Columbus: Ohio State University Press).

Ekers, R. D. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, ed. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 12.

Ekers, R. D., and van Gorkom, J. H. 1984, In Indirect Imaging, ed. J. A. Roberts (Cambridge: Cambridge University Press), p. 21.

Elmegreen, B. G., and Elmegreen, D. M. 1978, Ap. J., 220, 1051.

Elmegreen, B. G., and Lada, C. J. 1977, Ap. J., 214, 725.

Elmegreen, D. M. 1981, Ap. J. Suppl., 47, 229.

Emerson, D. T. 1976, M. N. R. A. S., 176, 321.

Ewen, H. I., and Purcell, E. M. 1951, Nature, 168, 356.

Fisher, J. R., and Tully, R. B. 1981, Ap. J. Suppl., 47, 139.

Fomalont, E. B., and Wright, M. C. H. 1974, In Galactic and Extra-Galactic Radio Astronomy, ed. G. L. Verschuur and K. I. Kellerman (New York: Springer-Verlag), p. 256.

Freeman, K. C. 1975, In Stars and Stellar Systems, Vol. 9, Galaxies and the Universe, ed. A. Sandage, M. Sandage, and J. Kristian (Chicago: University of Chicago Press), p. 409.

Gallouet, L., Heidmann, N., and Dampierre, F. 1973, Astr. Ap. Suppl., 12, 89.

Gerola, H., and Seiden, P. E. 1978, Ap. J., 223, 129.

Gerola, H., and Seiden, P. E. 1979, In Photometry, Kinematics, and Dynamics of Galaxies, ed. D. S. Evans (Austin: University of Texas), p. 475.

Gottesman, S. T. 1982, A. J., 87, 751.

Gottesman, S. T., Ball, R., Hunter, J. H., Jr., and Huntley, J. M. 1984, Ap. J., to appear.

Gradshteyn, I. S., and Ryzhik, I. M. 1980, Table of Integrals, Series, and Products (New York: Academic Press).

Haud, U., and Einasto, J. 1983, In IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel).

Hjellming, R. M. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, ed. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 4.

Hjellming, R. M., and Basart, J. P. 1982, In An Introduction to the NRAO Very Large Array, April 1982 Edition, ed. R. M. Hjellming (Socorro: National Radio Astronomy Observatory), Chapter 2.

Hodge, P. W. 1969, Ap. J. Suppl., 18, 73.

Hogbom, J. A. 1974, Astr. Ap. Suppl., 15, 417.

Hohl, F. 1971, Ap. J., 168, 343.

Hohl, F. 1972, J. Comp. Phys., 9, 10.

Hohl, F. 1976, A. J., 81, 30.

Hohl, F. 1978, A. J., 83, 768.

Hunter, J. H., Jr., Ball, R., Gottesman, S. T., and Huntley, J. M. 1984, in preparation.

Huntley, J. M. 1977, Ph. D. thesis, University of Virginia.

Huntley, J. M. 1980, Ap. J., 238, 524.

Hunter, J. H., Jr., Ball, R., and Gottesman, S. T. 1984, M. N. R. A. S., 208, 1.

Huntley, J. M., Sanders, R. H., and Roberts, W. W., Jr. 1978, Ap. J., 221, 521.

Jackson, P. D., and Kellman, S. A. 1974, Ap. J., 190, 53.

Jensen, E. B., Talbot, R. J., Jr., and Dufour, R. J. 1981, Ap. J., 243, 716.

Kalnajs, A. J. 1965, Ph. D. thesis, Harvard University.

Kalnajs, A. J. 1972, Ap. J., 175, 63.

Kerr, F. J., and Westerhout, G. 1965, In Stars and Stellar Systems, Vol. 5, Galactic Structure, ed. A. Blaauw and M. Schmidt (Chicago: University of Chicago Press), p. 167.

Kraus, J. D. 1966, Radio Astronomy (New York: McGraw-Hill).

Lada, C. J., Blitz, L., and Elmegreen, B. G. 1978, In Protostars and Planets, ed. T. Gehrels (Tucson: University of Arizona Press), p. 341.

Lau, Y. Y., and Bertin, G. 1978, Ap. J., 226, 508.

Lin, C. C., and Lau, Y. Y. 1979, Stud. Appl. Math., 60, 97.

Lin, C. C., and Shu, F. H. 1964, Ap. J., 140, 646.

Lin, C. C., and Shu, F. H. 1966, Proc. Nat. Acad. Sci., 55, 229.

Lin, C. C., and Shu, F. H. 1967, In IAU Symposium 31, Radio Astronomy and the Galactic System, ed. H. van Woerden (London: Academic Press), p. 313.

Lin, C. C., Yuan, C., and Shu, F. H. 1969, Ap. J., 155, 721.

Lindblad, B. 1956, Stockholm Obs. Ann., 19, Number 2.

Lindblad, B. 1963, Stockholm Obs. Ann., 22, Number 5.

Lynden-Bell, D., and Kalnajs, A. J. 1972, M. N. R. A. S., 157, 1.

MacMillan, W. D. 1958, Theoretical Mechanics: The Theory of the Potential (New York: Dover).

Mark, J. W.-K. 1974, Ap. J., 193, 539.

Mark, J. W.-K. 1976, Ap. J., 205, 363.

Matsuda, T., and Isaka, H. 1980, Prog. Theor. Phys., 64, 1265.

McKee, C. F., and Ostriker, J. P. 1977, Ap. J., 218, 148.

Mestel, L. 1963, M. N. R. A. S., 126, 553.

Mihalas, D., and Binney, J. 1981, Galactic Astronomy (Second edition, San Fracisco: Freeman).

Miller, R. H. 1971, Ap. Space Sci., 14, 73.

Miller, R. H. 1976, Ap. J., 207, 403.

Miller, R. H. 1978, Ap. J., 223, 122.

Miller, R. H., and Smith, B. F. 1979, Ap. J., 227, 407.

Morris, M., and Rickard, L. J. 1982, Ann. Rev. Astr. Ap., 20, 517.

Mueller, M. W., and Arnett, W. D. 1976, Ap. J., 210, 670.

Muller, C. A., and Oort, J. H. 1951, Nature, 168, 356.

Okamura, S. 1973, Pub. Astr. Soc. Japan, 30, 91.

Ostriker, J. P., and Peebles, P. J. E. 1973, Ap. J., 186, 467.

Roache, P. J. 1976, Computational Fluid Dynamics (Revised edition, Albuquerque: Hermosa).

Roberts, M. S. 1968, A. J., 73, 945.

Roberts, M. S. 1975, In Stars and Stellar Systems, Vol. 9, Galaxies and the Universe, ed. A. Sandage, M. Sandage, and J. Kristian (Chiago: University of Chicago Press), p. 309.

Roberts, W. W., Jr., Huntley, J. M., and van Albada, G. D. 1979, Ap. J., 233, 67.

Roberts, W. W., Jr., Roberts, M. S., and Shu, F. H. 1975, Ap. J., 196, 381.

Rogstad, D. H., and Shostak, G. S. 1971, Astr. Ap., 13, 99.

Rogstad, D. H., Wright, M. C. H., and Lockhart, I. A. 1976, Ap. J., 204, 703.

Rots, A. H. 1980, Astr. Ap. Suppl., 41, 189.

Rots, A. H. 1983, A Short Guide for VLA Spectral Line Observers (Edition 5.2, Socorro: National Radio Astronomy Observatory).

Rots, A. H., and Shane, W. W. 1975, Astr. Ap., 45, 25.

Sancisi, R. 1981, In The Structure and Evolution of Normal Galaxies, eds. S. M. Fall and D. Lynden-Bell (Cambridge: Cambridge University Press), p. 149.

Sancisi, R., and Allen, R. J. 1979, Astr. Ap., 74, 73.

Sancisi, R., Allen, R. J., and Sullivan, W. T., III. 1979, Astr. Ap., 78, 217.

Sandage, A. 1961, The Hubble Atlas of Galaxies (Washington: Carnegie Institution of Washington).

Sandage, A., and Tammann, G. A. 1981, A Revised Shapley-Ames Catalog of Bright Galaxies (Washington: Carnegie Institution of Washington).

Sanders, R. H. 1977, Ap. J., 217, 916.

Sanders, R. H., and Huntley, J. M. 1976, Ap. J., 209, 53.

Sanders, R. H., and Prendergast, K. H. 1974, Ap. J., 188, 489.

Sanders, R. H., and Tubbs, A. D. 1980, Ap. J., 235, 803.

Schempp, W. V. 1982, Ap. J., 258, 96.

Schulman, L. S., and Seiden, P. E. 1978, J. Stat. Phys., 19, 293.

Schulman, L. S., and Seiden, P. E. 1982, J. Stat. Phys., 27, 83.

Seiden, P. E. 1983, Ap. J., 266, 555.

Seiden, P. E., and Gerola, H. 1979, Ap. J., 233, 56.

Seiden, P. E., Schulman, L. S., and Elmegreen, B. G. 1983, Bull. AAS, 15, 944.

Seiden, P. E., Schulman, L. S., and Gerola, H. 1979, Ap. J., 232, 702.

Sellwood, J. A. 1980, Astr. Ap., 89, 296.

Shu, F. H. 1970, Ap. J., 160, 99.

Siefert, P. T., Gottesman, S. T., and Wright, M. C. H. 1975, In La Dynamique des Galaxies Spirales, ed. L. Weliachew (Paris: Centre National de la Recherche Scientifique), p. 425.

Sorensen, S.-A., Matsuda, T., and Fujimoto, M. 1976, Ap. Space Sci., 43, 491.

Spitzer, L., Jr. 1978, Physical Processes in the Interstellar Medium (New York: Wiley).

Sramek, R. A. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, ed. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 2.

Stark, A. A. 1977, Ap. J., 213, 368.

Strom, K. M., and Strom, S. E. 1978a, A. J., 83, 73.

Strom, S. E., and Strom, K. M. 1978b, A. J., 83, 732.

Strom, K. M., and Strom, S. E. 1978c, A. J., 83, 1293.

Thompson, A. R., Clark, B. G., Wade, C. M., and Napier, P. J. 1980, Ap. J. Suppl., 44, 151.

Thompson, A. R., and D'Addario, L. R., eds. 1982, Synthesis Mapping, Proceedings of the NRAO-VLA Workshop (Green Bank: National Radio Astronomy Observatory).

Toomre, A. 1963, Ap. J., 138, 385.

Toomre, A. 1964, Ap. J., 139, 1217.

Toomre, A. 1969, Ap. J., 158, 899.

Toomre, A. 1977, Ann. Rev. Astr. Ap., 15, 437.

Unwin, S. C. 1980a, M. N. R. A. S., 190, 551.

Unwin, S. C. 1980b, M. N. R. A. S., 192, 243.

van Albada, G. D., and Roberts, W. W., Jr. 1981, Ap. J., 246, 740.

van de Hulst, H. C. 1945, Neder, Tijd. Natuurkunde, 11, 210.

van den Bergh, S. 1960, Ap. J., 131, 215.

van den Bergh, S. 1978, In IAU Symposium 77, Structure and Properties of Nearby Galaxies, ed. M. Berkhuijsen and R. Wielebinski (Dordrecht: Reidel), p. 247.

van den Bergh, S. 1981, In The Structure and Evolution of Normal Galaxies, ed. S. M. Fall and D. Lynden-Bell (Cambridge: Cambridge University Press), p. 201.

van der Hulst, J. M., Ondrechen, M. P., van Gorkom, J. H., and Hummel, E. 1983, In IAU Symposium 100, Internal Kinematics and Dynamics of Galaxies, ed. E. Athanassoula (Dordrecht: Reidel), p. 233.

van der Kruit, P. C. 1981, Astr. Ap., 99, 298.

van der Kruit, P. C., and Searle, L. 1981a, Astr. Ap., 95, 105.

van der Kruit, P. C., and Searle, L. 1981b, Astr. Ap., 95, 116.

van der Kruit, P. C., and Searle, L. 1982a, Astr. Ap., 110, 61.

van der Kruit, P. C., and Searle, L. 1982b, Astr. Ap., 110, 79.

van Gorkom, J. H. 1982, In Synthesis Mapping, Proceedings of the NRAO-VLA Workshop, ed. A. R. Thompson and L. R. D'Addario (Green Bank: National Radio Astronomy Observatory), Chapter 15.

Warner, P. J., Wright, M. C. H., and Baldwin, J. E. 1973, M. N. R. A. S., 163, 163.

Weliachew, L., and Gottesman, S. T. 1973, Astr. Ap., 24, 59.

Weliachew, L., Sancisi, R., and Guelin, M. 1978, Astr. Ap., 65, 37.

Welter, G. L. 1982, Astr. Ap., 105, 237.

Winter, A. J. B. 1975, M. N. R. A. S., 172, 1.

BIOGRAPHICAL SKETCH

John Roger Ball was born July 6, 1956, to Lee and Kate Ball, in Pikeville, Kentucky. He grew up, for the most part, in Durant, Oklahoma, with his parents and his sister, Mary Elizabeth, to whom this volume is dedicated. He graduated from Durant High School in 1973, and, the following fall, entered the University of the South at Sewanee, Tennessee. He earned a B.A. in physics from Sewanee in May, 1977, and began graduate work in astronomy at the University of Florida in September, 1977. He was married to Sandy Lemlich on March 27, 1983. He once had some hobbies, but can no longer remember anything about them. The Ph.D. degree is expected in August, 1984.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

James H. Hunter, Jr., Chairman
Professor of Astronomy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Stephen T. Gottesman
Professor of Astronomy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Haywood C. Smith, Jr.
Associate Professor of Astronomy

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gary G. Ihas
Associate Professor of Physics

This dissertation was submitted to the Graduate Faculty of the Department of Astronomy in the College of Liberal Arts and Sciences and to the Graduate Council, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1984

Dean for Graduate Studies and Research

UNIVERSITY OF FLORIDA
3 1262 08553 2132