Near Infrared Surface Photometry of Spiral Galaxies

James Malcolm Stewart

Doctor of Philosophy
University of Edinburgh
1984

This thesis has been composed by me and consists entirely of my own work, except where specifically indicated in the text.

February 1984

TABLE OF CONTENTS

ABSTRACT 1

Chapter 1 INTRODUCTION

1.1 Importance of Infrared Observations of Galaxies 2
1.2 Observing Molecular Gas Distributions in Galaxies 3
1.3 Infrared Surface Photometry Techniques 6

Chapter 2 MECHANISMS FOR NEAR INFRARED LUMINOSITY IN GALAXIES

2.1 Introduction 9
2.2 Main Sequence Stars 11
2.3 Giant Stars 14
2.4 Supergiants 19
2.5 Carbon Stars 21
2.6 Emission from Dust and Gas 24
2.7 Observed Infrared Emission from Galaxies 27

Chapter 3 THE INSTRUMENT

3.1 Advantages of Simultaneous Channels 33
3.2 Optical Design 33
3.2.1 Basic Concept 36
3.2.2 Spherical Aberration 39
3.2.3 Chromatic Aberration 47
3.2.4 Beamsplitter and Filters 48
3.2.5 Mark II Optical Design 50
3.3 Electronic Design 50
3.3.1 Preamplifier Circuit 54
3.3.2 Gain and Frequency Response 56
3.3.3 Noise Performance 61
3.4 Mechanical Construction and Alignment 66
3.5 Laboratory Tests 66
3.5.1 Optical Tests 68
3.5.2 Electrical Tests 73

3.6 Performance on the Telescope 79
3.6.1 Beam Profiles 79
3.6.2 Sensitivity 81
3.7 Data Processing System 82
3.7.1 Description 82
3.7.2 Stability 83
3.7.3 Computer Configuration 86

Chapter 4 THE OBSERVATIONS

4.1 General Procedure 88
4.2 Chopping and Beamswitching 88
4.2.1 Observations with Beamswitching 90
4.2.2 Avoidance of Beamswitching 92
4.3 Effect of Chopper Risetime 93
4.4 Standards 93
4.4.1 Filter Systems 98
4.4.2 Use of Galactic Nuclei as Local Standards 101
4.4.3 Mark I Cryostat 104
4.4.4 Mark II Cryostat 106
4.5 Reduction of Beamswitched Data 106
4.5.1 Zero Points 107
4.5.2 Colour and Brightness Errors 109
4.6 Reduction of Rastered Data 111
4.6.1 Sky Background Level 111
4.6.2 Combination of Data from the Two Beams 115
4.6.3 Estimate of Errors 117
4.7 Features of the Data 119
4.7.1 Airmass Corrections 119
4.7.2 Self Consistency of Data 121
4.7.3 Pointing Errors 122
4.8 Technical Conclusions 124

CHAPTER 5 THE DATA

5.1 The Galaxies Observed 126

5.2 NGC 7331 128
5.2.1 The Data 128
5.2.2 Colour Variations 135
5.2.3 Comparison with Optical Data 143
5.3 NGC 4565 148
5.3.1 The Data 148
5.3.2 Colour Variations 152
5.3.3 Comparison with Optical Data 154
5.4 NGC 5907 157
5.4.1 The Data 157
5.4.2 Colour Variations 157
5.4.3 Brightness Distribution 160
5.5 NGC 7814 161
5.6 NGC 4216 163
5.7 Summary of Data 166

Chapter 6 INTERSTELLAR DUST

6.1 Origin of the Reddening Curve 167
6.2 Extinction Geometry Effects 172
6.2.1 Background 172
6.2.2 Homogeneous Luminosity and Dust Distribution 174
6.2.3 Thin Sheet of Extinction Embedded in Luminosity 178
6.2.4 Effect of Clumping 181
6.3 Observations of Extinction in the Milky Way 186
6.4 Previous Determinations of Total Extinction in the Milky Way 190
6.5 Interstellar Extinction Law in Galaxies 192
6.6 Effect of Scattering in the Infrared 195
6.7 Determination of Colour Excess Ratio 198
6.7.1 NGC 7331 199
6.7.2 NGC 4565 202
6.7.3 NGC 5907 205
6.8 Determination of Ratio of Total to Selective Extinction 205

6.9 Implications for the Extinction Curve in Galaxies 211

Chapter 7 DISTRIBUTION OF INTERSTELLAR MATERIAL

7.1 Dust Distribution in NGC 7331 216
7.2 Optical Depth of Dust in NGC 7331 220
7.3 Dust and Gas Densities in NGC 7331 223
7.4 The Interstellar Medium in NGC 7331 and the Milky Way 227
7.5 Atomic or Molecular Gas 229
7.6 Other Galaxies 234
7.7 Molecular Gas Distribution in Galaxies 235
7.8 A Model for the NGC 7331 Radial H_2 Distribution 238
7.9 Molecular Rings in Galaxies 244
7.10 Conclusions 248

Acknowledgements 250

Bibliography 251

Appendix A TABULATED DATA 260

Appendix B PUBLISHED PAPER inside back cover

ABSTRACT

Very little surface photometry of normal galaxies has yet been performed in the near infrared. The likely sources of infrared emission in galaxies are discussed paying particular regard to mechanisms which could cause variations in the observed colours. Late type giants and, in the particular galaxies selected, extinction from dust are shown to be the major contributors to the appearance of the galaxies in the near infrared.

The problems of measuring accurate surface colours in the presence of large brightness gradients are discussed. It is shown that simultaneous measurements at two wavelengths are necessary, requiring the construction of a special infrared photometer with two detectors, a beamsplitter and a common focal plane aperture. The design criteria of the system are discussed and the performance of the system in the laboratory and on the telescope is analysed.

Techniques for mapping galaxies are discussed and it is shown that beamswitching may be replaced by measurements on the sky. Good instrument stability and performance of the chopping secondary mirror have been shown to be necessary.

Five normal spiral galaxies were observed and the colour variations analysed. The effect on the colours of the distribution of dust and stars are discussed. The near infrared extinction properties of dust in these galaxies have been shown to be similar to those of dust in the Milky Way.

The distribution of reddening is related to the distribution of interstellar gas, predominantly molecular, in NGC 7331. A central deficiency of molecular gas in NGC 7331 has been deduced, analagous to the "molecular ring" in the Milky Way. At least 2 percent of the mass in the inner 9 kpc is in the interstellar medium (ISM). The advantage is demonstrated of higher spatial resolution than is currently available in CO data. The surface density of the ISM is shown to be relatively constant between 3 and 9 kpc, similar to the distribution in NGC 2841.

1. Introduction

1.1 Importance of Infrared Observations of Galaxies

Galaxies have been studied most extensively at optical wavelengths, not because these wavelengths most effectively portray the structure of the galaxy, but because of the efficiency of optical techniques in collecting information. Since galaxies were first observed in the infrared (Johnson 1966a) it has been known that much of the mass of galaxies is in the form of late type giants which radiate predominantly in the near infrared. Optical data do not represent this component accurately but are sensitive to young objects e.g. O and B stars. These trace the spiral structure but not the underlying component representative of the dominant mass.

Elliptical galaxies are relatively homogeneous and so optical and infrared observations see the same types of objects at different points in the galaxy. This is not true of spiral galaxies which have distinct and varied components e.g. disks, bulges, spiral arms and lenses. Spiral galaxies may therefore present a different appearance in the infrared representing more closely the distribution of the dominant mass component.

Chapter 2 discusses the different types of object which may contribute to the brightness and colours within galaxies, so that the observed colour variations may be understood. More attention is given to the study of colour variations within the galaxies than to the brightness variations for which observations to larger radii are required to decompose the brightness distributions into their constituent components.

The observations of 5 galaxies are presented in Chapter 5. These galaxies were selected to be normal undisturbed spiral galaxies of various Hubble types. Extinction from dust was found to be the major source of colour variations and this allows the properties of dust in other galaxies to be compared with the Milky Way. The two colour (J-K and H-K) observations of NGC 4565, NGC 5907 and NGC 7331 allow reddening vectors and ratios of total to selective extinction to be derived (Chapter 6). Single colour (J-K) observations of NGC 4216 and NGC 7814 are also presented and are interpreted in terms of their dust distribution.

The use of infrared wavelengths offers an important advantage in studying extragalactic dust. Because extinction at K is about one tenth that at V (Johnson 1963), greater column densities may be observed before the colours become saturated. The extinction evident in the optical appearance of spiral galaxies indicates that the optical colours are very likely to be saturated. The properties and distribution of dust in galaxies are important because dust plays an important role in star formation. Dust shields shields clouds of gas from optical and ultraviolet radiation. It also gains energy from the gas by collisions and radiates this energy at long infrared wavelengths at which the dust is optically thin. The dust therefore cools the gas as the cloud contracts aiding in star formation.

Stars embedded within clouds can change the composition and size of grains by evaporating volatile mantles. The extinction properties depend on the size, shape and composition of the grains and so different extinction characteristics in other galaxies may indicate different star formation processes. Both the distribution of dust and its extinction properties may therefore provide information on the process of star formation within galaxies.

1.2 Observing Molecular Gas Distributions in Galaxies

The existence of material in the interstellar medium has been known since Trumpler (1930) showed that the light from globular clusters suffered extinction from dust. The major component of the interstellar medium (ISM) was believed to be neutral hydrogen, HI, which could be detected and mapped by the 21 cm hyperfine transition. The HI distribution in the Milky Way was originally mapped using a 7.5m dish (see Kerr, 1968) producing a resolution of about 2 degrees. Much greater resolution is required to map the HI distribution in other galaxies and this is achieved by interferometric techniques (also used to map sources within the Milky Way). For example, synthetic apertures of 25 x 40 arcsecond have been used to map spiral galaxies (Bosma 1981).

Molecules, particularly metal oxides and carbon based molecules, have long been known to exist in the atmospheres of late type stars and so it is natural to expect that they should also exist in the ISM. The most common molecule is likely to be H_2, which was first detected in the ISM by Carruthers (1970) using satellite observations of electronic transitions in the ultraviolet. Spitzer et al (1973) observed the molecule in absorption towards 15 stars and in some of these H_2 was shown to be comparable to HI in column density (Jenkins and Savage 1974).

The development of millimetre wave techniques in astronomy has allowed many molecules to be studied by means of rotational transitions (e.g. Morris and Rickard 1982), but because H_2 is a symmetric diatomic molecule it cannot radiate energy by means of dipole rotational transitions (Herzberg 1950). A quadrupole transition is observed but only when the gas is shocked (Shull and Beckwith 1982). The ultraviolet transition is not suitable for the measurement of densities, unless small, because of the large extinction in the ultraviolet.

For these reasons H_2 is generally mapped by observing CO. This is a very stable molecule (its dissociation energy is 11.1 eV) made up of common elements and so it is the next most abundant molecule to H_2. The $J = 0 \rightarrow 1$ transition occurs at 2.6 mm and, although it is usually saturated, it is widely used for mapping the distribution of molecular gas.

The largest dishes used for molecular line work until very recently have been the 14 m Five College Radio Astronomy Observatory and the 11 m National Radio Astonomy Observatory dishes. These have diffraction limited beam sizes of about 1 arcminute, which for a distance of 20 Mpc corresponds to a dimension of 6 kpc. Observations of galaxies, particularly if inclined, therefore often suffer from poor spatial resolution, although in a few cases the velocity information can be used to provide greater resolution (Scoville et al 1983). Greater resolution is likely to be achieved in the future by creating interferometers and by the construction of larger dishes, but present molecular line data are restricted to resolutions of 50 arcsecods or larger.

An alternative method of mapping molecular gas at higher resolution is to make use of its known association with dust (e.g. Dickman 1978) and to map the distribution of interstellar reddening. Because the large column densities are likely to saturate optical colours, infrared colours are more suitable. At 3 um and longer wavelengths thermal emission from dust affects the colours of galaxies (e.g. Glass 1976) and so the 1 to 2 um region is most suitable. To achieve a long wavelength baseline, J-K ([1.2]-[2.2]) is likely to be the most suitable colour to measure.

Mapping the extinction also has the advantage that it is not luminosity dependent and so densities do not depend on the temperature of the ISM. However the geometrical distribution does affect the colours, making the derived densities lower limits; these effects are discussed in Chapter 6.

One of the most interesting features of the molecular gas distribution in the Milky Way is the 5.5 kpc molecular ring (e.g. Scoville and Solomon 1975). The use of the infrared extinction technique allows the possibility of similar features being observed in other galaxies, providing clues to the orgin of the feature. NGC 7331 is of similar Hubble type to the Milky Way and also has a similar HI distribution (Bosma 1981) and rotation curve (Rubin et al 1965). This galaxy was therefore selected for extensive observation. The distribution of the ISM in NGC 7331 is discussed and compared with the Milky Way and other galaxies in Chapter 7.

The earlier data of NGC 7331, presented in this thesis, showed the existence of a feature in NGC 7331 similar to the molecular ring of the Milky Way (Telesco, Gatley and Stewart 1982 : Appendix B). These data and further observations of NGC 7331 are examined in Chapter 7 in the light of the subsequent confirmation of this feature by CO measurements (Young and Scoville 1982c). The advantages of the higher resolution of the infrared data are demonstrated.

1.3 Infrared Surface Photometry Techniques

Optical surface photometry of galaxies can be performed efficiently by photographic techniques and recently by the use of solid state arrays (e.g. Murphy et al 1983). Many points can be measured simultaneously and under identical observing conditions. The number of resolution elements in a photographic image is determined largely by the atmospheric seeing. A galaxy subtending 50 square arcminutes on the sky can provide 2×10^5 resolution elements (pixels) if the seeing is 1 arcsecond. Even though the quantum efficiency of photographic emulsions is only a few percent, a large amount of information can be obtained. Solid state arrays such as charge coupled devices (CCDs) are now commonly used in the optical and have typically 3×10^5 pixels operating at high quantum efficiency and low noise.

Much work is currently being devoted to infrared arrays, but instruments offering good performance are not yet in general use. Surface photometry at infrared wavelengths must therefore be performed measuring points consecutively. This is time consuming but with a sensitive system on a large telescope (giving typically a signal to noise of 1 on a K=15 source in 1 second) a measurement accurate to 0.05 magnitudes may be made of a surface brightness of 17 mag's/arcsecond in a 5 arcsecond aperture in 1 minute. This surface brightness corresponds to a radial distance of about 6 kpc in a relatively nearby (20 Mpc distance) galaxy.

Consecutive measurements, in addition to increasing observing time, can also seriously degrade the accuracy of colour measurements. Brightess gradients in inclined galaxies can be in the order of 0.2 mag's/arcsecond and so small differences in the beam positions of the two measurements can generate spurious colours. To achieve a colour accuracy of 0.05 mag's in such a brightness gradient, requires that the positions of the two measurements should agree to better than 0.25 arcseconds. This is 4 times smaller than the telescope encoder resolution used until recently at UKIRT and can be an order of magnitude smaller than the pointing stability in the presence of guiding errors and windshake.

To determine colours accurately, therefore, measurements must be made simultaneously at two wavelengths. The construction of a cryostat measuring K and one other wavelength simultaneously is described in Chapter 3. The use of a common focal plane aperture ensures also that there are no systematic pointing differences between the two channels, which would cause a spurious colour proportional to the brightness gradient to be measured.

Good correspondence both in beam shape and size is required for the accurate measurement of surface colours. This is not important when measuring stars, but has been suggested (Jones et al 1982) as the cause of systematic differences between the colours of galaxies measured by Frogel et al (1978) and by Griersmith et al (1982).

Two versions of the cryostat were constructed. The Mark I version used 0.25 mm. diameter detectors with which it was difficult to obtain flat beam profiles. The simultaneity of measurement at the two wavelengths, however, allowed colour variatons to be mapped in galaxies. The Mark II version used 0.5 mm detectors and enabled well matched and flat beam profiles to be obtained. Surface brightnesses and colours suitable for comparison with other work were therefore attainable with this system, although the lack of a grid of red standard stars means that small colour transformations may be necessary for the slightly inclined filters.

The instrument was used to make two dimensional colour maps in J-K and H-K of three spiral galaxies and one dimensional scans of two further galaxies. The colour variations were the main object of study. Information was also obtained on luminosity gradients, but since it does not extend to large radii, is not considered in so much detail.

Although two wavelengths were measured simultaneously, different points and colours (J-K and H-K) were measured consecutively. The variations in effective beam profile, caused by varying performance of the chopping secondary mirror, are discussed in Chapter 4. The data were obtained both with conventional beamswitching techniques and by mapping the entire area without beamswitching. These techniques are discussed in Chapter 4.

The results are presented in Chapter 5. Because of its favourable inclination and similarity to the Milky Way, NGC 7331 was observed most extensively. The other galaxies observed are discussed in less detail.

2. Mechanisms for Near Infrared Luminosity in Galaxies

2.1 Introduction

This study is concerned with luminosity profiles and colour variations within galaxies and so it is necessary to understand in what types of objects the luminosity originates. The profiles and colour variations may then be used to deduce important properties of the galaxies. It will be demonstrated that the near-infrared luminosity is more representative of the mass distribution than the optical luminosity and that the colours can provide less ambiguous information in some cases than the optical colours.

At optical wavelengths the available resolution allows many luminous components of galaxies to be identified directly from photographs, for example luminous stars, globular clusters and HII regions. In all but the very closest galaxies however, stars are not luminous enough to be identified individually and the sources of luminosity must be identified by studying the broadband and spectroscopic properties of the integrated light. It has long been recognised that the light from galaxies is composite in nature and cannot be assigned a mean stellar spectral type (e.g. Johnson 1966a), since in different spectral regions the various sources contribute to the luminosity in different amounts. This also means that the dominant near-infrared sources of luminosity may be different to the optical sources.

Because infrared measurements have generally been made at low resolution a point at a time, only the integrated properties of galaxies have been determined. The sources of infrared luminosity can be tentatively identified by extrapolating the optical luminosity into the infrared, using models of the stellar content, and identifying other possible sources of luminosity of similar type to objects observed in the Galaxy. The expected radiation from these sources can be compared with infrared measurements to place constraints on the components of other galaxies. The total luminosity, the broad band colours and the spectroscopic properties can serve to identify the sources of luminosity. In addition, models of star formation rates and initial mass functions can also place constraints on the possible sources of luminosity.

Many of the data in this study have been obtained at three wavelengths, J, H and K, and this allows a two colour plot to be used to discriminate between different luminosity sources and to identify sources of colour variations. The possible sources of luminosity will be discussed paying attention to their location on the two colour diagram. The properties of the sources will be deduced from observations mostly of objects within the Galaxy, although in a few cases extragalactic observations of luminous objects also provide useful information. The independent colour data obtained in the current observations are J-K and H-K. The two colour diagrams used will therefore be J-K against H-K instead of the more customary J-H against H-K so that the errors in each coordinate will be independent.

Different types of sources will lie at different positions on the two colour diagram. Mixtures of two sources will lie between them on the two colour diagram but not necessarily along the straight line connecting them. If one of the two colours is the same in the two sources, or if the ratios of the colours in the two sources are equal, then varying mixtures will lie on the connecting straight line. In the general case, however they lie on a curve on the two colour diagram below the connecting straight line. An example of this is the well known fact that fact that a mixture of radiation from black bodies of two temperatures is different in colour to the radiation from a black body of intermediate temperature, even though black bodies form a straight line sequence on the two colour diagram. In the same way the radiation from a mixture of stars will not correspond on the two colour diagram to any particular stellar type, but will be displaced below the sequence. Fig. 2.1 illustrates the efects of mixing sources in different proportions. Such effects will introduce scatter on the two colour diagram in addition to colour variations, described in the following sections, intrinsic to the sources themselves.

Likely relative contributions from various sources to the disk luminosity of other galaxies will be deduced from their density in the solar neighbourhood and their absolute K magnitudes. This allows the likely composition of disk of other galaxies to be deduced in an approximate manner. The composition of the bulges of other galaxies is better known than that of the disks due to the relative absence of dust and greater homogeneity of stellar content. Comparison with elliptical galaxies and direct observations of the bulges both in the optical and infrared allow good estimates to be made of the sources of luminosity.

2.2 Main Sequence Stars

Stars spend most of their active lives on the main sequence and so, although they are more luminous at other stages in their evolution, a significant proportion of their energy output will occur when they are on the main sequence. This proportion will be less for the infrared than optical region of the spectrum, since evolved stars tend to have cooler photospheres than the main sequence stars from which they evolved.

Johnson (1966a) compiled a list of the intrinsic optical and infrared colours of representative stars, so that he could model the stellar content of galaxies. He considered main sequence stars of type B to G and giants of types K and M. He did not consider late type main sequence stars because their numbers were poorly known on account of their low luminosity and because studies (cf Arp, 1965) had shown that giants of this spectral type would be likely to dominate the luminosity of galaxy bulges. Johnson (1966b) did however compile a list of optical and infrared colours of main sequence stars between spectral types O5 and M8.

Unfortunately at this time Johnson did not use the H filter and so only the J-K colour was available in the near infrared. Veeder (1974) observed late main sequence stars in H and K (but not J) allowing their location on the two colour diagram to be determined. Studies of the lower main sequence, which is more important in infrared measurements than in optical, have also been made by Glass (1974), Mould and Hyland (1976) and Persson et al (1977). Late type dwarfs, like giants, lie above the black body locus in the two colour plot due to blanketting effects, particularly H- which has an opacity minimum in the H passband. Convection effects in dwarfs cause them to occupy a different region on the two colour plot to giants; the reduced temperature gradient causes a turnover in the J-H colour at later types.

In order to understand the source of infrared emission in early type galaxies, Frogel et al (1978) observed representative main sequence stars of type G5 and later in J,H and K and combined these observations with those of Johnson (1966b) and Veeder (1974) to allow a stellar sequence to be plotted on the (J-H),(H-K) diagram. Whittet and van Breda (1980) obtained colours for early type main sequence stars. These data and those of Koorneef (1983a) have been transformed onto the Johnson filter system (Section 4.4.1) to provide a stellar sequence on the two colour diagram.

This sequence (Koorneef 1983b) is similar to that given by Frogel et al after transformation, although there is some discrepancy in the position of spectral types along the sequence. The filter system used by Koorneef is similar to the UKIRT system (Section 4.4.1) and so Koorneef's data provide a framework useful for assessing the observed colours of the galaxies. Fig. 2.2 shows the location of this sequence on the J-K against H-K two colour diagram.

The likely contribution from main sequence stars of various spectral types may be roughly estimated by combining the number densities of stars in the solar neighbourhood with their absolute K magnitudes and scale height perpendicular to the disk. The absolute K magnitudes may be derived in many cases from the absolute V magnitudes (Allen 1973, p200) using the colours of Johnson (1966b) and Frogel et al (1978). Elias (1978a) has made such a compilation from which the absolute K magnitude integrated through the disk in an area of 1 kpc^2 may be estimated for each spectral interval (Table 2.1):

$$M_K \text{ (in 1 kpc}^2\text{)} = M_K \text{ (star)} + 2.5 \log(2.n.\beta)$$

where n is the number of stars kpc^{-3} in solar neighbourhood, and β is the scale height in kpc.

Elias did not include dwarfs later than K3 because their space density is poorly known. If a space density of 0.02 pc^{-3} (Sanduleak, 1976) and a scale height of 300 pc are assumed for early M dwarfs, then the integrated K magnitude is -12.7. For later dwarfs the uncertainties are even greater, but it is clear that the K and M dwarfs will dominate the K light compared to earlier dwarfs.

	M_K of star	log(n)	scale height	M_K of 1 kpc^2
B0,1	-3.12	2.90	0.04	-7.63
B2,3	-1.33	3.78	0.04	-8.04
B5	-0.53	4.04	0.04	-7.89
B8-A0	0.34	5.28	0.07	-10.73
A2-A5	1.33	5.49	0.10	-10.65
F0-5	1.87	6.04	0.14	-11.85
F8-G2	3.14	6.34	0.25	-11.96
G5	3.51	6.45	0.30	-12.06
G8-K3	4.15	6.95	0.35	-12.84
K5	4.3			-12.7
M0	5.7			
M2	6.2			
M5	7.3			
M8	11			

Table 2.1. The estimated integrated K magnitude due to dwarfs in an area 1 kpc^2 through the disk in the solar neighbourhood.

The sum of the listed luminosities amounts to an absolute K magnitude of -14.2. A typical distance of the galaxies observed is 20 Mpc and at this distance the observed K magnitude would be about 17.5 if the galaxy were observed face-on and the density of stars were similar to that of the solar neighbourhood.

Characteristics other than the spectral type can also affect the location of main sequence stars on the two colour plot, and should be considered as possible causes of colour variations if main sequence stars provide a significant proportion of the radiation. Mould and Hyland (1976) found that the (J-H),(H-K) diagram separated late type dwarfs into two populations differing by 0.1 magnitudes in

J-H. The bluer population consists of halo stars and stars with old disk motions. The models of Mould (1976) predicted that metal deficient stars would show such an effect due to higher pressures in the more transparent atmospheres, but Mould and Hyland could not come to any conclusion as to whether the old disk stars showed this effect on account of higher gravity or metal deficiency. They also identified some stars with J-H excesses of about 0.05 magnitudes and interpreted these as being stars with low gravity still contracting onto the main sequence.

Persson et al (1977) performed a similar study on late type dwarfs and found that in very late types the decrease in J-H does not continue past about H-K = 0.25. They did not confirm all the details of the Mould model but did come to the same conclusion that the metal deficient halo stars do have bluer J-H colours than the disk stars.

These metallicity effects will be difficult to recognise in the present study in the presence of other mechanisms for colour variations and the likely dominance of giants, but may be seen as scatter in the two colour diagram.

2.3 Giant Stars

Giant stars have been well studied in the infrared, being well suited to early observations because of their high infrared luminosity. Johnson (1966a and 1966b) measured giants from G5 to M6 at J and K and these data are still being used today as representative. His data did not include H measurements but these have been added by Lee (1970) for M giants and by Frogel et al (1978) for a complete representative grid with which to compare the colours of early type galaxies. Koorneef (1983b) has transformed photometric data, largely from Lee (1970) onto the Johnson system and these are shown in Fig. 2.2. These are consistent with those of Frogel et al (1978) when transformations are taken into account.

They occupy locations well away from the black body locus,

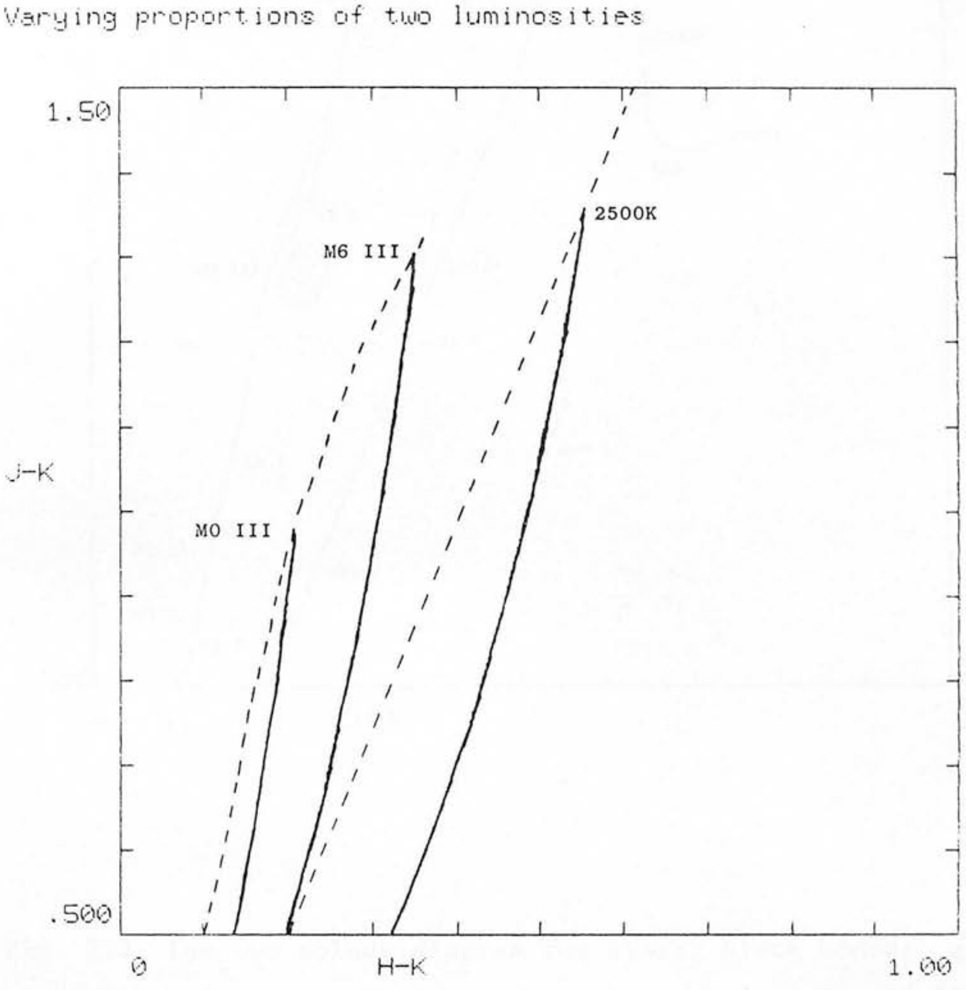


Fig. 2.1 The effect of mixing sources on the two colour diagram. The heavy lines represent the loci as the colour of an A0 star (or 11000K black body) is mixed with the colours of late type stars and a relatively cool black body. The loci are not straight nor do they follow the stellar or black body sequences.

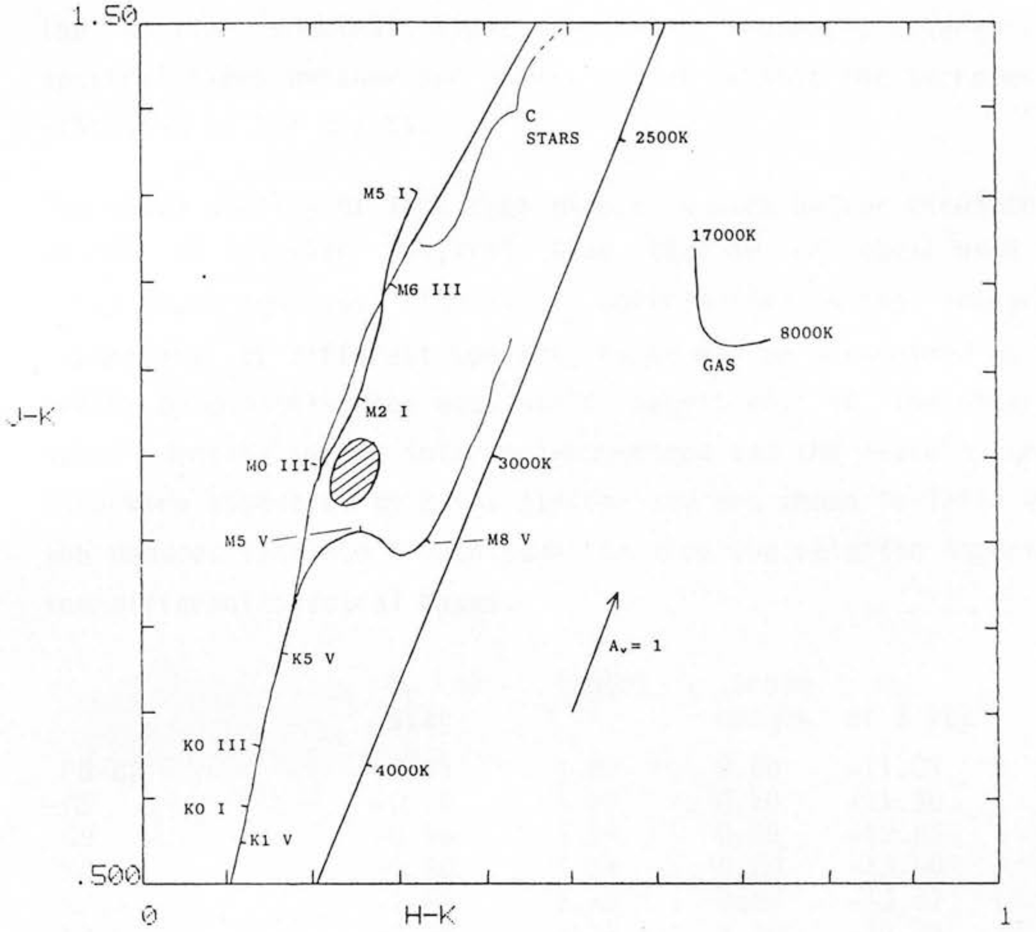


Fig. 2.2 The two colour diagram for stars, black bodies, gaseous emission and bulges of spiral galaxies. The stellar sequences are from Koorneef (1983b) except for the C star sequence which is from Cohen et al (1981). The gaseous emission locus is from Thuan (1983) and the reddening vector (A_V=1) is from Koorneef (1983b). The black body locus was calculated using the filter transmissions from Selby (1981). The hashed area represents the colours of the bulges of spiral galaxies (Griersmith et al 1982).

especially for later types, due to the blanketting in their atmospheres. They occupy very similar positions to the dwarfs for the earlier spectral types, but the sequences diverge at later spectral types because the giants do not exhibit the turnover in J-H exhibited by the dwarfs.

The space density of late type giants is much better known than for dwarfs of similar spectral type because of their much greater intrinsic brightness. The likely contribution to the integrated K luminosity of different spectral types may be determined as for the dwarfs by combining the absolute K magnitudes of the stars, the space density in the solar neighbourhood and the scale height. The data were assembled by Elias (1978a) and are shown in Table 2.2 with the deduced absolute K magnitudes to show the relative importance of the different spectral types.

	M_K of star	log(n)	scale height	M_K of 1 kpc^2
F8-G2	0.45	4.60	0.50	-11.05
G5	-0.08	4.60	0.50	-11.58
G8	-0.56	5.15	0.25	-12.68
K0,1	-0.80	5.54	0.20	-13.66
K2,3	-1.66	5.18	0.20	-13.62
K4,5	-3.36	4.23	0.30	-13.38
M0	-4.14	3.43	0.30	-12.16
M1	-4.40	3.08	0.30	-11.55
M2	-4.76	3.08	0.30	-11.91
M3	-5.23	3.08	0.30	-12.38
M4	-6.04	2.95	0.30	-12.86
M5	-6.90	2.95	0.30	-13.72
M6	-7.90	2.40	0.30	-13.35
M7	-8.90	2.04	0.30	-13.45
M8+	-9.90	1.60	0.30	-13.35

Table 2.2 The estimated integrated K magnitude due to giants in

These data, although only a rough guide, show that K and M giants dominate over earlier types. In addition, it can be deduced that giants will provide more light than dwarfs since the sum of the listed luminosities for giants amounts to an absolute K magnitude of -15.9 in an area 1 kpc^2. At a distance of 20 Mpc this is equivalent to an observed K magnitude of 15.6. These figures are 1.7 magnitudes brighter than the estimated K radiation from dwarfs and so giants can be expected to provide the majority of the light although dwarfs still provide an appreciable fraction. This analysis is relevant only to the disks of galaxies, but since the bulges are not sites of recent star formation the dominance of giants over dwarfs will be more pronounced.

Late type giants and dwarfs may be distinguished on the two colour diagram, but a less ambiguous method exists in the form of the CO index (Baldwin et al, 1973). This measures the strength of the bandhead of the first overtone of CO, which depends on luminosity and is thus much stronger in giants than dwarfs. Frogel et al (1978), while deriving their grid of representative stellar colours, also measured the CO index as a function of spectral type and luminosity class and demonstrated it to be a powerful luminosity discriminant for stars of type K0 and later. Observations of galaxies (Section 2.7) using the CO index confirm the dominance of giants in the integrated infrared light.

The effect of metallicity on the location of giants in the two colour diagram has been investigated by Cohen et al (1978), who have measured the infrared colours of globular clusters. They found that the globular cluster giants have redder J-H colours, by up to 0.15 magnitudes, for a given H-K colour than the field giants measured by Johnson (1966a) and Frogel et al (1978). This effect was also noticed by Glass and Feast (1973) who suggested that the cause was the reduced opacity of species relative to H-, although the model calculations of Cohen et al (1978) show that this would move the colours in the other direction. They suggest that the effect is due to neglected molecular species and may not be a simple function of metallicity since they do not see any ordering with respect to measured metallicities. It is clear however that different populations of giants may occupy slightly different locations on the two colour diagram.

Star counts at 2.4 um along the Galactic plane towards the bulge of the Galaxy (Kawara et al 1982) indicate that 63 percent of the light is contributed by K and M giants. Similar conclusions have also been reached by observations towards the Galactic anticentre (Hayakawa et al 1978) confirming that, both in bulge and disk regions, late type giants dominate the near infrared luminosity.

2.4 Supergiants

Supergiants, although not numerous, are very luminous and so must be considered as a potentially significant source of infrared luminosity. Because they are so short-lived, their locations will be confined to sites of recent star formation and so they may be important in interpreting the infrared colours of small regions of a galaxy even though their contribution to the integrated luminosity may be small.

Although very luminous, observations of supergiants have been very difficult because they are so rigidly confined to the plane of the Galaxy. Lee (1970) made what remains the standard study of supergiants in the infrared, a spectral region well suited to their study due to the high extinction they suffer. Although their intrinsic colours cannot usually be directly observed, their spectral type can, and Lee plotted supergiants of similar spectral type on a two colour diagram thus immediatley deducing an average reddening vector. For a very few stars intrinsic colours could be determined allowing the other stars of the same spectral type to be dereddened.

The intrinsic colours are based on only a few stars, but the locus they describe on the two colour diagram is plausible in that it lies very close to the giant locus. CO measurements (Baldwin et al, 1973, Frogel et al, 1978) show that the CO index is stonger in supergiants than giants but is not an effective luminosity discriminant due to variations produced by metallicity variations. This supports the view that the atmospheres of giants and supergiants are similar in appearance in the infrared. Class Ia supergiants lie below other luminosity classes in the two colour diagram due to an apparent flux deficiency at H. Their locations are intermediate between those of other supergiants and dwarfs.

Although it is primarily late type supergiants which have been considered as significant sources of infrared emission, the properties of earlier type supergiants should also be considered. Bouw and Parsons (1971) found that, within a single luminosity class, the absolute visual magnitudes of early type supergiants are relatively flat with increasing spectral type and comparable to the visual magnitudes deduced by Lee. Since the M supergiants have intrinsic V-K colours of at least 4 magnitudes, they are likely to dominate the early types unless they are much less mumerous. Humphreys (1970) found that the ratio of blue to red (K5 or later) supergiants is 8.8 for stars more massive than 20 solar masses and 2.4 for stars less massive. It is unlikely therefore that early type supergiants can be a significant source of near-infrared emission.

Because of their high luminosity, supergiants may be observed in nearby galaxies, allowing their absolute magnitudes to be determined more easily. Glass (1979a) has observed individual late type supergiants in the Magellanic Clouds. He found that their bolometric magnitudes decrease with increasing spectral type and that they form an increasing sequence in the order Milky Way, SMC and LMC. However Aaranson (1981) suggests that this may be a selection effect. The decreasing bolometric magnitudes with spectral type, if real, are not likely to make the early type supergiants significant in the near-infrared light. The intrinsic colours of these stars may be determined unambiguously since the extinction to the Magellanic Clouds is accurately known. Glass finds that the near infrared colours are compatible with those given by Lee.

Elias (1978a) quotes a representative absolute K magnitude of -10 for M type supergiants, although there is a large scatter in this due to the different luminosity classes and spectral types included. Lee (1970) deduces that the absolute K magnitudes of luminosity classes Ia and Ib are -11.2 and -8.3 respectively. At a distance of 20 Mpc 1000 supergiants, of absolute K magnitude -10, would be required in the beam to produce an apparent K magnitude of 14, but this number would be much less if very late type supergiants were present. Star counts at 2.4 µm (Kawara et al 1982) along the Galactic plane towards the bulge indicate that late type supergiants conribute only 7 percent of the light.

The M supergiants were proposed by Lee (1970) as a tracer of spiral structure and he identified spiral features in their longitudinal distribution, although the data are ambiguous. It is clear however that they do indicate recent star formation and may contribute to the integrated starlight in the spiral arms of other galaxies. They may show up as regions of enhanced emission on the underlying continuum of old disk stars. The findings of Persson et al (1983) support this possibility. They find that the V-K colours of the youngest (10^7 to 10^8 years) clusters in the Magellanic Clouds are 2 magnitudes redder than predicted by the star formation models of Struck-Marcell and Tinsley (1978) and conclude that this is due to the models neglecting the presence of red supergiants.

Observations at H and K in the Scutum region (Kawara et al 1982) led Mikami et al (1982) to conclude in the 5 kpc molecular ring region (Scoville and Solomon 1975) there is enhanced emission from late type supergiants. Kawara et al (1982) suggest that the effect may be local. Hamajima et al (1981) conclude that it is due to increased transparency of the interstellar medium in this direction since the space density of M giants is enhanced. The extinction measurements of Ichikawa et al (1982) support this explanation. However it is possible that supergiants could make significant contributions to the light in certain regions of galaxies with recent star formation. Because of their location on the two colour diagram, JHK colours cannot distinguish between such population changes and reddening.

2.5 Carbon Stars

The discussion of Section 2.3 concerned oxygen rich giants, which greatly dominate in number carbon rich giants in the Galaxy. This situation does not necessarily hold for other galaxies and so the properties of carbon stars should be considered.

Carbon stars have very complex spectra, containing features of molecules based on carbon rather than oxygen. These cause such effective blanketting that there is little or no continuuum in the spectrum to define a useful colour temperature. Because they cannot be classified as easily as oxygen rich giants, distances, intrinsic colours and absolute magnitudes are more difficult to determine. In addition, in the Milky Way they are much less numerous than oxygen rich giants (Blanco et al, 1978).

Their properties have been more easily assessed by observing them in the Magellanic Clouds and a few other nearby galaxies, for which distances and extinction can be accurately determined. Cohen et al (1981) and Blanco et al (1980) have measured the intrinsic colours of C stars in the Magellanic Clouds and the Milky Way and find that they form a continuation of the M giant sequence towards redder colours in both J-K and H-K.

They find differences towards the redder end of the sequence, where the galactic C stars are redder in J-H for a given H-K than the stars of the Magellanic Clouds. At the very red H-K colour of 0.6 (but typical for a C star), the difference is about 0.06 for LMC stars and 0.10 for SMC stars. This suggests a metallicity effect since the sequence SMC-LMC-MW increases in metallicity. Cohen et al (1981) suggest that this is the case and that the colours themselves are more a function of relative abundances than of temperature.

Several estimates have been made of the absolute magnitudes of C stars in the Milky Way and other galaxies. Cohen et al (1981) find that the typical absolute K magnitude of C stars in the Magellanic Clouds and the Milky Way is -8.5, making them about a magnitude brighter in K than an M5 giant. The brightest C star observed by Aaranson and Mould (1980) in the Fornax galaxy has an absolute bolometric magnitude of -5.65 and an absolute K magnitude of -8.6. At a distance of 20 Mpc, 580 such stars would need to be in the beam to produce an observed K magnitude of 16.

C stars are known to be rare in the Milky Way as shown by the work of Nassau and van Albada (1949), who found C stars only towards the anticentre of the Galaxy, and Blanco et al (1978), who performed star counts in the area of Baade's Window. They estimated the ratio of C stars to late M stars to be only 0.003 in the nuclear bulge and so on this basis C stars would not contribute significantly to the infrared emission of other galaxies. However in the LMC they found that the ratio was about 2:1 and in the SMC they found almost no M stars but many C stars. This demonstrates the dangers in extrapolating Milky Way star counts to other galaxies, although it can be argued that C stars and M giants are similar types of objects.

The uncertainties in estimating the proportion of C stars to M stars are further demonstrated by Persson et al (1983), who do not find any difference, between the Magellanic Clouds, in the proportion of K light contributed by C stars. They suggest that the findings of Blanco et al (1978) were affected by incompleteness in their surveys. They conclude that C stars are present in intermediate age clusters, but not in young clusters less than 10^9 years old.

C stars have also been observed in a few other nearby galaxies. Frogel et al (1980) and Aaranson and Mould (1980) have observed them in the Fornax and Sculptor dwarf galaxies and also find that their relative numbers in these galaxies, the Magellanic clouds and the Milky Way can be interpreted as a function of metallicity. They find the C stars have systematically different colours in different galaxies,consistent with the observations of Cohen et al (1981), and also interpret these differences as a metallicity effect.

On the (J-K),(H-K) two colour diagram, C stars continue the giant sequence farther into the red (Aaranson and Mould 1980, Cohen et al 1981). Cohen et al derive a linear relationship

$$(J-H) = 0.72\ (H-K) + 0.66$$

for the colours of galactic C stars over the range $0.3 < H-K < 0.8$. Transformed onto the UKIRT filter system (Section 4.4.1) this is plotted on the two colour diagram in Fig. 2.2.

Although C stars are much redder than typical M giants, it is possible that a late population of M giants could produce similar colours and so the (J-K),(H-K) diagram cannot unambiguously distinguish these two types of stars. In addition, extinction will redden early M giants towards the location of C stars in the two colour diagram and other information such as the optical appearance of obscuration or deviations in the luminosity profiles is needed to distinguish between C stars and extinction reddening.

The galaxies considered in this study are normal type spirals which can be expected to have undergone similar processing of material in stellar interiors as the Milky Way. They should not be metal poor as are the Magellanic Clouds and dwarf galaxies and so large numbers of C stars would not be expected.

2.6 Emission from Dust and Gas

In addition to stars, infrared radiation is produced by non-stellar sources in galaxies and may be detectable. The mechanisms, thermal emission from hot dust and emission from ionised gas, occur in the same types of objects and so will be considered together. At longer infrared wavelengths, 10 μm and beyond, some galaxies are known to emit the majority of their radiation as thermal emission from dust (Rieke and Lebofsky, 1979). The temperatures inferred from the 10um to 30um emission is several hundred Kelvin (Price, 1981), which implies that the dust is close to the luminosity source.

The total luminosity required to produce a detectable K flux can be calculated for different black body temperatures. Table 2.3 shows the number of solar luminosities required to give a K magnitude of 16 if the source is at a typical galactic distance of 20 Mpc. For this purpose a zero magnitude K star is assumed to generate a flux density of 4.1×10^{-14}W/cm^2/um at the effective wavelength of the K filter (Allen and Cragg, 1983), and a solar luminosity is taken as 3.83×10^{26}W (Allen 1973, p160).

Temperature (K)	$L/L_\odot$ at 20 Mpc to give K=16
2000	6.2×10^{6}
1500	6.1×10^{6}
1200	7.6×10^{6}
1000	1.1×10^{7}
700	1.4×10^{8}
500	4.7×10^{8}
300	3.0×10^{9}
200	2.0×10^{15}

Table 2.3 The luminosity required at a distance of 20 Mpc to produce a detectable K signal.

In order for thermal emission from dust to be significant at 2 μm, its temperature must be at least several hundred Kelvin. This implies that the dust must be in close proximity to stars since the radius R at which a grain would attain a temperature T, in the optically thin case, is given by

$$R = \sqrt{L \,/\, 16\pi\sigma T \, Q_a \,/ Q_e \, .\exp(-\tau)} \,/T^2$$

where σ is the Stefan Boltzmann constant and L is the luminosity. For temperatures at which the radiation would be detectable at K, a very large amount of dust must be close to the sources of luminosity. For example if one assumes Q_a equals Q_e and that T is 300K then R is 1 A.U. or 5×10^{-6}pc for a solar luminosity.

Ionised gas, such as is found in HII regions, emits radiation which can be detected in the near infrared (Willner et al 1972). This radiation originates in free-free and bound-free transitions in ionised hydrogen and helium, two photon emission from the 2s-1s transition of HI and emission from spectral lines. Such emission is more commonly observed in the radio spectrum, where there is negligible emission from thermal blackbody sources and the radiation is not extincted by dust, but it can also be detected in the infrared.

The location of gaseous emission on the two colour diagram is shown in Fig. 2.2, using the calculations of Thuan 1983 for a range of representative temperatures. Also shown is the location of black body thermal emission, calculated by integrating over the measured bandpasses of the UKIRT filters (Section 3.2.4).

Infrared observations of stars with circumstellar dust shells (Merrill and Stein, 1976) show that some stars do have shells emitting at a few hundred Kelvin, but fall far short of the required total luminosity to be detectable. The near infrared emission from T Tauri stars has been attributed to emission from gas (Rydgren, 1976), but it is likely that thermal emission from dust may be more significant in these objects. However the total luminosity of circumstellar shells and T Tauri stars are not likely to be sufficient to be detectable at 2 um or shorter wavelengths. Planetary nebulae also have insufficient luminosity (Willner et al 1972) to contribute significantly to the K flux.

HII regions show thermal emission from dust (Wynn-Williams and Becklin, 1974) but this is too cool, of the order of 100K, to provide detectable emission at 2 um. In the radio part of the spectrum they emit by free-free transitions and the spectrum of this emission should remain flat down to the near infrared. However the extinction associated with most of these HII regions is so large, even at 2 μm, that the near infrared radiation is thermalised to longer wavelengths, as demonstrated by the representative spectra of Wynn-Williams and Becklin. Although the intrinsic flux density is typically 10 Janskys at 2 μm, less than one tenth of this emerges from the region.

HII regions are confined to the plane of the Milky Way and so it is possible that the extinction may be much less if the regions were viewed from outside. Some HII regions have been observed in other galaxies and have been found to provide more flux in the near infrared. Several HII regions have been observed in M101 (Blitz et al, 1981); the flux density at H and K is consistent with that extrapolated from radio wavelengths and is interpreted as free-free emission. Thermal emission from hot dust dominates the spectrum at 10 and 20 μm but the blackbody curve, extrapolated to 2 μm, indicates much less flux than is observed. Unless another hotter dust component is present the near infrared emission is due to gaseous emission. The brightest of the observed regions provides 3.5 mJy at 2 μm; placing this at a distance of 20 Mpc, rather than

6.8 Mpc as for M101 (Aaranson et al, 1980), would generate an observed K magnitude of 15.5.

NGC 604, a large complex of HII regions in M33, has also been observed in the near infrared by Israel et al (1982b). The complex consists of 10 components of typical size 10 pc, embedded in an envelope of size 225 pc. It could therefore occupy much of the area seen in a 4 arcsecond aperture positioned on a galaxy of distance 20 Mpc. The detected emission is 21 mJy at H and K, which corresponds to a K magnitude of about 11.1. At a distance of 20 Mpc, rather than the 720 kpc of M33 (de Vaucouleurs, 1978), the K magnitude of this complex would be approximately 18.4.

An object such as NGC 604 would not be detectable in the observations of this study, but the M101 HII regions if arranged in a complex within the beam could be detectable. It is possible therefore that gaseous emission, or perhaps hot dust emission, could provide a significant proportion of the measured flux at a few points in the disks of the galaxies observed.

Large scale thermal emission from dust has been observed in the Milky Way by Price (1981) and Boisse et al (1981). Oda et al (1979) have determined the 2.4 μm flux distribution and shown it to be due to stellar emission, presumably from K and M giants, with little contribution by thermal emission from dust. Price (1981) has shown that the 4.2 μm flux distribution is very similar to the 2.4 μm distribution and quite unlike the longer wavelength distribution. It, therefore, is also stellar in origin and so it can be concluded that only a very small component of the shorter K radiation can be due to emission from dust. The same situation is likely to apply in other spiral galaxies.

2.7 Observed Infrared Emission from Galaxies

Galaxies were first observed in the near infrared by Johnson (1966a) whose aim was to determine the dominant sources of infrared emission. He concluded that in the galaxies he observed (ellipticals, spirals, irregulars and perturbed galaxies) the emission comes predominantly from K and M giants, although he also proposed a very late type population to explain the strong fluxes at L.

Glass (1976) has observed a large sample of spiral galaxies at J,H,K and L and finds that their colours are consistent with a dominant population of K and M giants, although he concludes that giant rich and dwarf rich populations cannot be distinguished by JHK photometry. In some cases he found that there was evidence for emission from warm dust affecting the L data, which he concluded was of limited use in studying the stellar emission. The J,H and K fluxes he interpreted as originating in stellar emission.

Measurements of the CO absorption, either by spectrophotometry or by narrow band colour indices, do discriminate between giants and dwarfs. These have been used to analyse the components of early type galaxies by e.g. Baldwin et al (1973) and Frogel et al (1978), who find that the infrared light of elliptical and SO galaxies is dominated by K and M type giants. Spiral bulges, although having different kinematics, have been shown to be very similar in stellar content to early type galaxies (Aaranson 1977). Using the CO index he has shown that the infrared emission is dominated by K and M giants rather than dwarfs. He was able to describe the optical and infrared colours by a simple mixture of early type dwarfs and late type giants. Even in the extremely dusty spiral, NGC 253, late type giants have been demonstrated to be responsible for the near infrared flux by the presence of a strong CO feature in near infrared spectrophotometry (Wynn-Williams et al, 1979).

Much of the interest in broad band colours of galaxies has been concerned with metallicity variations both within and between galaxies. Early type galaxies have been observed more extensively in the near infrared than spirals, largely because extinction and different populations in the disk do not complicate the study of such metallicity variations.

Radial colour variations have been studied in early type galaxies by Strom et al (1978) and Frogel et al (1978) using multi-aperture techniques. They find that the U-V and V-K and J-K colours redden with decreasing aperture. Frogel et al also measured the CO index as a function of aperture and found little correlation. By comparision with infrared observations of globular clusters, whose metallicities have been determined by optical spectroscopy, they conclude that the colour variations in galaxies are due to increasing metallicity with decreasing radial distance. The CO index is affected less by metallicity than luminosity, since the band heads in giant stars are saturated, and the results show that giants dominate the emission in all apertures.

The effect of metallicity on the infrared colours of early type galaxies and globular clusters has also been investigated by Aaranson et al (1978). They find that the optical and infrared colours of the globular clusters become redder with increasing metallicity due to an increase in line blanketting, a shifting of the horizontal branch from the red to the blue and a cooling of the giant branch. They find a strong correlation with galactic luminosity of the U-V, V-K and J-K colours and the CO index. Although the latter is primarily sensitive to the stellar luminosity, it also has a dependence on metallicity. They derive a relationship between J-K and metallicity

$$J-K = 0.85(+0.03) + 0.14(+0.02)\,[M/H]$$

for $-2.2 < [M/H] < 0$. Their investigation did not include H measurements and so no metallicity vector can be derived for the two colour diagram. This effect however is only evident in a large sample and individual data will have a large scatter due to variations in the mix of stellar spectral types.

Although most studies of colour variations within and between galaxies have been concerned with ellipticals and early type galaxies, spiral galaxies have also been observed. Aaranson (1977) has made extensive multi-aperture observations of spiral galaxies using apertures in which the measured light originates mostly in the bulges. His results show that spiral bulges are very similar in composition to early type galaxies (Frogel et al 1978), although the dynamics are different.

In the bulges of spirals, Aaranson (1977) found the optical colours redden with decreasing aperture, suggesting increasing metallicity as for the early type galaxies. He did not detect any radial variations in either the CO index or the near infrared colours. The latter was surprising since it was observed in early type galaxies, but perhaps demonstrates the problems associated with multi-aperture work of spiral galaxies.

Interpreting radial colour variations in spiral galaxies is difficult because of extinction problems and differing bulge disk ratios in the multi-aperture measurements. Griersmith et al (1982) have attempted to circumvent this problem by making multi-aperture observations of a large number of early type spiral galaxies, which are bulge dominated over a larger aperture range and suffer less internal extinction than Aaranson's sample of galaxies. They generally confirm the conclusions of Aaranson, finding that the U-V, V-K and J-K colours all redden with decreasing aperture, but to a lesser extent than in elliptical and SO galaxies.

In the multi-aperture observations of both Aaranson and Griersmith et al there are some galaxies with anomolous positions in the two colour diagram. Aaranson identified some of these "K excess" galaxies with very inclined systems such as NGC 5907 where extinction may be significant. In others, such as NGC 1068, there is evidence for thermal emission from warm dust although normal spirals do not show this effect. The location of the galaxies observed by Griersmith et al on the two colour diagram are shown in Fig. 2.2. These colours have been corrected for galactic reddening and redshift and have been interpolated to a standard aperture. The normal galaxies generally lie towards the bottom of the sequence whereas emission line objects are redder in colour, due largely to the effects of extinction (Griersmith et al 1982).

For most of the infrared excess galaxies of the sample of Griersmith et al, the infrared colours can be explained by a moderate (visual extinction about unity) amount of extinction. The optical colours require the addition of a young population of stars and gaseous emission in addition to the extinction. This is consistent with the infrared colours since the only young stars likely to contribute to the infrared flux are supergiants with very similar colours to the K and M giants. The radiation from the gas is much more significant in the optical due to the line emission. They rule out very late M giants or C stars as the origin of the infrared excess since they would have no affect on the anomolous optical colours.

Both these studies demonstrate the difficulties encountered in interpreting multi-aperture measurements of spiral galaxies even when some care is taken in their selection. In addition the flux in successive annuli becomes a smaller proportion of the total as the radius is increased causing a dilution of the colour variations. It is clearly preferable to perform surface colour photometry to investigate radial variations.

A small amount of near infrared surface photometry has been performed on spiral galaxies. Hackwell and Schweizer (1983) demonstrated that infrared mapping can identify components of galaxies not prominent at optical wavelengths. In an H map of NGC 1566, they showed that important of components of galaxies may be more prominent in the infrared than in the optical. The H map showed the presence of a bar which was not obvious in the optical data (Sandage and Tamann 1981). Spiral features were also present indicating that the density of old stars was enhanced in at least some types of spiral arms.

NGC 5128 has been mapped at J, H and K (Harding et al 1981) in a 30 arcsecond aperture showing that previous estimates of the dust density in the dark lane were low due to saturation at optical wavelengths. Higher resolution mapping at K (Adams et al 1983) showed evidence for stellar emission associated with the dust lane. In the galaxy itself colour gradients were detected, which became bluer in V-K with increasing radius, consistent with the multi-aperture measurements of normal galaxies.

However little near infrared surface photometry has been performed on normal spiral galaxies. Infrared colours can provide information on dust characteristics and distribution, which may not be apparent from optical data due to saturation. Colour variations due to metallicity and population gradients may also be apparent.

3. The Instrument

3.1 Advantages of Simultaneous Channels

Colour photometry in the infrared is normally performed by making sequential measurements through two filters. The colour of the source is then calculated from the ratio of these two measurements. For such colours to be accurate the following requirements must be met.

(a) If the source is resolved in the beam, the profiles must be flat over a region larger than the angular size of the source.

(b) If the source is extended, the beam widths and shapes in the two channels must not differ significantly. The permissible error is dependent on the brightness gradient of the source.

(c) The source must not vary between between consecutive measurements.

(d) Instrumental and atmospheric factors must be the same for the two measurements. These include atmospheric transparency, seeing if the apertures are small, and chopper characteristics.

For two classes of object these criteria may not be met. Some objects, such as cataclysmic variables, show rapid fluctuations in brightness of the order of tens of seconds. If 20 seconds or a minute were required to attain an acceptable photometric accuracy, then the intrinsic brightness of the source could vary significantly between consecutive observations in different filters and the ratio of these measurements would have little meaning. If the two measurements are made simultaneously then the ratio gives a valid estimate of the average colour during the integration.

Extended objects such as galaxies or emission nebulae, can have steep brightness gradients so that the measured brightness is very dependent on the position. Guiding errors and windshake can cause the effective position measured to be slightly different between the two observations and so again the ratio of these two measurements may not provide an accurate estimate of the colour of the object. In the inner regions of galaxies, brightness gradients of 0.2 magnitudes per arcsecond are typical (e.g. NGC 5907, Chapter 5). In such cases even very small positional errors will cause large colour errors. Even using 22 bit resolution telescope encoders the pointing accuracy is limited to 0.31 arcseconds and so the measured colour could only be accurate to 0.06 magnitudes. Windshake and seeing fluctuations can have a much larger effect, making colour accuracies of better than 0.1 mag's difficult to achive in regions of large brightness gradients.

Various other instrumental effects can also cause consecutive observations to provide less accurate colour information even when observing unresolved sources which do not vary with time. If a small aperture is being used in windy conditions, then the source may occasionally be removed from the beam and the photometric accuracy is therefore affected. Colours derived from such data will have errors equal to or greater than the sum of the errors associated with each filter. If simultaneous measurements are made however, each wavelength is affected equally and so the derived colour will have much higher accuracy than the individual measurements of brightness.

A two channel cryostat operating in the near infrared has been used by Elias (1978) to search for sources embedded within dark clouds. A three channel system has been used by Kawara et al (1982) to perform systematic star counts. For these purposes close correspondence of beams is not so important and separate apertures can, if necessary, be used. While still in development the cryostat described here was successfully used to search for embedded sources within the Orion Molecular Cloud (Lonsdale et al 1982) in spite of poorly matched profiles. For accurate brightness and colour surface photometry, however, care must be taken to ensure close correspondence between the beams.

The radiation falling on the detector is modulated by the chopping secondary mirror, so that the detector sees the source for less than half the observing time. The proportion of the time the detector sees the source depends on the rise time of the chopper, the degree of "droop" in the chopper waveform when the mirror should be stationary and the aperture size being used. If the chop amplitude is large and the aperture is small then this proportion can be significantly less than 50 percent. For typical observations a figure of about 40 percent is often achieved. If this figure changed to 38 percent between consecutive observations then a colour and brightness error of 5 percent would be introduced. Simultaneous measurements prevent the colour error being affected in this way.

The observations in this study are concerned with colour variations within galaxies. Some of them were made when the chopper performance was poor and variable, or when the wind caused the image of the star to move by several arcseconds. Nevertheless accurate colours were achieved which were not affected by pointing errors or chopper variations due to the simultaneity of the measurements. This is demonstrated by the data taken with the galactic nuclei on the edge of the aperture (Appendix A). Statistical errors in the measured brightnesses are typically 5 percent in such cases, whereas the colour errors are less than 1 percent.

In addition to the important improvement in the accuracy of the measured colours, there is the obvious saving in observing time since the measurements do not need to be repeated through a second filter. This time saving is maximum when the integration time required to attain the desired accuracy is the same for each filter. As shown in Section 3.6.2 this is approximately true for the J, H and K filters of the system which was built.

The system was initially designed around the available Santa Barbera Research Center (SBRC) 0.25 mm. diameter detectors. Good data were obtained with this system but it suffered from limited aperture sizes and profiles which were not flat-topped. For these reasons, the system was redesigned around larger 0.5 mm. detectors which became available from Cincinnati Electronics (CE). These two versions are referred to as the "Mark I" and "Mark II" cryostats respectively.

3.2 Optical Design

3.2.1 Basic Concept

In order to ensure that both detectors of a two channel system see the same area on the sky, a common focal plane aperture is desirable. The alternative, to have a separate aperture for each channel, would have required very careful alignment to measure colours in the large brightness gradients of galaxies.

Having a common focal plane aperture however, requires that a beamsplitter is positioned between the aperture and the detectors. The beam size is therefore increased by an amount x/f, where x is the distance between the focal plane and the detector and f is the focal ratio of the telescope. At the time of construction two focal ratios, f/9 and f/35, were commonly used and a decision had to be made between them. For a distance x of 80 mm, the beamspreads for f/35 and f/9 focal ratios were 2.3 mm and 8.9 mm respectively. The available detectors were 0.25 mm in diameter and so a beamspread of 8.9 mm could not be accommodated using a standard Fabry lens system, since the lens would have had to operate faster than f/0.25. The only way an f/9 system could be designed therefore would be to employ re-imaging with consequent alignment problems and loss of light in the additional optical components. It was therefore decided to built the system to operate with the f/35 secondary mirror.

It is necessary to focus the radiation onto the detector using some form of optics inside the cryostat. The alternatives were off-axis mirrors or on-axis lenses. Mirrors have the advantage of being achromatic and so the system could be accurately aligned while warm using light. On the other hand the fast off-axis paraboloids required would have been expensive and in any case the alignment would still have been difficult because of the inaccessability of the detectors close to the fast optics. In the 1 to 3 μm region, zinc selenide has high transparency and low dispersion (Wolfe 1978). It is therefore well suited as a material for manufacturing lenses and many ranges of diameter and focal lengths may be obtained easily and cheaply from optical component manufacturers. The refractive index in the visible (2.61), is significantly different to that in the near infrared (2.45) and so alignment adjustments made using light are subject to some uncertainty. Because no off-axis components were involved it was not anticipated that this would present serious problems, and so the optical components chosen were zinc selenide lenses.

With an f/35 telescope beam, and a distance x of 80 mm, the size of the beam at the lens is 2.3 mm plus the size of the focal plane aperture. This had to be imaged onto the 0.25 mm detector without loss of light. To obtain flat beam profiles, as the image of the star moves in the focal plane aperture, it is required that the image on the detector does not change position so that variations in sensitivity across the detector have little effect. To achieve this it is customary to focus the secondary mirror, rather than the image of the star itself, onto the detector. Since the secondary mirror is at a large distance from the detector, the detector is placed at the focus of the lens. By simple geometrical optics, it can be shown that the size of the image in the plane of the detector is F/f, where F is the focal length of the lens and f is the focal ratio of the telescope. For an f/35 system therefore, to produce an image filling the detector requires a lens of focal length 8.75 mm.

Rather than considering rays falling onto the detector, the optical system can be traced backwards so that the detector presents an acceptance cone to the secondary mirror. A lens of focal length f/F presents an acceptance cone of exactly focal ratio f (f/35 in this case), neglecting aberrations. To ensure that all the light falls onto the detector, the system must be aligned on the telescope so that the acceptance cone is pointed accurately at the secondary mirror. This is possible, if difficult, with a single channel system but for a two channel system inevitably the acceptance cones of the two detectors would not be exactly colinear and a compromise would have to be achieved with neither detector gathering all the light from the secondary. In addition, any flexure in such a system would increase the signal in one channel and decrease it in the other causing a large error in the colour.

The Mark I cryostat optical system was designed therefore, so that one acceptance cone was fairly closely matched to the focal ratio of the telescope and the other acceptance cone was broader. This was to allow the alignment on the telescope to be carried out by pointing the narrower acceptance cone at the secondary mirror. The broader acceptance cone would then also include the secondary without any compromises being necessary. The focal lengths chosen for the lenses were 6.94 and 8.29 mm, producing acceptance cones of f/27.8 and f/33.1 for the two detectors. The parameters of the lenses, at 1.65 um, are given in Table 3.1.

	Lens 1	Lens 2	
radius of curvature =	10.00	11.95	mm.
focal length =	6.88	8.22	mm.
diameter =	12.70	12.70	mm.
centre thickness =	3.78	3.35	mm.
back focal length =	5.34	6.85	mm.

Table 3.1 The design parameters of the zinc selenide lenses used in the Mark I cryostat.

The focal lengths were calculated using the paraxial approximation i.e.

focal length = r/(n-1)

back focal length = r/(n-1) -t/n

where r is the radius of curvature of the front convex surface, n is the index of refraction (2.454 at 1.6 μm) and t is the centre thickness of the lens.

With simple on-axis lens systems of this type, two types of aberrations can be important, viz. chromatic aberration and spherical aberration. These are considered in the following sections.

The internal layout of the cryostat is shown in Fig. 3.1 and although the Mark II cryostat is shown, the Mark I layout was very similar. The focal plane aperture is defined by the aperture wheel near the window of the cryostat. A central beamsplitter transmits the K radiation to one detector and reflects other wavelengths through a filter wheel to the second detector. The baffling around the aperture wheel is shown on the diagram, but for clarity a second level of baffling is not shown. This encloses the optical components and the optical paths ensuring that the detectors see radiation only thorugh the focal plane aperture. The detector and lens for each channel are held in an assembly to which is attached the cold electronics. The external connections comprise electrical sockets for each channel and drive shafts for the filter and aperture wheels.

3.2.2 Spherical Aberration

Spherical aberration becomes important when rays enter the system a significant height above the optical axis and the paraxial approximation for tracing rays no longer applies. In the system described here, this is the case since for an aperture of 5 arcseconds on the sky the linear aperture in the focal plane is 3 mm, which becomes 5 mm at the first surface of the lens due to the beamspread. The lens therefore must operate at a very fast focal ratio. The aberration will be described first using some approximations and then using the results of a computer ray tracing program.

Although these figures refer to the f/35 system, spherical aberration would also be significant with an f/9 system even though the linear aperture would only be about 0.9 mm. This is because the optical throughput (etendu) of any system must be conserved, i.e.

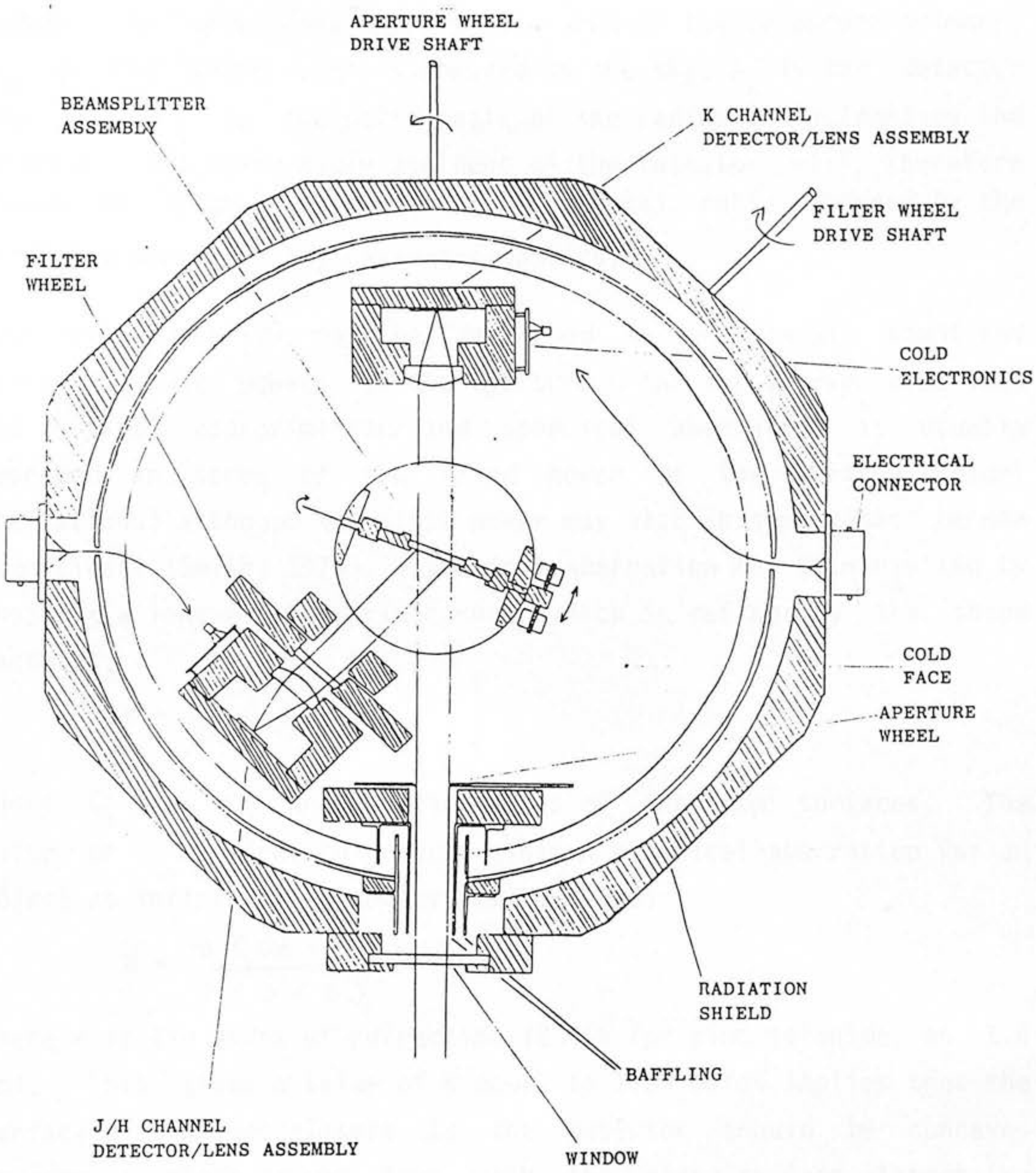


Fig. 3.1 The internal layout of the two channel In Sb cryostat. The aperture and filter wheels are driven by shafts extending through the cryostat wall (not shown). Also omitted from the diagram is a secondary stage of baffling around the optical components and light paths. The components mounted on the cold face are held at solid nitrogen temperature. These are enclosed by the radiation shield at liquid nitrogen temperature.

$$A_1\Omega_1 = A_2\Omega_2 = \text{constant}$$

where A is the area and Ω the solid angle at any point in the system. In this case A_1 is the area of the telescope primary, Ω_1 is the solid angle subtended on the sky, A_2 is the detector area and Ω_2 is the solid angle of the radiation incident on the detector. The solid angle incident on the detector will therefore always be large regardless of the focal ratio produced by the secondary mirror or any re-imaging optics.

Spherical aberation may be described by polynomial equations containing odd powers of the aperture. The first power describes the paraxial approximation and spherical aberration is usually desribed in terms of the third power of the aperture (Seidel aberrations) although the fifth power may also in some cases become significant (Smith, 1978). Spherical aberration may be minimised by choosing a lens of appropriate shape which is defined by the shape factor K,

$$K = \frac{C_1}{C_1 - C_2}$$

where C_1 and C_2 are the curvatures of the two surfaces. The value of K required to produce minimum spherical aberration for an object at infinity is given by (Smith, 1978)

$$K = \frac{n(2n+1)}{2(n+2)}$$

where n is the index of refraction (2.456 for zinc selenide at 1.6 μm). This gives a value of K equal to 1.63 which implies that the surface of the lens closest to the detector should be concave. However a plano-convex lens with the plane surface closest the detector (K = 1) is not too far from the optimum lens shape. The angular blur β due to third order spherical aberration is given by

$$\beta = \frac{r^3 \phi^3}{4(n-1)^2}\left[n^2 - (2n+1)K + \frac{(n+2)}{n} K^2\right]$$

where r is the semi-aperture and ϕ is the reciprocal focal length.

Calculating β for the optimum lens and the plano-convex lens gives:

$$\beta = 0.143 \; r^3 \phi^3 \qquad \text{(optimum)}$$
$$\beta = 0.228 \; r^3 \phi^3 \qquad \text{(plano-convex)}$$

The linear blur may be calculated from

$$b = \beta . f$$

where f is the focal length of the lens. Thus

$$b = 0.228 \; r^3/f^2 \qquad \text{(plano-convex)}$$

This linear blur is tabulated below for different semi-apertures in the focal plane assuming that the focal length of the lens is 6.94 mm (the shorter of the two) and that the beamspread adds 1.14 mm to the semi-aperture.

Semi-aperture (mm)	Sky aperture (arcseconds)	Blur (mm)
0	0	0.007
0.5	1.6	0.021
1.0	3.2	0.046
1.5	4.8	0.087
2.0	6.4	0.147
2.5	8.0	0.228

Table 3.2 The linear blur due to third order spherical aberration for the plano-convex lens of focal length 6.94 mm with the object at infinity.

Since the detector diameter is 0.25 mm and the size of the image calculated from the paraxial approximation is 0.20 mm (the focal length divided by the focal ratio), using this system it should be possible for apertures of 4 or 5 arcseconds to be used without significant loss of radiation.

This analytical approach is subject to certain assumptions. The object is assumed to be at infinity and on-axis, whereas the secondary mirror is being imaged onto the detector and subtends an angle corresponding to the f/35 beam. Fifth order aberrations are neglected and the thickness of the lens (about half the focal length) may also have effects.

To avoid these approximations and also to estimate the effect of alignment errors on beam profiles, a ray tracing program was written using the technique described by Nussbaum and Phillips (1976). The program considers rays in the meridional plane and estimates the proportion falling on the detector for an off-axis object in the focal plane aperture emitting rays in an f/35 beam. These rays were assumed to originate in a beam of angle 1.6 degrees distributed in such a way in the meridional plane as to take into account the circular nature of the beam. The circular shape of the detector was neglected but this assumption will have only a small effect on the computed profiles, tending to overestimate slightly the number of rays at the periphery of the image likely to miss the detector.

The souce was assumed to be scanned across the aperture and so this represents the beam profile that could be obtained on the sky assuming a small stellar image with no aperture stop in the focal plane of the telescope. The important computed quantity is the distance off-axis a stellar image can be in the focal plane without radiation being lost by not falling on the detector. The effect on the beam profile of an off-axis detector and an out of focus detector were also computed.

Fig. 3.2 shows the beam profile that would be obtained scanning a star image across the focal plane. The signal is plotted as a function of the distance in millimetres the source is off axis. It is seen that no signal is lost until the image is about 1.8 mm off axis, corresponding to an aperture on the sky of 5.8 arcseconds,

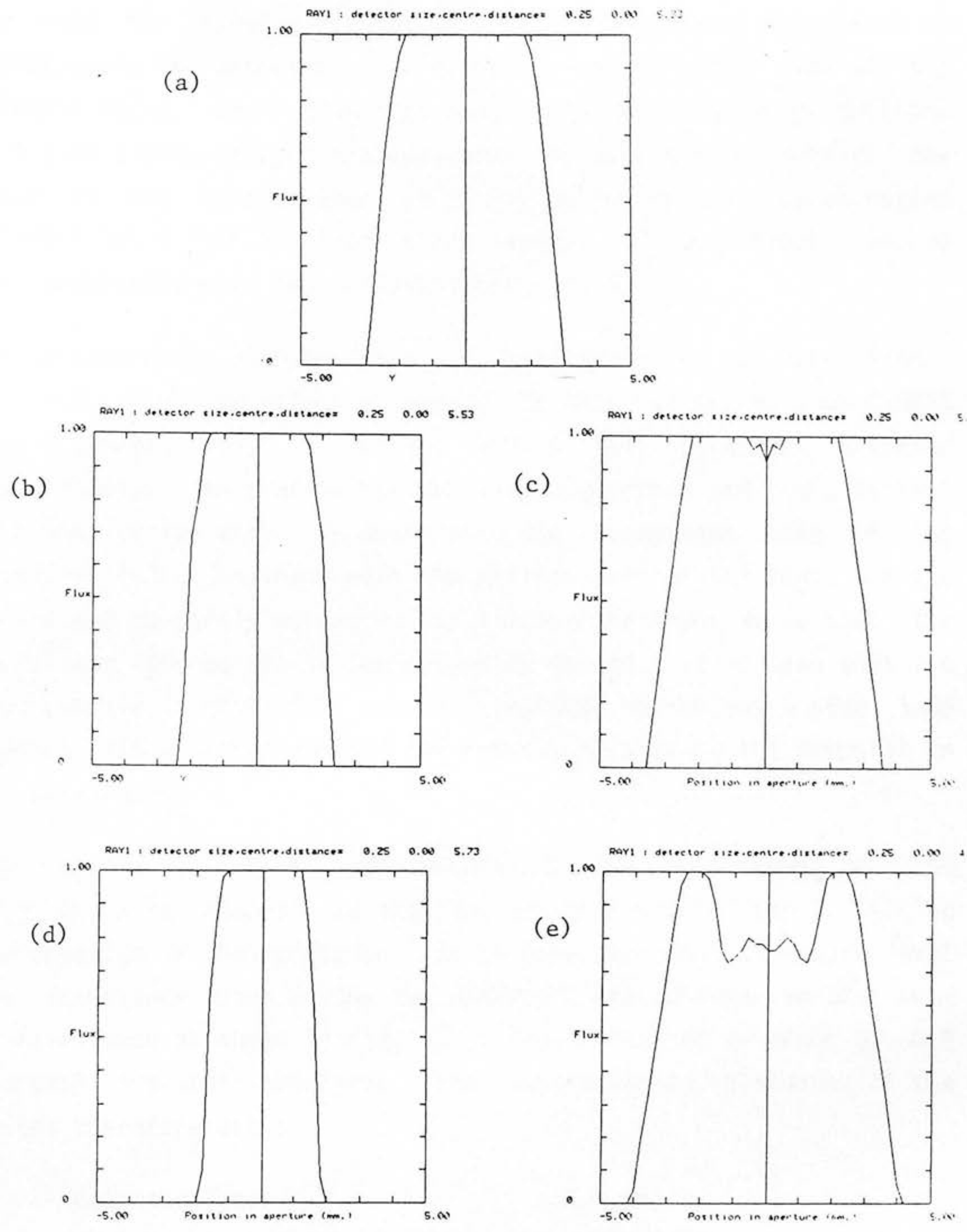


Fig. 3.2 The computed effect on the sky beam profiles of the Mk I cryostat of detector/lens focussing errors. (a) represents the correctly focussed profile. (b), (c), (d) and (e) represent focussing errors of .2, −.2, .4 and −.4 mm respectively. Each box shows the detector response as the source is moved off-axis between −5 and 5 mm in the focal plane.

consistent with the analysis above. As the source goes beyond this region the light falling on the detector declines rapidly indicating the need for careful alignment. Fig 3.2 also shows the effect of malfocussing the detector by an amount 0.2 mm on either side of the paraxial focus. The region seen completely by the detector declines to 1.5 mm either side of the axis when the detector is outside the focus of the lens. When it is 0.2 mm inside the covered region broadens but a hole develops in the centre. These effects become more pronounced when the focussing error is 0.4 mm.

The detector must also be closely on-axis so as not to lose light. Fig. 3.3 shows the effect of moving the detector off-axis by 0.0625 and 0.125 mm, i.e. a quarter and a half detector diameter respectively. The profile becomes very asymmetric and light is lost over most of the area. In such cases the acceptance cone of the detector is not colinear with the optical axis of the lens, but the effect can be partly corrected for tilting the cryostat so that the acceptance cone points at the secondary mirror. It is seen that the beam profile is still flat over an aceptable region but a side lobe appears which could confuse the results of tilting the cryostat on the telescope.

This calculation shows that errors in the positioning of the detector with respect to the lens can be corrected for by tilting the cryostat on the telescope. It is more important to ensure that the acceptance cones of the two detectors are colinear so that beam profiles such as shown in Fig. 3.3d (equivalent of an error of 0.5 degrees) are not obtained. The manufacturing tolerances of the system therefore are :

detector focus : 0.1 mm
detector transverse position : 0.05 mm
colinearity of acceptance cones : 0.3 degrees

Although these tolerances are very tight, it was considered that

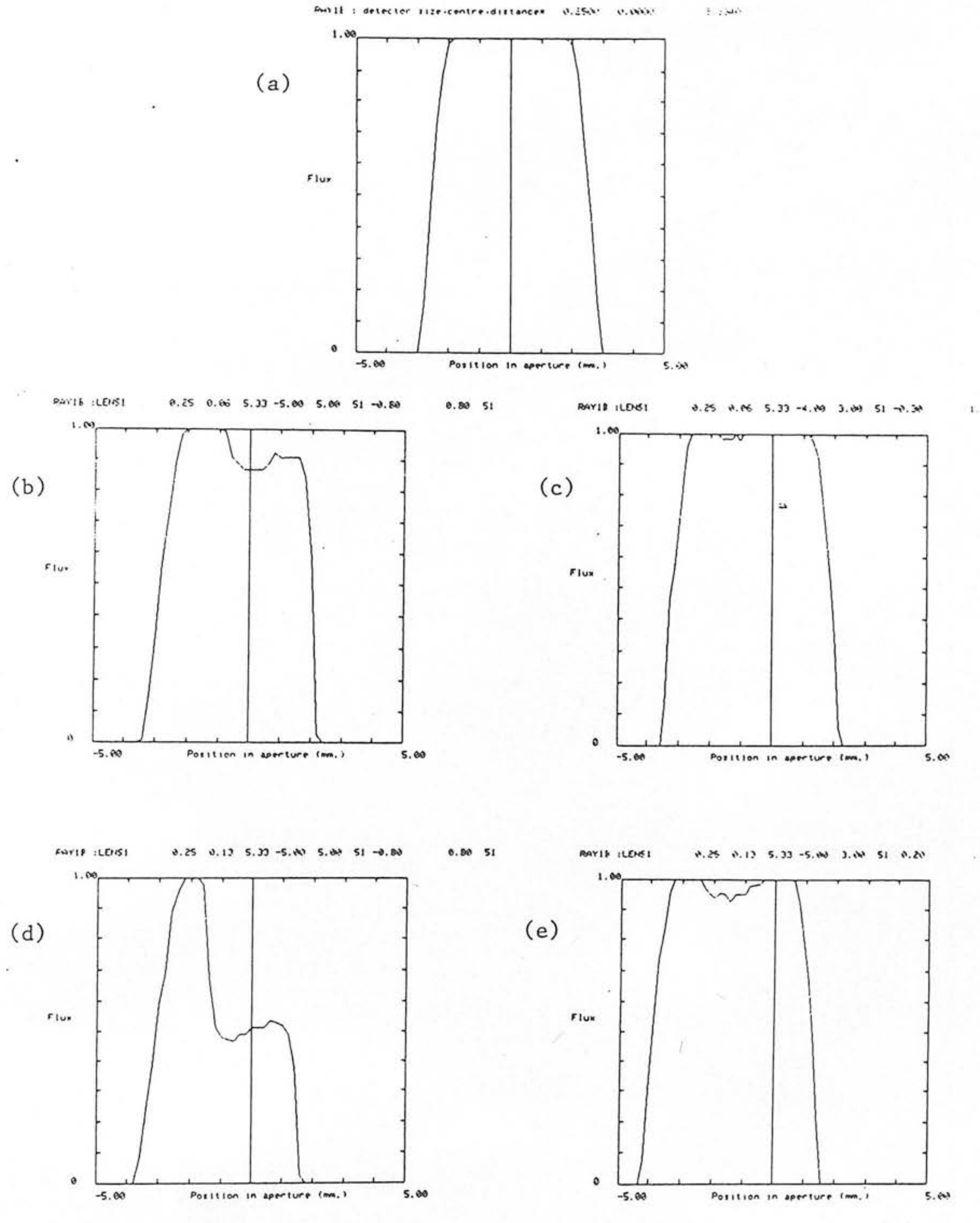


Fig. 3.3 The effect on the sky beam profiles of the Mk. I cryostat of displacing the detector from the optical axis of the lens. (a) is the on-axis profile and (b) and (d) are the profiles when the detector is 0.0625 and 0.125 mm. off-axis respectively. The profiles can be partially restored by tilting the cryostat as shown in (c) and (e).

such a design was feasible using the available 0.25 mm detectors.

3.2.3 Chromatic Aberration

Because refractive indices are wavelength dependent, different wavelengths of radiation are focussed at different positions. A well-focussed image at one wavelength will therefore be blurred at another. Wolfe (1978) shows that the refractive index of zinc selenide has a low dispersion for infrared wavelengths, although it changes rapidly in the optical. Using the wavelength dependent refractive index, the focal lengths of the lenses may be calculated. The results for the faster lens used at J and H are shown below. The blur sizes, calculated assuming that the system is focussed at 1.65 μm, are small compared with those due to spherical aberation and the differences in focal length are less than the manufacturing tolerances. The blur size is also indicated for a visible wavelength, such as was used for initial alignment and it is seen that the spot is typically larger than the detector diameter.

Wavelength um	Refractive index	Focal length mm	Blur sizes mm
0.6	2.613	6.201	(0.306)
1.25	2.470	6.804	0.032
1.65	2.454	6.875	0.000

Table 3.3 The effect of wavelength on the refractive index of zinc selenide and the focal length of the lenses. The blur size is calculated for a 5 arcsecond aperture, assuming that the system is focussed at 1.65 μm.

The effect on the beam profiles is important since it is necessary that the beam sizes are the same. The difference in focal length between the J and H wavelengths is 0.07 mm which does not produce a significant change in beam profiles (Fig. 3.2) on the sky. Chromatic aberration can therefore be neglected compared to the significant spherical aberration.

Of more inconvenience is the large difference of refractive index between optical and infrared, wavelengths. At 0.6 µm, the wavelength of a helium neon laser, the focus of the lens is at a significantly different position to the infrared focus. This has implications for the alignment procedure as explained in Section 3.4.

3.2.4 Beamsplitter and Filters

Near infrared filters are multi-layer interference filters and so that part of the spectrum which is outside the passband is reflected. This means that such a filter may also be used to as a dichroic beamsplitter, with other filters placed in the reflected beam so long as they do not overlap in wavelength with the transmitted radiation.

In order to place components in the reflected beam, without causing obscuration, the beamsplitter must be placed at non-normal incidence. The cryostat has an octagonal cross section which means that the most convenient angles for the reflected beam are 45 and 90 degrees; this allows the filter shaft to be perpendicular to the cryostat wall. To reduce wavelength shifts due to the non-normal incidence a 45 degree deflection was selected, so that the filter was placed at 22.5 degrees to the incident beam. The pass bands of interference filters shift to shorter wavelengths when the filters are used at non-normal incidence, but data supplied by OCLI indicate that this shift will be in the order of 0.5 to 1 percent of the central wavelength, when the beamsplitter is placed at 22.5 degrees.

The filters used are a set from the batch purchased from Optical Coating Laboratory Inc. (Ocli). These filters are used in the other UKIRT cryotats and in the infrared systems of other observatories such as the NASA Infrared Telescope Facility and the Anglo Australian Observatory. The nominal bandpasses and measured transmissions of these filters are (Ocli data) listed in Table 3.4.

J	1.10 - 1.40 μm	transmission = 67 %
H	1.50 - 1.80 μm	85
K	2.00 - 2.40 μm	75

Table 3.4 The nominal bandpasses and measured average transmissions for the filters (Ocli data).

The transmission profiles have been measured cold by Selby (1981) and the effective wavelength defined as

$$\lambda_{eff} = \frac{\int \lambda F(\lambda)\ d\lambda}{\int F(\lambda)\ d\lambda}$$

was also measured, where $F(\lambda)$ is proportional to the filter tansmission. These results shown below are close to the nominal figures.

	wavelengths		
	short	long	effective
J	1.089	1.370	1.228
H	1.490	1.782	1.643
K	1.969	2.388	2.182

Table 3.5 The measured filter bandpasses and effective wavelengths (Selby 1981).

Allen and Cragg (1983) found that the J filter was rather poorly defined and that the material used for other optical components and the water vapour in the atmosphere can shift the effective wavelength. However for sapphire and zinc selenide they did not find any significant wavelength shift and so the effective wavelengths for the three filters in this cryostat should be close to the filters in the other cryostats.

It is assumed that the reflectance off the beamsplitter does not vary significantly within the J and H pass bands. If this were the case colour equations would need to be applied to the data to allow comparisons with data measured on other systems. To verify that this is unnecessary, either the filter pass bands should be measured in the laboratory with a monochromator, or well defined red and blue standard stars should be observed. At the time of construction and use neither of these options was available. The effects are likely to be small, however, and are of little importance when studying colour variations.

3.2.5 Mark II Optical Design

Useful data were obtained with the Mark I cryostat but, as demonstrated in Section 3.6.1, the beam profiles were not flat enough to perform brightness photometry accurate enough for observations of cataclysmic binaries, for example. This was not the case for surface photometry of galaxies, since the extended nature of the source tended to smooth out the small peaks in the beam pofiles. Another requirement for the new system was for larger apertures to reach lower surface brightness regions of extended objects and to allow observations to be carried out in the presence of windshake.

Two 0.5 mm diameter detectors became available for use in mid 1982 and this offered the opportunity to upgrade the cryostat. Using the principle of conservation of optical throughput it is clear that the lens using these larger detectors could be operated at a slower focal ratio than with the 0.25 mm detectors. Since the transverse spherical aberration varies inversely as the square of the focal length of the lens, larger apertures and flatter profiles should be obtainable with such a system. Table 3.6 shows the calculated blur sizes for a lens of focal length 13.1 mm which will give an image of diameter 0.375 mm on the 0.5 mm detector.

Semi-aperture (mm)	Sky aperture (arcseconds)	Blur (mm)
0	0	0.002
1.0	3.2	0.013
2.0	6.4	0.041
3.0	9.6	0.094
3.5	11.2	0.133
4.0	12.8	0.180

Table 3.6 The linear blur due to third order spherical aberration for the plano-convex lens of focal length 13.1 mm with the object at infinity.

These calculations suggest that apertures up to 11 arcseconds on the sky should be attainable with no loss of light through spherical aberration. Calculations using the ray tracing program confirmed this and also showed that the manufacturing tolerances could be considerably relaxed.

The predicted profiles for a perfectly focussed and aligned system are shown in Fig. 3.4. The image may be almost 4 mm off-axis before light is lost. Defocussing by 0.2 mm, reduces this figure to about 3.3 mm (10.6 arcseconds on the sky) as the detector is moved away from the lens. If the detector is moved inside the focus, then the area covered in the focal plane is actually increased. Moving the detector away from the focus of the lens, however means that the image will move on the detector as the star image moves in the focal plane. This is not desirable as variations in detector responsitivity across its area will then affect the signal. Moving the detector off-axis by 0.10 mm produces the profile shown in Fig. 3.5b, which can be corrected by tilting the cryostat (Fig. 3.5c). The required manufacturing tolerances of the 0.5 mm system may then be summarised as shown below. These are for an 11 arcsecond beam rather than the 5 arcsecond beam of the 0.25 mm detector system.

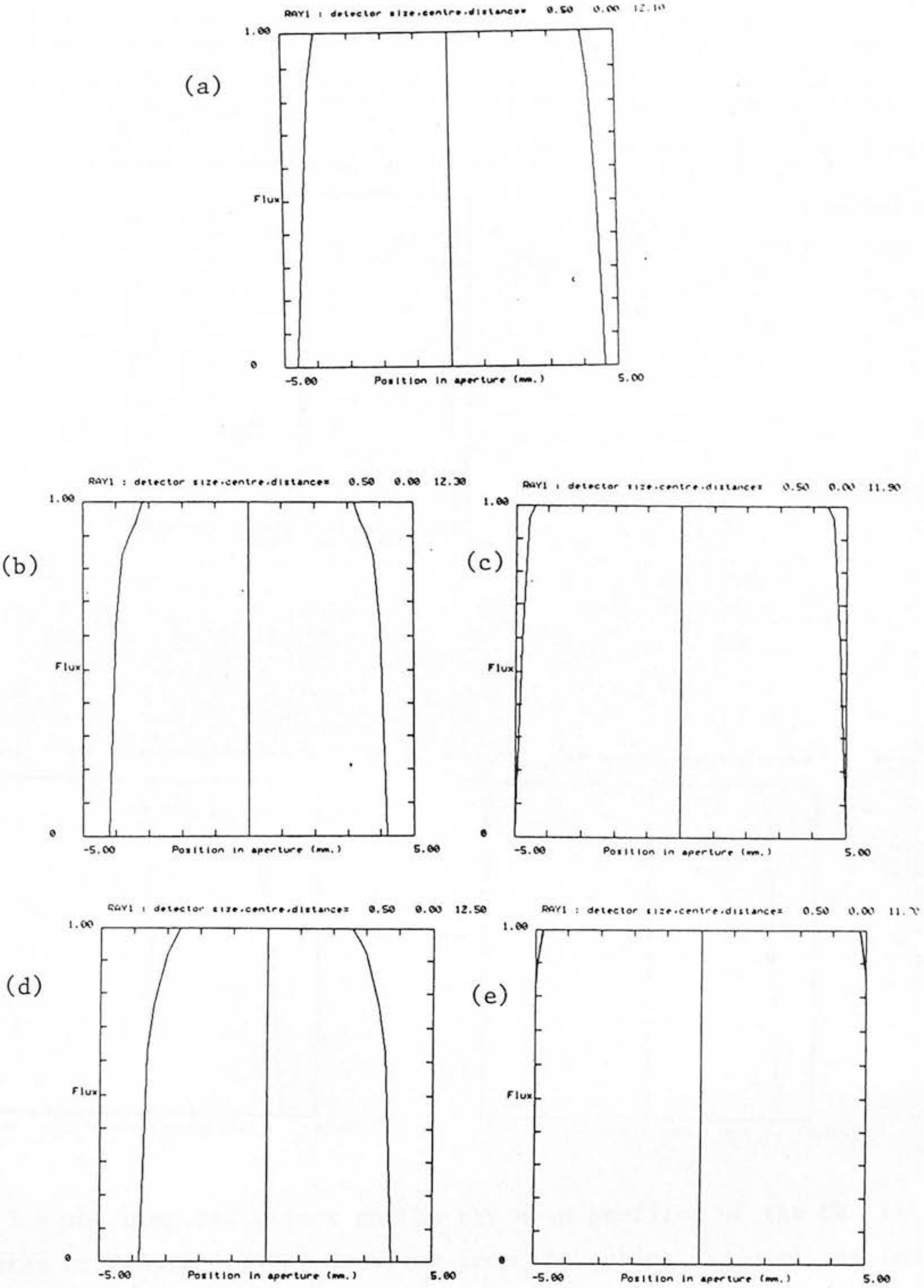


Fig. 3.4 The computed effect on the sky beam profiles of the Mk. II cryostat of detector/lens focussing errors. (a) is the correctly focussed profile and (b), (c), (d) and (e) represent focussing errors of .2, -.2, .4 and -.4 mm. respectively.

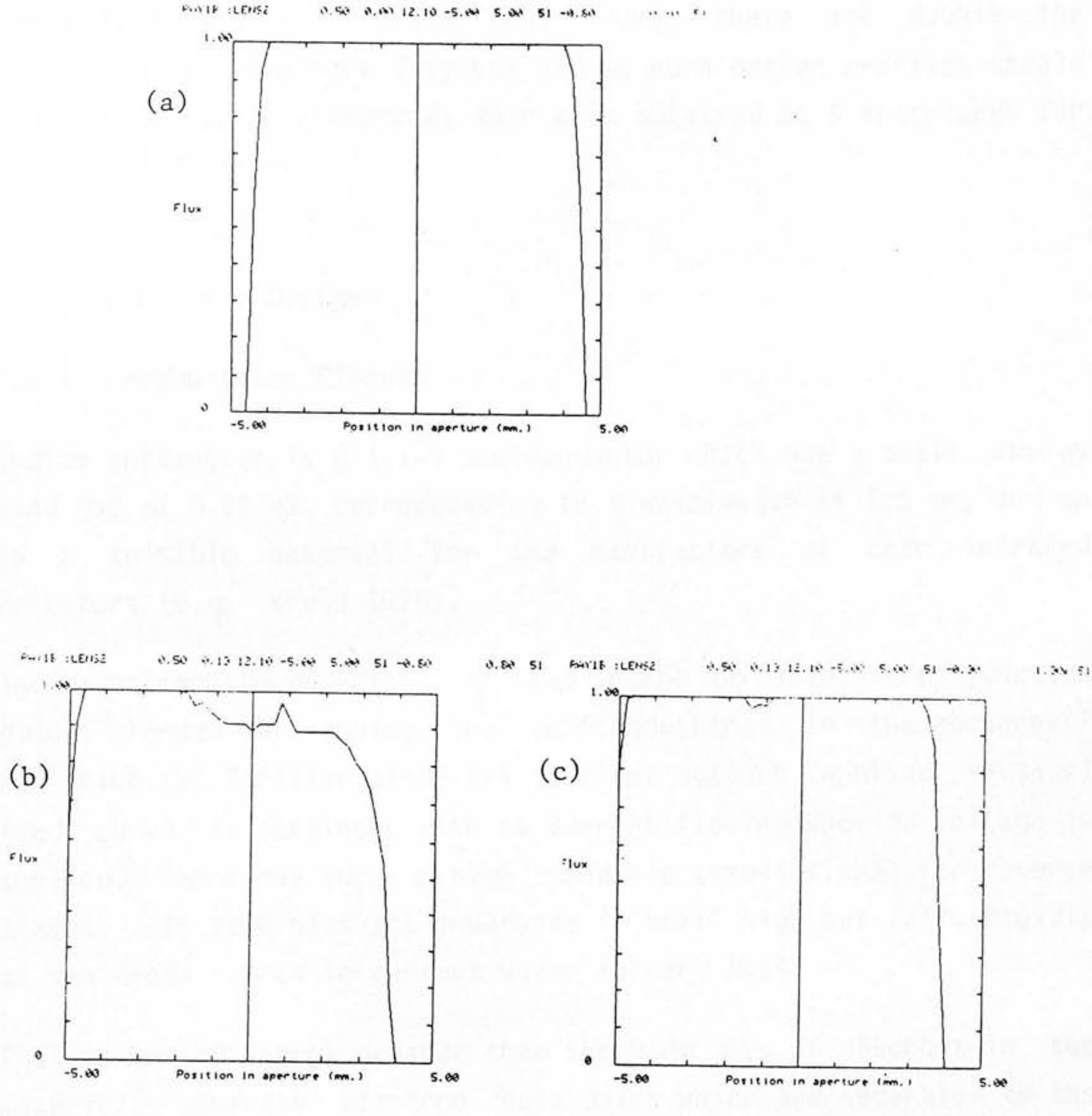


Fig. 3.5 The computed effect on the sky beam profiles of the Mk. II cryostat of displacing the detector from the optical axis of the lens. (a) is the on-axis profile and (b) represents the effect of displacing the detector by 0.125 mm. This can be corrected by tilting the cryostat as shown in (c).

detector focus	: 0.2 mm
detector transverse position	: 0.1 mm
colinearity of acceptance cones	: 0.3 degrees

Apart from the colinearity requirement, these are double the tolerances of the Mark I system and so much better profiles should be expected for 11 arcseconds than were obtained at 5 arcseconds for the Mark I cryostat.

3.3 Electonic Design

3.3.1 Preamplifier Circuit

Indium antimonide is a III-V semiconductor which has a small energy band gap of 0.22 eV, corresponding to a wavelength of 5.5 um, and so is a suitable material for the manufacture of near infrared detectors (e.g. Kruse 1970).

Indium antimonide detectors, as used in the infrared, are junction diodes formed by doping the semiconductor. In the absence of radiation the familiar diode I-V (current against applied voltage) load curve is obtained, with no current flowing when no voltage is applied. The curve shows a high impedance (small slope) for reverse biases. At zero bias the impedance is still high but falls rapidly as the diode starts to conduct under forward bias.

Photons having energy greater than the band gap, if absorbed in the material, generate electron hole pairs which are separated by the electric field in the junction created by the diffusion of carriers. The connection of an external circuit allows these carriers to flow generating an electrical current and the load curve is displaced along the I axis so that a current flows at zero bias.

The incident radiation may be determined either by measuring the voltage across the diode if the external circuit has resistance, or by measuring the current with zero voltage across the detector. The latter method is chosen because the current is linear with the number of detected photons and because minimum noise is obtained (by thermodynamic pinciples) when no energy sources are present other than random thermal motions.

To measure the current and ensure that the detector is zero biased a transimpedance amplifier with negative feedback is used in a virtual ground configuration so that the input node is held at ground potential. A circuit commonly employed is that described by Hall et al (1975) using two field effect transistors (JFETs) in a long tailed pair cofiguration followed by an operational amplifier.

As is shown below, the gain of this circuit (output voltage divided by input current) is given by R_f, the value of the feedback resistor. In order to make the Johnson noise from R_f negligible, its value must be large. A value of 4×10^{11} ohms was used; this was made comparable with but not higher than the detector resistance, since sources of noise other than Johnson noise become significant. Ceramic chip resistors are used because they operate at cryogenic temperatures and because the shunt capacitance, which limits the frequency response of the Hall circuit, is small. These resistors have poor linearity with voltage and so the sensitivity of the circuit can depend on the background and amplifier operating point. Also the frequency response is determined by poorly defined capacitances, including the shunt capacitance across the feedback resistor and the distributed capacitance to ground along it.

These disadvantages may be avoided by employing capacitive feedback, a version of which is described by Wyatt et al (1974). In this circuit, the gain is determined by the values of capacitors which are intinsically more linear than resistors, since their values depend only on the physical dimensions and separation of the conducting surfaces and on the dielectric constant of the intervening material. The values of resistors, on the other hand, depend on the density and mobility of current carriers, which in turn depend on the temperature and presence of electic fields in the material. The capacitor in the feedback circuit also reduces the effects of the smaller stray capicitances.

The gain, frequency response, linearity and noise performance of this circuit are analysed below and compared with the performance of the Hall circuit. The performance is investigated analytically and analysed using a general purpose circuit analysis progam, SPICE2 (Vladimirescu et al, 1981). The circuit used is shown in Fig. 3.6; this is equivalent to the Hall circuit if R_1 is made very small, R_2 is made very large and C_f becomes the stray shunt capacitance.

3.3.2 Gain and Frequency Response

The circuit may be simplified by assuming that it consists of a current source (the detector), an amplifier of infinite input impedance and voltage gain, and the feedback components R_f, C_f, R_1, R_2 and C_2.

Z_2 is the impedance of C_2 and R_2 i.e.

$$Z_2 = R_2 + 1/j\omega C_2$$

Z_f is the impedance of R_f and C_f i.e.

$$Z_f = \frac{R_f}{1 + j\omega R_f C_f}$$

The fed back voltage is

$$V_f = V_o Z_2/(Z_1 + Z_2)$$

,where V_o is the output voltage

The circuit is a virtual ground amplifier meaning that the input is held to ground potential, which follows from the fact that the output voltage is finite and the amplifier gain is very large. This means that the source sees a low input impedance although almost no current flows into the amplifier. (The input impedance is in fact Z_f/A_0 where A_0 is the open loop voltage gain). The source current from the detector, I_d, is therefore equal to the current flowing through the feedback impedance Z_f.

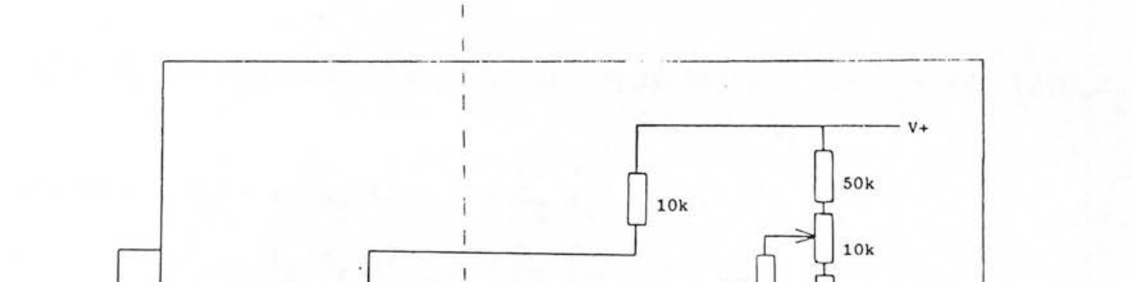


Fig. 3.6 The preamplifier circuit used for each channel. The circuit is based on that of Hall et al (1975) with the addition of capacitive feedback (Wyatt et al 1974). The components to the left of the dotted line are inside the cryostat at solid temperature and those to the right of the dotted line are at ambient temperature.

Therefore

$$I_d = \frac{V_f}{Z_f}$$

$$= V_o \frac{Z_2}{Z_1 + Z_2} Z_f$$

or

$$V_o = I_d Z_f \frac{Z_f + Z_2}{Z_2}$$

Assume $R_1 >> R_2$ and that Z_2 is capacitive i.e. $\omega << 1/(C_2R_2)$

Therefore

$$V_o = I_d Z_f (1 + j \omega C_2 R_1)$$

$$= \frac{I_d R_f (1 + j \omega C_2 R_1)}{1 + j \omega C_f R_f}$$

$$= I_d \left\{ \frac{1/R_f + \omega^2 C_2 C_f R_1}{1/R_f^2 + \omega^2 C_f^2} + j \omega (C_2 R_1/R_f - C_f) \right\}$$

Assume $\omega >> 1/(C_fR_f)$, and $R_fC_f = C_2R_1 = 1/\omega_0$

Therefore

$$V_o = I_d C_2 R_2/C_f$$

This expression shows that the transimpedance gain C_2R_1/C_f is independent of frequency and also independent of the value of the feedback resistor R_f. The gain of the uncompensated amplifier, and also the d.c. gain of the compensated amplifier, is R_f which is equal to C_2R_1/C_f by the assumption above. The frequency response of the amplifier is therefore flat from d.c. to the frequency at which Z_2 starts to become resistive i.e. when

$$0 < \omega < 1/C_2R_2$$

The easiest way to achieve the condition $C_fR_f = C_2R_1$ is to select R_f, C_f and C_2 and to trim R_2 to obtain a square wave output for a chopped input (or to fix R_2 and trim C_2). This also gives the value of R_f whilst cold which may be useful as a check of the temperature characteristics of the feedback resistor.

At high frequencies Z_2 becomes capacitive causing there to be very little feedback. The inclusion of R_2 helps ensure high frequency stability and also defines the upper frequency limit of the flat part of the gain. This may also be useful when the signal is detected using phase sensitive detection methods which do not reject the harmonics (which have signal to noise inferior to the fundamental).

Therefore, for frequencies $\omega > 1/C_2R_2$

$$V_o = I_dR_1/(j\omega C_fR_2)$$

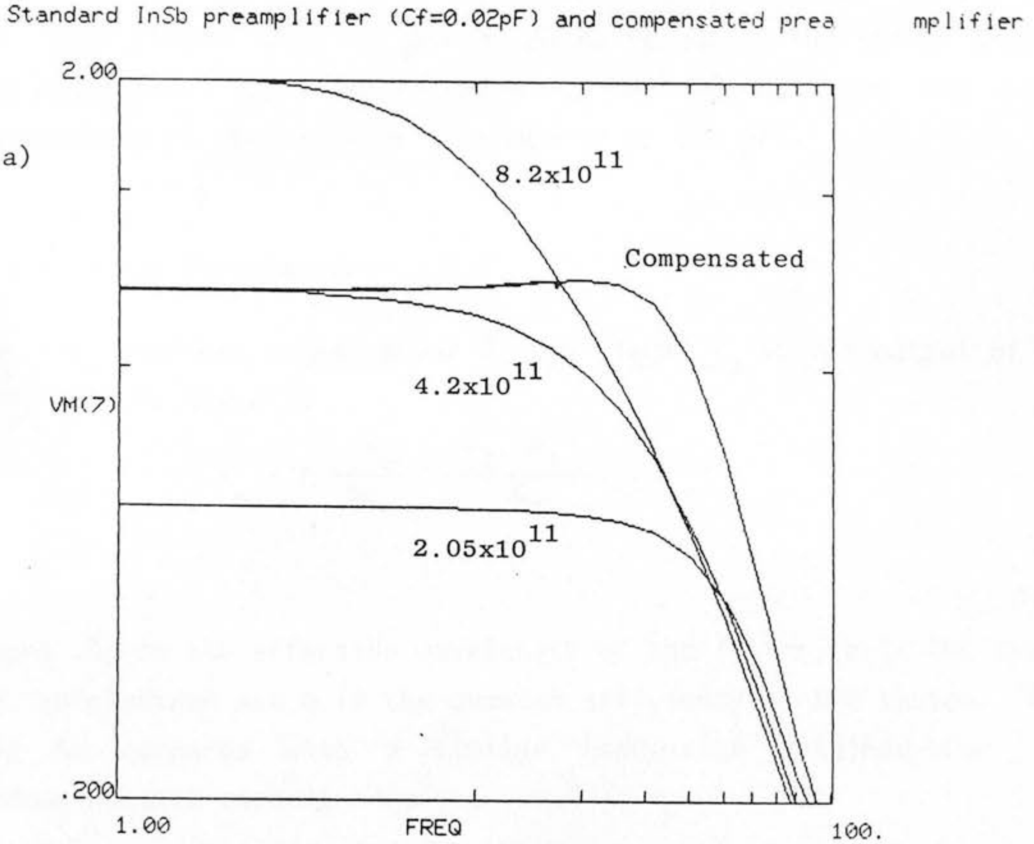
Assuming a large voltage gain in the operational amplifier, the uncompensated circuit may be shown to have gain described by

$$V_o = I_dR_f/(1 + j\omega C_fR_f)$$

where R_f now represents the stray capacitances associated with the feedback resistor (typically 0.05 pF for an Eltec Model 104 ceramic resistor). For a feedback resistance of 4.1×10^{11} ohms, this implies a frequency response which is 3 dB down at 8 Hz. The frequency respose and hence the gain at typical chopping frequencies are thus very dependent on both stray capaitances and the value of the non-linear feedback resisitor.

Both the compensated and uncompensated circuits have been fully analysed using the SPICE2 program. The JFET parameters were taken from the 2N 6484 data sheets and included typical values for the transconductance and cpacitances. The operational amplifier data was taken from the OP05 data sheets (National Semiconductors). The feedback resistor shunt capacitance was taken to be 0.05 pF and distributed capacitances along it and to ground were ignored. The latter was small since the resistor and associated wiring were not held close to the ground plane. The distibuted capacitances along the resistor have the effect mainly of altering the slope of the gain beyond the 3 dB point.

The calculated frequency response of the uncompensated cicuit is shown in Fig. 3.7 for different values of feedback resistor. Also shown is the frequency response of the compensated circuit for the same feedback resistor values. The gain of the compensated circuit is seen to be independent of the value of the feedback resistor at these frequencies and the frequency response is flatter. The fall off in response is determined by the selected values of R_2 and



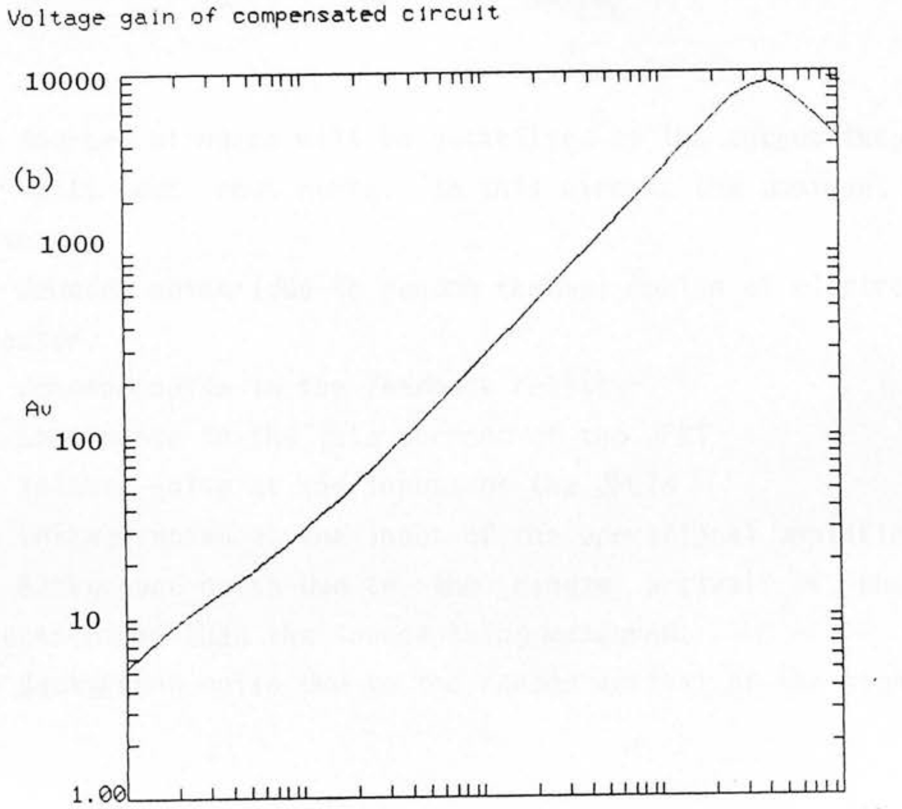


Fig. 3.7 (a) The frequency response of the compensated circuit and uncompensated circuits (the value of the feedback resistor is indicated) in the presence of stray shunt capacitance of 0.05 pF.

(b) The voltage gain of the compensated circuit. These responses were calculated using the SPICE2 program.

C_2. The slight rise in gain at 20 Hz is due to the finite gain of the operational amplifier (assumed to be 20,000) and the shunt capacitance of the detector (assumed to be 100 pF).

3.3.3 Noise Performance

For an incident signal power P, the signal V_s at the output of the amplifier is given by

$$V_s = P \frac{\lambda_e}{hc} \cdot \frac{C_2 \; R_1}{C_f} \cdot q$$

where λ_e is the effective wavelength of the filter, e is the charge on an electron and q is the quantum efficiency of the system. This may be compared with a similar expression obtained for the uncompenstated circuit

$$V_s = P \; q \frac{\lambda_e}{hc} \cdot \frac{R_f}{1 + j\omega C_f R_f}$$

The sources of noise will be quantified by the output they generate in volts per root Hertz. In this circuit the dominant sources of noise are

(a) Johnson noise (due to random thermal motion of electrons) in the detector.

(b) Johnson noise in the feedback resistor

(c) Shot noise in the gate current of the JFET

(d) Voltage noise at the inputs of the JFETs

(e) Voltage noise at the input of the operational amplifier

(f) Background noise due to the random arrival of photons from objects other than the source being measured.

(g) Background noise due to the random arrival of the signal photons

(a) Johnson Noise

The noise from the detector and feedback resistor ideally is Johnson noise which must be present from thermodynamic principles if the resistance is not infinite. Johnson noise is given by

$$V_N = \sqrt{4\ k\ T\ R} \qquad V/\sqrt{Hz}$$

where T is the temperature, R is the resistance and k is Boltzmann's constant.

Johnson noise in the detector can be minimised by cooling it. As well as reducing T in the equation, this also has the effect of increasing R by freezing out minority carriers in the semiconductor. The resistance of detectors increases exponentially as they are cooled below 77K (Wimmers and Smith 1982). The temperature is lowered by creating a low pressure above the liquid nitrogen, which becomes solid forcing the temperature down towards its triple point of 63K. Further increases in resistance are obtained by exposing the detector to short wavelength radiation (about 1μm through the J filter), which affects the surface states on the semiconductor. Using these techniques the resistance of the detectors could be increased to between 10^{11} and 10^{12} ohms. The Johnson noise of the detector then represents a design target for the preamplifier noise performance.

The value of the feedback resistor was such that Johnson noise from the feedback resistor, although higher than that from the detector, was not the dominant noise source of the system. For a value of 4.1×10^{11} ohms and a temperature of 60K, the Johnson noise is 37 $\mu V/\sqrt{Hz}$.

(b) Photon Noise

The photon noise is estimated by assuming that the arrival times of the photons follow a Poisson distribution. For a background power P the output noise voltage is then

$$V_N = e\ \frac{C_2\ R_1}{C_f} \sqrt{\frac{P\lambda}{hc}}$$

The photon noise of the source is fundamental and is not a limiting factor at these wavelengths. Stray background noise can be avoided by cooling the surroundings of the detector to liquid nitrogen temperature (77K), so that they emit a negligible number of photons at wavelengths shorter than 5.5 um, and ensuring that no extraneous radiation enters the cryostat. Inevitably the detector does see some background, due to wings on the acceptance cone and imperfect baffling. This background generates a signal, seen as a d.c. output from the amplifier V_{dc}, from which may be calculated the noise voltage V_N.

$$V_N = V_{dc}\sqrt{e\,\frac{C_2\,R_1}{C_f}}$$

$$= 2.56 \times 10^{-4}\ V_{dc} \qquad V/\sqrt{Hz}$$

For this to be less than the Johnson noise, V_{dc} must be less than 21 mV.

(c) Shot Noise

The current flowing into the gates of the JFETs generate a noise contribution due to the statistical fluctuations in the flow of electrons also assumed to obey a Poisson distribution. For a current I the noise voltage at the output is

$$V_N = \sqrt{Ie}\ \frac{C_2\,R_1}{C_f} \qquad V/\sqrt{Hz}$$

For this to be less than the Johnson noise the gate current must be less than 5.1×10^{-14} amps. This criterion is not met at room temperature (where it is 15 pA according to the manufacturer's specification), but since this current is an exponential function of temperature, at liquid nitrogen temperatures it can be expected to be negligible. At these temperatures other leakage conductances become important. Rieke et al (1981) have measued the leakage current of this type of JFET to be 2×10^{-15} A at 77K.

(d) Voltage Noise

Noise souces in the JFETs manifest themselves as noise voltages which are normally referred to the input. These sources include channel noise (Radeka, 1967), generation noise and donor trapping (Haslett, 1972) and are functions of frequency and temperature. Much work has been performed to select suitable JFETs for infrared systems both in ground and space sytems. Rieke et al (1981) have determined a noise voltage of 10 nV/$\sqrt{\text{Hz}}$ for the Intersil 2N 6484. However noise performance varies between individual devices and often shows large fluctuation with temperature (e.g. Haslett, 1972), making it difficult to predict actual performance.

Although the voltage gain of the circuit is small at d.c. (about unity if R_d is large), the capacitance associated with the detector (about 30 pF for the 0.25 mm detectors and 120 pF for the 0.5 mm detectors) causes the voltage gain to be large at a typical chop frequency of 10 Hz. If the detector is represented by a parallel combination of resistance, R_d, and capacitance, C_d, and the input voltage noise by V_N then (see Fig. 3.6)

$$V_i = V_o/A = (V_N + V_o) \cdot \frac{Z_2}{Z_1 + Z_2} \cdot \frac{Z_d}{Z_d + Z_f}$$

,assuming that R_1 and $Z_2 \ll Z_f$ and Z_d. Therefore,

$$V_o = V_N \frac{(R_1 + Z_2) \cdot (Z_d + Z_f)}{Z_2 \, Z_d - (R_1 + Z_2)\,(Z_d + Z_f)/A}$$

Since A is large,

$$V_o = V_N \left(1 + \frac{R_1}{Z_2}\right)\left(1 + \frac{Z_f}{Z_d}\right) \qquad \text{(Eq. 3.1)}$$

At typical chopping frequencies Z_2, Z_f and Z_d are capacitive. Therefore

$$V_o = V_N \, (1 + j\omega C_2 \, R_1) \, (1 + C_f/C_d)$$

$$\simeq V_N \, j\omega R_1 \, C_d \, C_2/C_f$$

$$= V_N \, j\omega R_f \, C_f \qquad \text{(if } R_f \, C_f = R_1 \, C_2\text{)}$$

Without compensation this would be

$$V_o = V_N \left(1 + \frac{C_d}{C_f} \right)$$

Although it appears that the noise from the compensated cicuit is greater by a factor of $R_f C_d C_f/(C_d+C_f)$ this is also the factor by which the gain of the transimpedance amplifier is increased (equations 3.1 and 3.2) since C_d is large compared to C_f, and so the frequency compensation does not increase the effect of voltage noise. Using a value for C_d of 120 pF the voltage gain is about 3000 and so the voltage noise of the JFETs must be less than about 12 nV/$\sqrt{\text{Hz}}$ if this source of noise is to be less important than the Johnson noise.

Such noise levels are achievable but good performance is likely to be dependent on the operating temperature and ideally the JFETs should be temperature controlled using a method such as that described by Storey (1981). Problems were encountered with this method and the JFETs were operated without temperature control.

Voltage noise from the operational amplifier may also be significant if the voltage gain of the first stage is small. The JFET is operated in source follower mode and so the gain will be approximately unity if $g_{fs}.R$ is larger than unity, where g_{fs} is the transconductance of the JFET and R is the load resistance in the source of the JFET. The nominal g_{fs} is 2000 minimum, which should rise at lower temperatures (Sesnic, 1972), and so the gain of this stage should be approximately unity. The source follower configuration prevents the Miller effect becoming important (Watson, 1970) and so the input impedance of the JFET should still be large at 10 Hz. The quoted input noise of the operational amplifier is 10 nV/$\sqrt{\text{Hz}}$ and is not subject to low temperature effects being external to the cryostat. The voltage noise from this source should therefore be no larger than that from the JFETs although a JFET buffer stage with voltage gain would make this more certain.

3.4 Mechanical Construction and Alignment

The components were mounted in a standard HD 3 cryostat manufactured by Infrared Laboratories Inc. . This cryostat is compact and light, so minimising the risk of flexure. It has two canisters for storing cryogens and is often used in applications using liquid helium. The inner can is thermally connected to a flat cold face and the outer can is connected to a radiation shield surrounding the cold face and inner can.

The cooled components (detectors, feedback resistors and capacitors, JFETs, lenses, filters, beamsplitter and aperture wheel) were all mounted on a baseplate which was screwed firmly to the cold face. This baseplate was removable to facilitate alignment. Most of the cooled metal components were manufactured from oxygen free copper which has good thermal conductivity at cryogenic temperatures.

Accurate manufacture of the detector and lens assemblies were essential to the success of the instrument, since no facilities for adjustments were included in the design. The critical dimensions were the distance between the detector and the back of the lens, and the distance the detector was away from the optical axis of the lens. It was also important to ensure that the optical axis of the lens was parallel to the cold face.

The detector/lens assemblies, shown in Fig. 3.1 for the Mark II design, were manufactured from a block of oxygen free copper. The front and rear faces, the surface against which the lens was held, and the base were machined flat and perpendicular. The cavity in which the lens rested was made a tight fit to the edge of the lens so that the optical axis could not shift laterally. This cavity was made concentric to the circular aperture at the rear of the lens assembly, in which the detector assembly (also circular) was a tight fit. To allow alignment with a laser, the top of the lens assembly was made removable.

The detector was positioned in the centre of the circular detector assembly using a travelling microscope to achieve a positional accuracy of about 0.05 mm. To ensure accurate positioning of the detector behind the lens, the height of the detector surface above the mating surface of the detector assembly was measured using the travelling microscope. The distance from the rear of the lens to this mating surface was also measured and the sum of these gave the back focal position. To adjust this to the design value, the rear suface of the lens assembly was rubbed down using fine grade abrasive paper on an flat table. This procedure was believed accurate to 0.1 mm. The detector/lens assembly therefore formed a unit with an acceptance cone parallel to the base and perpendicular to the front face.

The beamsplitter could be rotated in two axes and was held by set screws. These could be adjusted to move the beamsplitter through known angles. The J and H filters were mounted in a wheel placed between the beamsplitter and the reflected channel detector/lens assembly. These were moved from outside the cryostat by shafts made from G10, a low thermal conductivity glass fibre composite material, to allow operation with liquid helium. Liquid nitrogen was used throughout and so the filter shaft was later replaced with a stainless steel tubular shaft to afford greater stiffness.

To ensure low noise operation it was important to eliminate thermal background and stray light. The optical paths were surrounded by baffling so that the detectors could not see the radiation shield and could only see out the window through the focal plane aperture. As a further precaution, the filter and aperture shafts, which could have a large temperature gradient, were covered in aluminised Mylar to reduce their thermal emissivity. These shafts also thermally connected to the radiation shield by copper braiding so that as much temperature drop as possible took place between the cryostat outer wall and the radiation shield.

The optical components were positioned on the base plate using a laser for alignment. The laser, repesenting the optical axis of the telescope, was directed through a small focal plane aperture in such a way as to be parallel to the baseplate and approximately perpendicular to the aperture wheel. The K channel was aligned initially with the beamsplitter removed. Using an eyepiece, the position of the focussed laser light on the detector was examined. The detector/lens asembly was positioned so that this was central on the detector. Stray reflections in the lens also aided in this procedure.

Having positioned the K channel detector/lens assembly, the beamsplitter was inserted without moving the laser or baseplate. The only effect this had on the K channel was to move the beam laterally by approximately 0.7 mm, which had insignificant effect on the alignment. The detector/lens assembly of the reflected channel was then position in a similar manner. The beamsplitter angle was adjusted so that the stray reflections from the lens on the rear of the aperture wheel were concentric with the aperture. This ensured the beam was normal and central on the lens. The final positioning at this stage was made by moving the detector/lens assembly.

Optical alignment tests were carried out on the cooled cryostat as described in Section 3.5.1. The difference in direction between the acceptance cones so determined was used to calculate an adjustment in beamsplitter angle. The cryostat was then warmed up, the adjustment made and then the procedure was repeated. Errors in focus or positioning of detectors with respect to lenses necessitated the whole procedure being repeated. Many iterations were required with the Mark I cryostat, but the initial alignment was sufficiently accurate for the Mark II cryostat reflecting the less critical tolerances of the design.

3.5 Laboatory Performance

3.5.1 Optical Tests

Because of the close tolerances required for the colinearity of the acceptance cones of the two detectors, extensive iterative tests of the optical alignment were carried out in the laboratory. The tests will be described in detail for the Mark I system, for which they were more critical. The test results for the Mark II cryostat will also be described. The tests were designed to determine

(a) The alignment of the individual detector/lens assemblies.

(b) The degree of coverage by the detector of different sized focal plane apertures.

(c) The alignment of the two detector/lens assemblies with respect to each other.

Of these (c) was most important since it allowed compensation of the inevitable manufacturing and construction errors in the individual detector/lens assemblies and was necessary to ensure good correspondence between the two beams necessary for accurate colour measurements.

The alignment of the detectors with respect to their respective lenses was determined by the contruction and was not easily adjustable. However the detector/lens assemblies could be moved on the cold face and the beamsplitter could be rotated and tilted by predetermined quantities. This allowed correction of small errors in colinearity between the two acceptance cones. This method was not effective for large errors since rotating or tilting the beamsplitter causes the centre of the focal plane aperture to move with respect to the acceptance cone of the reflected channel. A movement of one degree by the beamsplitter would cause the axis of the accpetance cone to move in the focal plane by 1.4 mm, a significant proportion of the 3.6 mm region covered by the detector (Section 3.2.5). Large movements of the beamsplitter would therefore cause the spherical aberration to degrade the profile of the deflected channel, necessitating a movement of the detector/lens assembly. The adjustment was therefore an iterative process.

The testing technique used was to scan a small chopped isotropic source across the acceptance cone of the cryostat and measure the signal in each of the two channels. The focal plane aperture was therefore evenly illuminated with parallel light from a distant source. To receive all the light from the secondary mirror there had to be a region 1.6 degrees wide, common to both channels, in which all the light incident on the focal plane aperture fell on both detectors. This became less easy to achieve as the focal plane aperture was enlarged, due to spherical aberation.

The expected profiles were computed using a version of the ray tracing program described in Section 3.2.2. The program calculated the proportion of the rays emanating from the focal plane aperture which fell on the detector for different angles to the focal plane. Fig. 3.8 shows the profiles that would be achieved for different apertures assuming that the alignment was perfect. As the aperture is increased the flat region becomes narrower and so light starts to be lost from the outer regions of the secondary mirror. In agreement with the calculations of Section 3.2.2, when the aperture becomes larger than about 3.8 mm the acceptance cone starts to become narrower than f/35 and light is lost.

The effect of offsetting the detector 0.0625 mm (one quarter the detector diameter) from the optical axis of the lens is shown in Fig. 3.9, for a 4 mm focal plane aperture. The profile becomes progressively more degraded as the aperture is increased in size. Also shown is the effect of moving the centre of the aperture 1 mm off the optical axis of the lens. These effects appear very similar and were observed during testing. Acceptable profiles can still be obtained if the centre of the aperture, the optical centre of the lens and the centre of the detector are colinear with each other even if they are not colinear with the optical axis of the lens. Fig 3.9(d) shows the calculated profile when these are displaced 0.5 degrees from the axis of the lens.

The colinearity of the acceptance cones was measured by determining the centre of each beam, as defined by the half power points. The angular error in each axis was then calculated, from which the beamsplitter adjustment was determined. The cryostat was then opened up and the beamsplitter was adjusted before cooling the cryostat down again and retesting. In practice, with the Mark I cryostat, many such stages were found to be necessary, with some adjustments to the detector focus and positioning also being

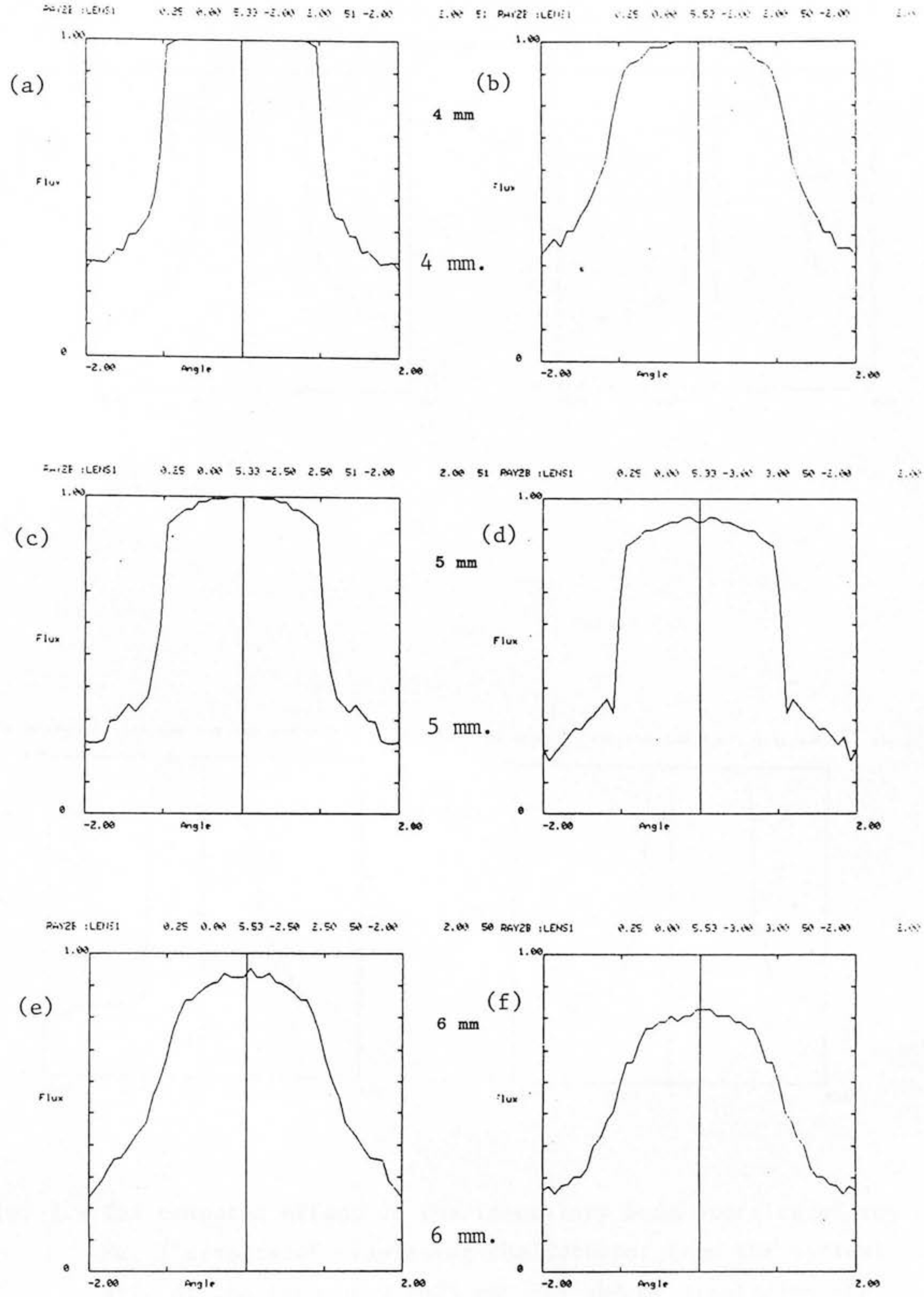


Fig. 3.8 The computed effect on the laboratory beam profiles of increasing the aperure size from 4 mm. (top) to 6 mm. (bottom). The left hand profiles are for a properly focussed system and those on the right are for a detector/lens focus error of 0.2 mm. Each figure shows the collecting efficiency as an isotropic source is moved through the acceptance cone between angles -2 and +2 degrees. The profiles are for the Mk. I cryostat.

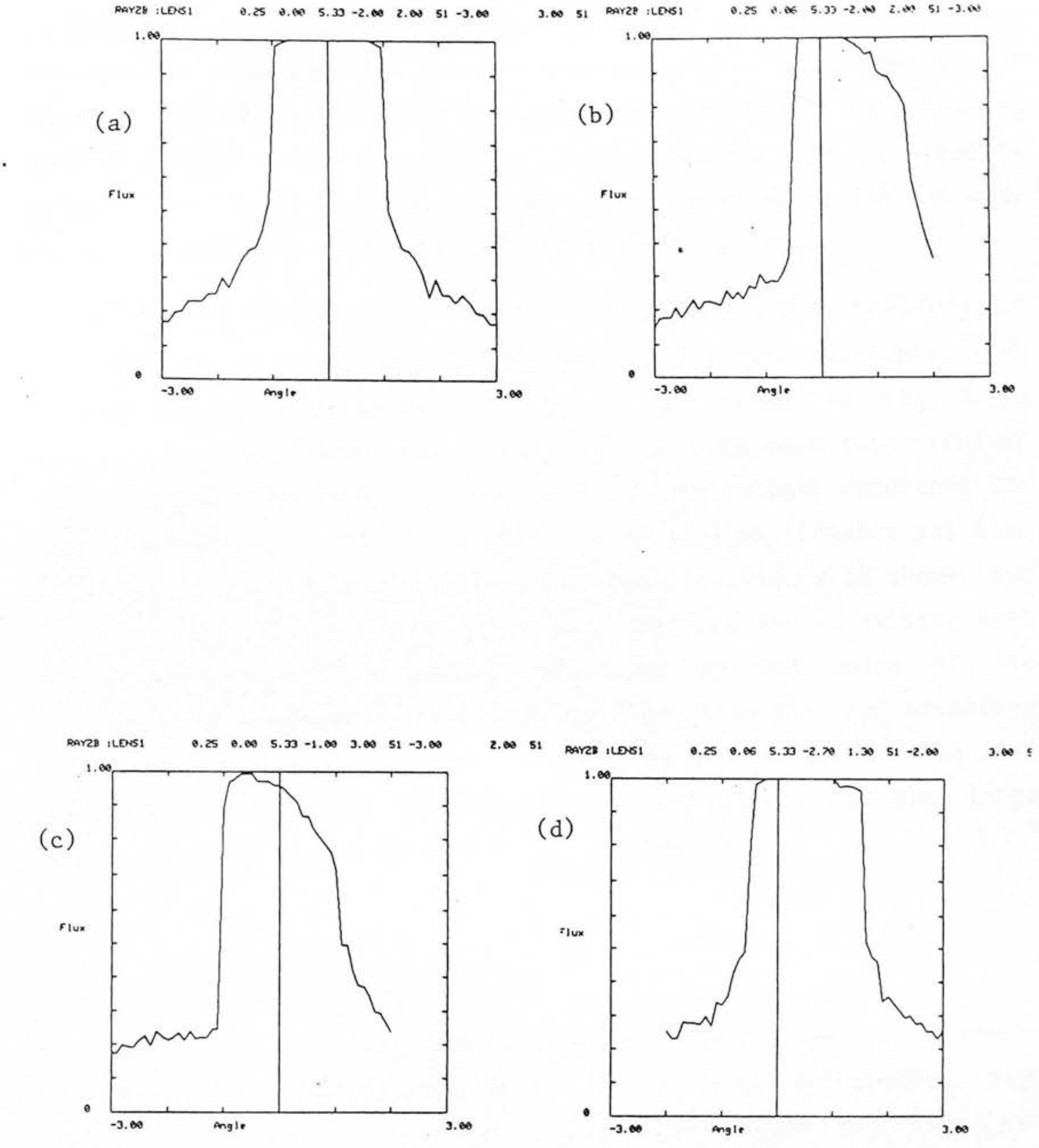


Fig. 3.9 The computed effect on the laboratory beam profiles of the Mk. I cryostatof displacing the detector from the optical axis of the lens by 0.0625 mm. (b) and of displacing the focal plane aperture by 1 mm. (c). (d) shows the effect of displacing the aperture by 0.7 mm. to compensate for the detector displacement of 0.0625 mm. A 4 mm. aperture is assumed.

required. Ideal profiles at the maximum aperture were not achieved, as seen from the examples shown in Fig. 3.10a, but the alignment was adequate to allow accurate colour and brightness measurements of galaxies to be performed. Because of the difficulty in obtaining good individual profiles, care was taken to ensure good colinearity between the two acceptance cones. The overlap region was wide enough to accept an f/30 beam between the -10% points.

The difficulty in obtaining flat profiles prompted the building of the Mark II cryostat when 0.5 mm detectors became available. The results of similar tests on this cryostat are shown in Fig. 3.11a and it is seen that this cryostat is much more successful at collecting all the radiation from an f/35 beam in both detectors to large apertures. This improvement in optical performance was also demonstrated at the telescope (Section 3.6.1). Fig. 3.11 shows the profiles obtained through the 7 mm. aperture and it is seen that the internal cryostat alignment, as shown by the shape of the profiles, is adequate to collect the light from the f/35 secondary mirror at the design apertures. The region of overlap of the two acceptance cones, about f/30 between the -10% points, was also large enough to accept the f/35 beam from the telescope.

3.5.2 Electrical Tests

Tests were performed on the system to check that the circuit performed as designed and that the noise was acceptable. High detector impedance is essential to good performance and this was checked by blanking the detectors off from external radiation, applying a small voltage bias to the detector and measuring the change in current. This could be done conveniently by applying a few millivolts to the gate of the second JFET. This change in voltage, ΔV_g, then appears across the detector and the change in the output voltage, ΔV_g, can be measured. The resistance of the detector is then

$$R_d = \frac{\Delta V_o}{\Delta V_o - \Delta V_g} . R_f$$

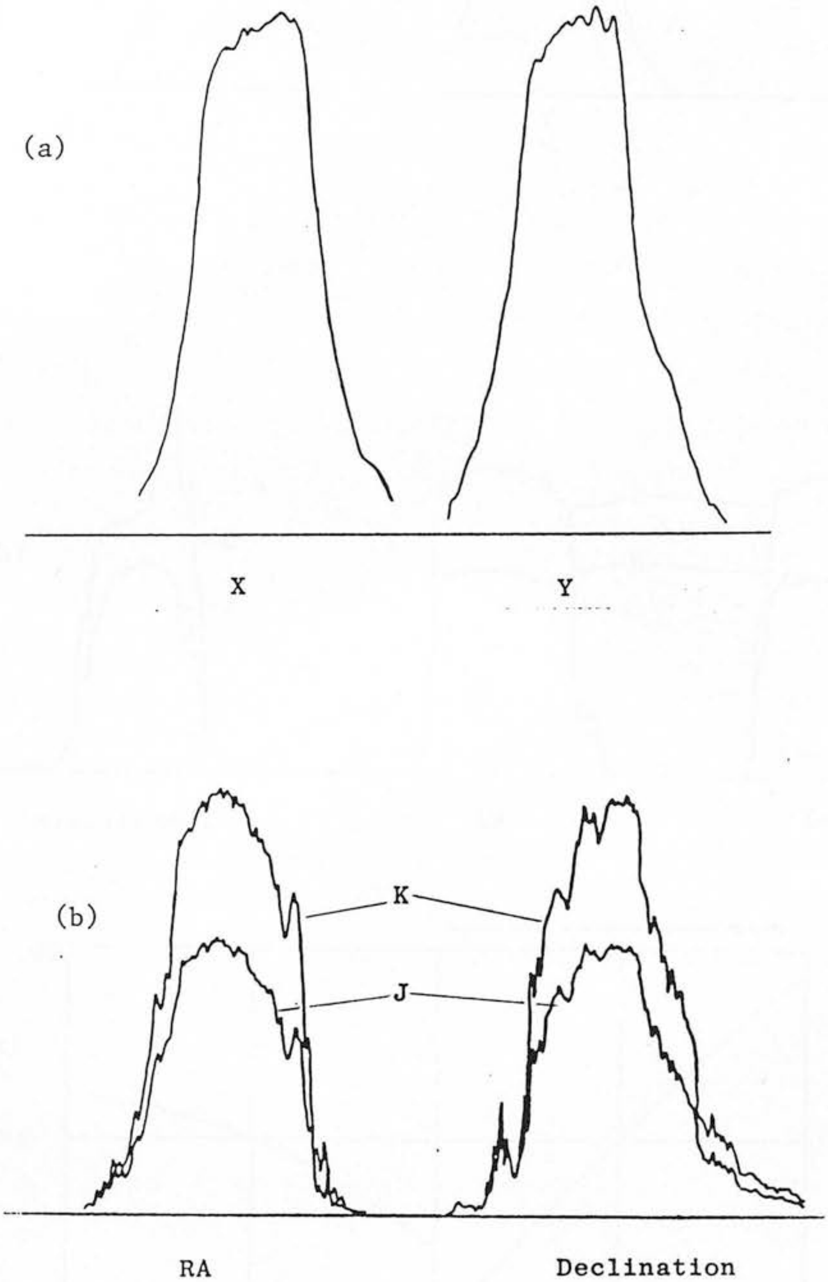


Fig. 3.10 The beam profiles obtained with the Mk. I cryostat. (a) shows the laboratory profiles as an isotropic source is scanned across the beam in X and Y. (b) shows the sky beam profiles as a star is scanned across the 4 mm. aperture.

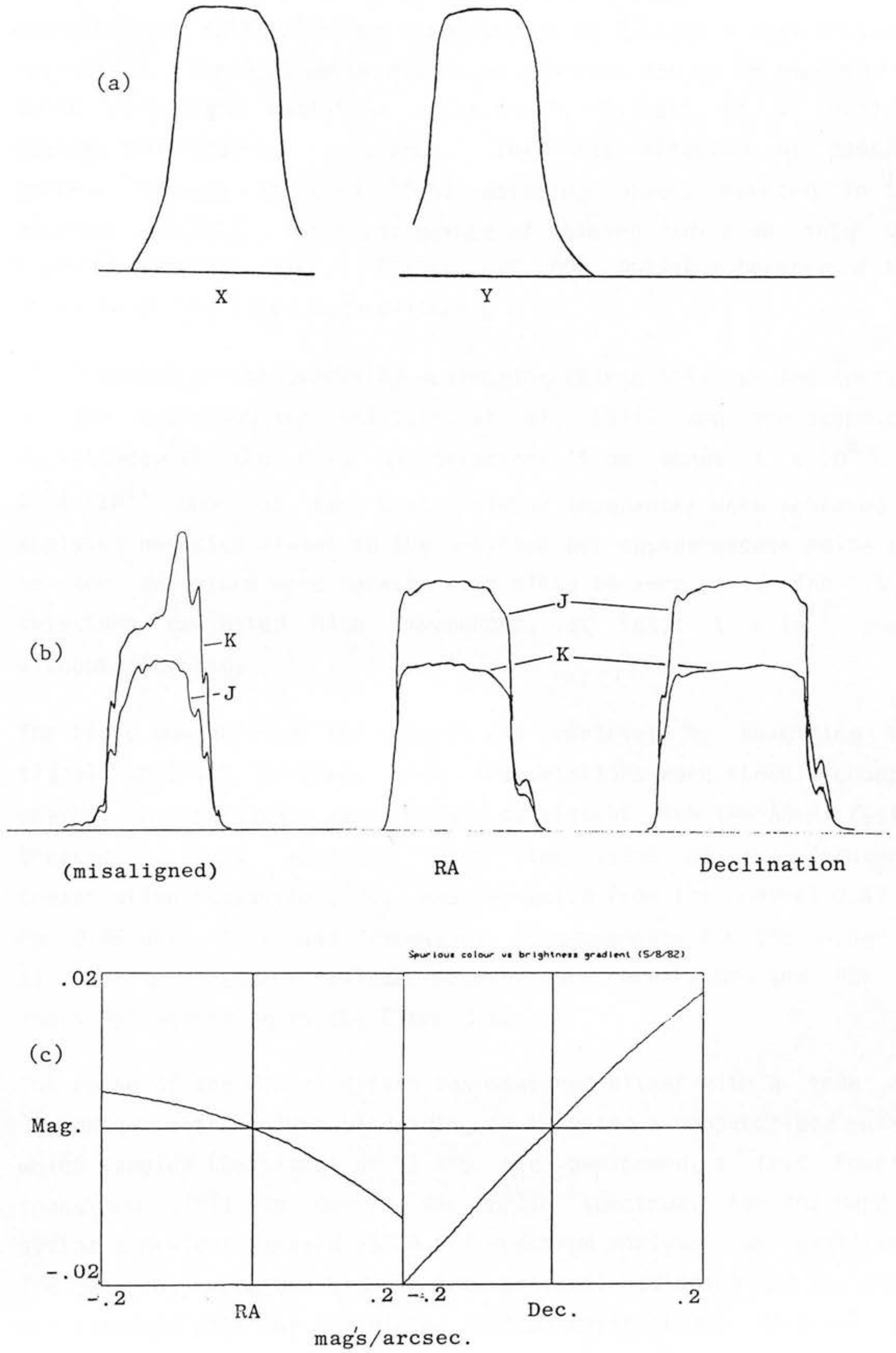


Fig. 3.11 The beam profiles obtained with the Mk. II cryostat. (a) shows the laboratory profiles and (b) the sky beam-profiles of the 7 mm. aperture. Also shown is the profile obtained before the cryostat was correctly tilted. (c) shows the small spurious colour that would be measured in the presence of large brightness gradients.

Both the 0.25 mm SBRC detectors and the 0.5 mm CE detectors were operated at solid nitrogen temperatures to achieve a high detector resistance. The 0.25 mm detectors in addition had to be exposed to short wavelength radiation ("flashing", Gillett et al 1977) to achieve the required resistance. This was effected by passing current through infrared light emitting diodes mounted in the detector assembly. The usual method of shining radiation into the cryostat through the J filter was not possible because of the presence of the fixed beamsplitter.

The flashing process works by destroying charge sites on the surface of the semiconductor (Gillett et al, 1977) and increased the resistance of the 0.25 mm detectors from about 1 x 10^9 to 1 x 10^{11} ohms at zero bias. Higher impedances were achieved by applying negative biases to the detector but caused excess noise and so the detectors were oerated with close to zero bias. The 0.5 mm detectors exhibited high impedances, at least 1 x 10^{12} ohms, without flashing.

The basic operation of the circuit was verified by examining the signal waveform produced when the detectors were shown a chopped source. A good square wave output, consistent with the 30 Hz design breakpoint, was achieved when the value of the frequency compensation capacitor, C_2, was increased from the nominal 0.47 μF to 0.68 μF. This was necessary to compensate for the change in value of the feedback resistor between room temperature and 60K of about 35% according to the Eltec data.

The noise of the Mark I system was measured either with a true rms voltmeter with a 10Hz bandwidth, or by using a computerised method which sampled the signal at 1 kHz and performed a fast Fourier transform (FFT) to derive the noise spectrum. For the Mark II system a Hewlett Packard 3582A FFT spectrum analyser was available. The detectors were blanked off from external radiation and the noise was compared with the calculated Johnson noise of the detector and feedback resistor.

A typical noise spectrum (from the Mark I system and the computerised spectrum analyser) is shown in Fig. 3.12. The average signal level was subtracted from the data before the FFT was calculated and so the spectrum falls to zero at d.c. The noise spectrum is relatively flat as would be generated by Johnson noise in the detector or background noise. Voltage noise is multiplied by the voltage gain of the amplifier which rises at 6 dB/octave from low frequencies and so seen at the output it would also generate a fairly flat spectrum due to its 1/f nature. The measured noise is about 50 percent greater than the Johnson noise due to the measured detector impedance of typically 1 x 10^{11} ohms at zero bias (80 $\mu V/\sqrt{Hz}$), but was no worse than other systems.

The noise performance of the Mark II system might be expected to be better because of the higher impedance of the detectors, which reached impedances high compared to the value of the feedback resistor (4 x 10^{11} ohms) at solid nitrogen temperaure without flashing. However the detector area was four times larger and so the shunt capacitiance would also be expected to be four times larger. This was confirmed by Cincinnati Electronics who estimated the capacitance to be 80 pF and probably larger. The voltage gain of the amplifier is proportional to this capacitance (Section 3.3.3) and so the noise performance was not as good as that of the Mark I system.

The noise was measured by using the HP 5582A spectrum analyser and was found to be similar in the two channels. The noise spectrum was not flat and the noise was found to increase with frequency as shown below.

Frequency (Hertz)	Noise ($\mu V/\sqrt{Hz}$)
1	60
2	71
5	100
10	141

Table 3.7 . The measured noise of the detectors of the Mark II cryostat at solid nitrogen temperature with the aperture blanked off.

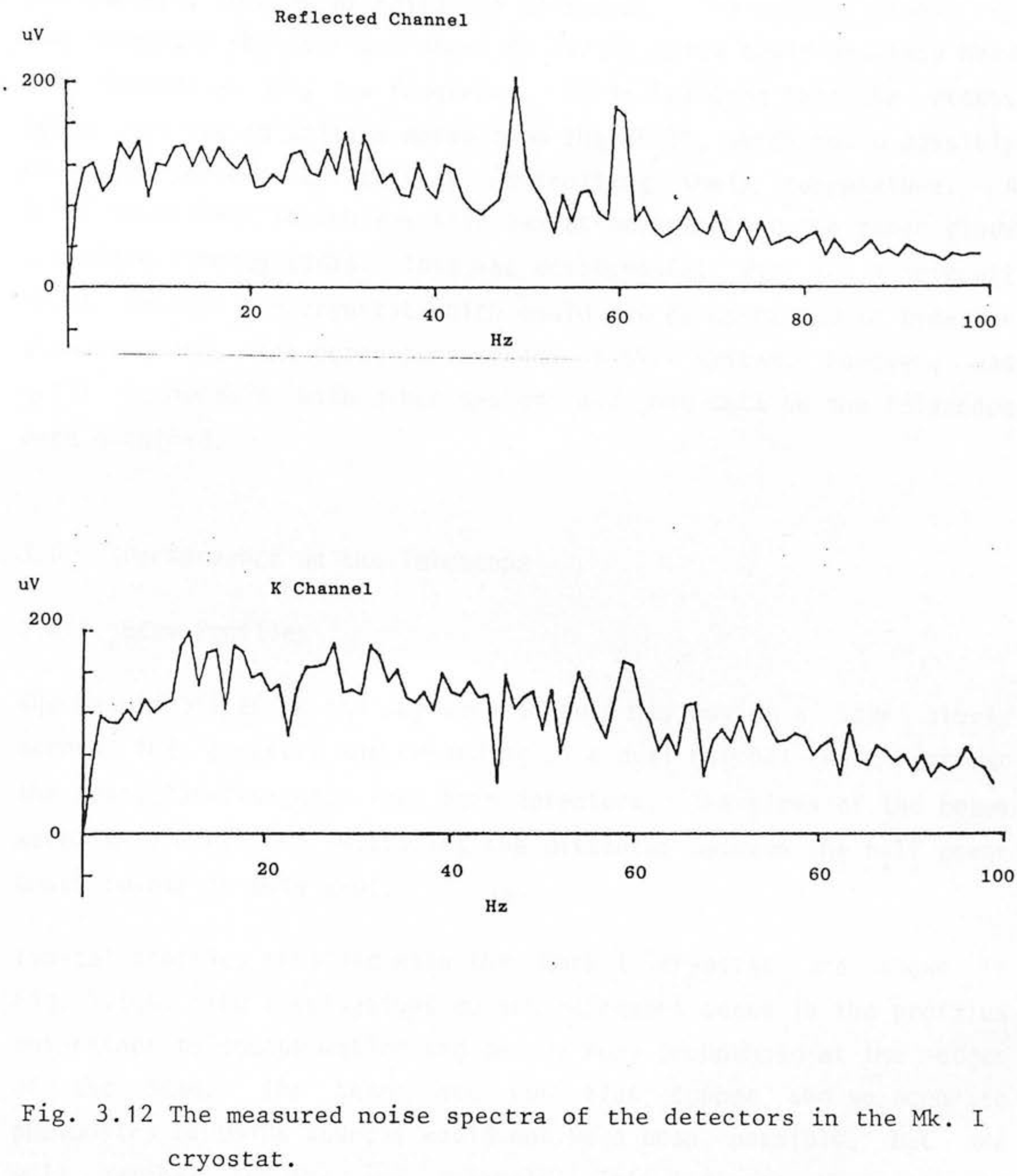


Fig. 3.12 The measured noise spectra of the detectors in the Mk. I cryostat.

These noise levels are in excess of those measured for the Mark I cryostat and the different shape of the noise spectrum shows that the dominant sources of noise are different. The Johnson noise of the feedback resistor was about 40 $\mu V/\sqrt{Hz}$ which could possibly have been reached at very low frequency. It is believed that the excess noise was due to voltage noise from the JFETs, which could possibly have been lessened by actively controlling their temperature. A brief experiment to achieve this was attempted using the zener diode technique (Storey 1981). This was unsuccessful due to a circuit fault inside the cryostat which could not be corrected in time for the observing. The noise performance of the system, however, was still comparable with other systems and good data on the telescope were obtained.

3.6 Performance on the Telescope

3.6.1 Beam Profiles

The beam profiles on the sky were measured by moving a star slowly across the aperture and recording on a dual channel chart recorder the demodulated signals from both detectors. The sizes of the beams were determined by measuring the distances between the half power power points in both axes.

Typical profiles obtained with the Mark I cryostat are shown in Fig. 3.10b. The fluctuations do not represent peaks in the profiles but rather telescope motion and become more pronounced at the edges of the beam. The beams are not flat topped and so accurate photometry of point sources would not have been possible, but are well matched in the two channels. This indicated that accurate colours could be obtained and that surface photometry of extended objects could be performed, although brightnesses in magnitudes per square arcsecond could not be assessed. For studying brightness and colour variations within an object this is not important.

The Mark II cryostat produced much flatter profiles in the laboratory and this is reflected in the flat beam profiles measured on the sky. The beams were also well matched and no difference in the position of their half power points was detected. Typical profiles are shown in Fig. 3.11b, which also shows a profile measured with the cryostat not accuratley pointing at the secondary mirror. This is very similar to the computed profile of Fig. 3.3(d) indicating that the optics are not perfectly aligned inside the cryostat but this is well within the tolerances as demonstrated by the profiles of the cryostat properly pointed at the secondary. The fluctuations at the edge of the beam are caused by small telescope motions and do not represent real features in the profiles.

Using the beam profiles measured on 5 August 1982, the spurious colour measured in the presence of large brightess gradients has been calculated. Fig. 3.11c shows the result of convolving these profiles with a linear brightness gradient and taking the difference in magnitudes between the two channels. For the large brightness gradient of 0.2 mag's/arcsec, the spurious colour is not more than 0.02 magnitudes compared to about 0.1 magnitudes which is likely to be obtained by consecutive measurements (Section 3.1).

The beams were measured to have widths, between the half power points, of 10.1 ±0.3 arcseconds in both axes. The beams were well matched in the two channels. Differences in the response across the profiles did not amount to more than about 2 percent of the beam height. The profiles were flat to 5 percent over most of the area of the beam, the precise extent being difficult to determine because of the effect of telescope motions at the edge of the beam. Because of the quality of the profiles, surface brightnesses in magnitudes per square arcsecond can be estimated from the data. To derive these 4.7 ±0.2 magnitudes are added to the total measured magnitude.

When performing photometry at positions near bright objects such as galactic nuclei, it is important that the system should not exhibit side lobes or tails on the beam profiles. The design of the system is such as to make this unlikely. Apart from the window, the first component in the cryostat is the focal plane stop and subsequent optical components are well baffled. Stray reflections from lenses or filters cannot therefore cause tails or side lobes.

In order to verify that such effects were not present, observations were performed at positions about 20 arcseconds away from a star of K magnitude about 3. Integrations of 40 seconds duration did not produce statistically significant detections, implying an upper limit of K=16. No side lobes or tails were therefore detected, which would be significant in the steepest brightness gradient observed (0.2 mag's/arcsec in NGC 5907). Tails were produced however due to the finite transition time of the chopper (Section 4.3).

3.6.2 Sensitivity

The sensitivity of the system was determined by measuring the signal from a standard star and the noise from a measurement on the sky. The standard data acquisition program PHOTOM (Stewart 1983b) was used for this and an instrumental zero point (Section 4.5.1) in magnitudes, corrected for the amplifier range and integration time, was derived. For the noise measurements the integration time was set to one second and the time constant of the phase sensitive detector (lock-in amplifier) was set to 125 ms. A series of integrations (typically 40) was measured from which the mean and standard deviation were calculated. The mean represents the telescope and sky offset (small at these wavelengths) and the zero levels of the lock-in amplifier output and voltage to frequency converter. The standard deviation represents the noise in one second (in counts /mV /ms, where one count represented approximately 2 uV) from which a "noise magnitude" was calculated and printed.

noise magnitude = -2.5 log (2x (st. dev.))

The factor of 2 is included because the zero point represents a "right" minus "left" value, i.e. twice the signal. The noise at the output of the preamplifier in $\mu V/\sqrt{Hz}$ was also calculated and printed.

The senstitivity of the system, expressed as a "one second one sigma" magnitude was then the difference between the zero point and the noise magnitude. Typical values for the two system are shown below.

	Noise ($\mu V/\sqrt{Hz}$)		Noise Magnitude		Zero point Magnitude			1 sigma 1 second Magnitude		
	J,H	K	J,H	K	J	H	K	J	H	K
Mark I	8	12	-2.3	-2.7	-17.7	-17.7	-17.7	15.4	15.5	15.0
Mark II	18	22	-3.1	-3.4	-18.4	-18.7	-18.0	15.1	15.6	14.7

Table 3.8 Measured sensitivities on the telescope through a 5 arcsecond aperture (Mark I) and a 10 arcsecond aperture (Mark II).

The zero points show that the Mark II cryostat was more efficient in collecting radiation, a consequence of the better optical alignment and the more appropriate anti-reflection coating on the lenses. The noise was worse for the Mark II system due to the higher shunt capacitance associated with the larger detectors. The likely source of this noise is the JFETs and could probably have been decreased by temperature controlling the JFETs. However the sensitivity compared very favourably with other cryostats which did not offer the two channel capability.

3.7 Data Processing System

3.7.1 Description

The output of the detector preamplifiers consists of the signal from the source, modulated by the chopper, and noise. The use of a phase sensitive detector (or lock-in amplifier) allows effective discrimination of the signal from the unmodulated noise. The lock-in amplifier used, Ithaco Model 391A, uses the heterodyne principle, mixing the signal with an oscillator signal derived from the reference (chopper) frequency. This technique has the advantage that harmonics of the reference frequency are rejected. Although information is contained in the harmonics (since the modulating signal is approximately a square wave), the signal to noise ratio is worse than at the fundamental frequency if the noise spectral density is approximately flat.

The lock-in amplifier (one for each channel) generates an analogue signal proportional to the component of the signal in phase with the chopper (or at a fixed phase difference from it). This signal is then integrated over an interval of time to achieve an adequate signal to noise ratio. The integration and digitisation is performed by a voltage to frequency converter and a counter (Fig. 3.13). The digitised data are stored, analysed and displayed by the PDP 11/40 UKIRT Instrumentation Computer (Stewart 1983a).

The lock-in amplifier contains a preamplifier, whose gain is selected so that either the signal represents a significant proportion of full scale deflection (about 50 percent) or the noise from the detector preamplifier is visible in the output from the lock-in amplifier (about 5 percent with a time constant of 0.4 seconds). The output of the lock-in amplifier is an analogue signal in the range 0 to 10 volts, with zero input signal being represented by 5 volts. The voltage to frequency converter (VFC) is an Analog Devices Model 460 operated in a circuit to provide a conversion rate of 50 kHz/V. The signal fluctuations therefore correspond to at least 2.5×10^4 pulses/second and so quantisation errors are completely negligible.

The gating and integration of these pulses are performed by Camac modules (Borer 1008 counter/timer and Nuclear Enterprises 003/4 scalar). The timing is controlled by the 1 MHz quartz oscillator in the 1008 module which opens a gate for the specified number of microseconds. The counters are read by the computer which processes and displays the data and controls the observations.

3.7.2 Stability

The stability of the components in the data processing chain after detection is important, since the data are the results of calculating the differences of sets of measurements (Table 4.1). These comprise the signal plus a large zero offset viz. the lock-in

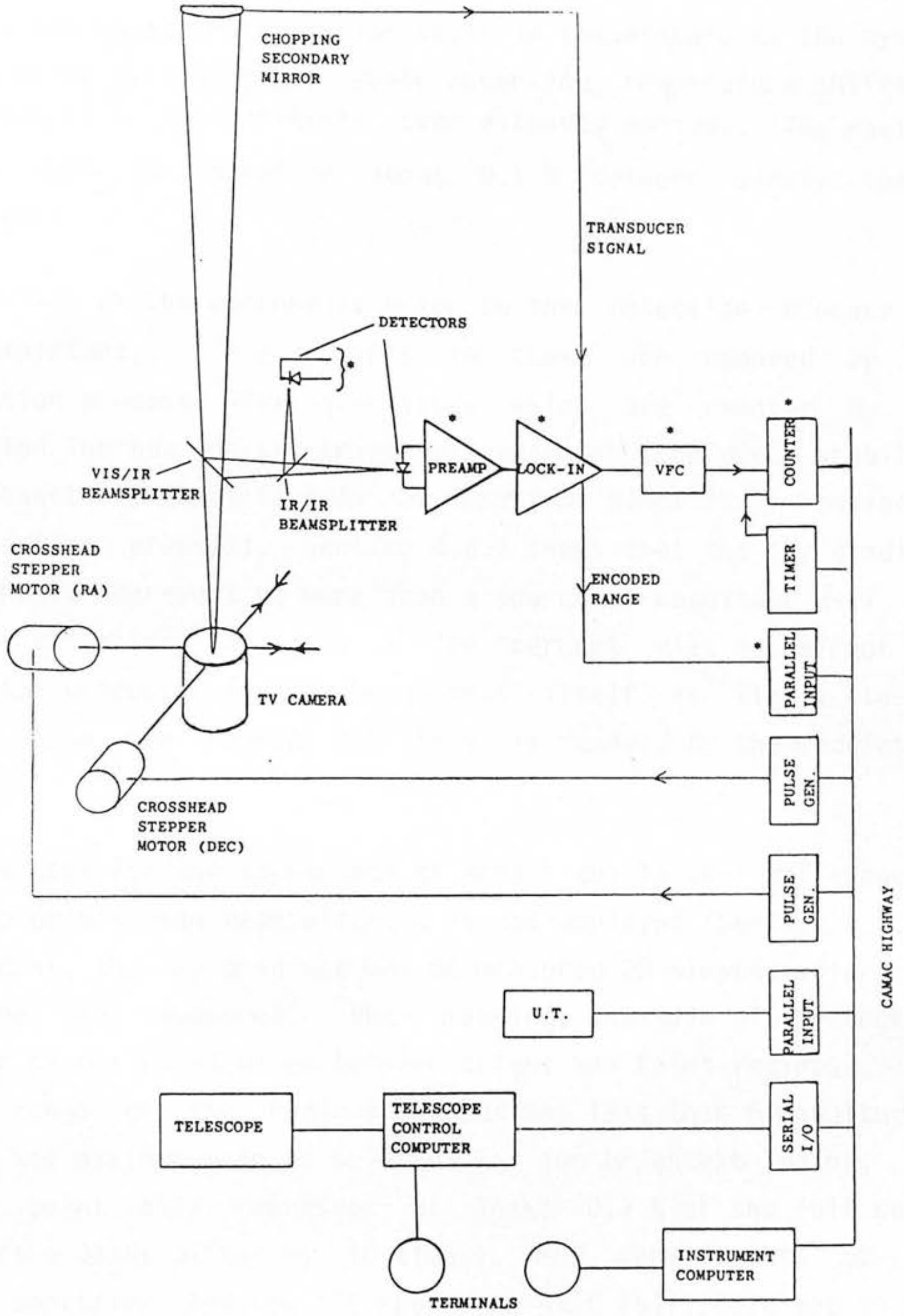


Fig. 3.13 The data acquisition and telescope configuration used to make the observations.

amplifier output for no signal and the VFC output offset. The temperature stabilities of both the lock-in amplifier and the VFC are both 0.01 %/°C, while the stability of the quartz timer is considerably better than this. The quoted figures for time varying drift are less than the temperature dependent drift. The VFC is housed in the Camac crate and can shift in temperature as the system heats up after switching on. While observing, temperature shifts of more than 10 °C are unlikely over extended periods. The maximum shift in gain is therefore about 0.1 % between widely spaced measurements.

The stability of the components prior to the detection process is less important, since shifts in these are removed by the demodulation process. The quantities which are removed by the subtraction include the sky brightness gradient (the d.c. stability of the detector preamplifier is not important since it is removed by the detection process). Section 4.6.1 shows that the sky gradient signal should represent no more than a source of magnitude H=17 and so gain variations of even a few percent will not affect the subtraction process. The sky brightness itself is likely to be brighter than the source but this is removed by the modulation process.

The worst case for the appearance of errors due to the differencing procedure occurs when beamswitching is not employed (Section 4.2.2). In this case, the sky gradient may be measured 20 minutes after the data have been measured. When mapping, the gain of the lock-in amplifier cannot be adjusted between bright and faint regions. The dynamic range of the regions mapped was less than 5 magnitudes. Assuming the maximum gain is selected for the brightest point, the faintest point will represent at least 0.3 % of the full scale reading (the gains differ by 10 times). The zero points of the lock-in amplifier and the VFC represent half full scale and so the shift in zero point will be in the order of 0.007 %/°C. The faintest point could therefore be subject to an error of 2 %/°C of temperature shift.

Most of the areas of the maps are significantly brighter than the faintest point. The integration times were chosen so that this faint point would be expected to have a signal to noise of about 10 relative to the noise generated by the detector and preamplifier. Temperature shifts of up to 5 degrees should therefore not degrade the data significantly. While beamswitching, such temperature shifts cannot occur over the measurement interval of 20 seconds. When beamswitching is not performed, the measurements of the source and sky brightness gradient can differ by 20 minutes and so significant temperature shifts are possible but not likely.

3.7.3 Computer Configuration

The instrument is controlled by the standard UKIRT data acquisition program PHOTOM (Stewart 1983b). This program accepts data from the two counters connected through the data acquisition chain to the two detectors. Observations were sequenced in pairs for beamswitching and statistics were performed on the differences. Zero points were derived from observations of standard stars and used to calculate actual magnitudes and colours for the observations of the sources. Zero point shifts and airmass corrections were applied subsequently.

To allow account to be taken of the lock-in amplifier gain, the range switches on each were encoded and connected to the computer through a parallel interface. Universal Time was obtained by accessing a clock slaved to the time signal from a satellite.

Control of the telescope was achieved either by moving the crosshead carrying the television camera or by sending commands directly to the telescope control computer. The TV camera lies on the same optical path as the infrared detectors; the light is split by a dichroic beamsplitter. Accurate telescope movements were achieved by moving the TV camera a known amount in declination and RA and moving the telescope to bring the image on the TV screen back to its original posision. The TV camera movement was effected by sending strings of pulses from the computer to two stepper motors on the crosshead. The scale was 165 motor steps per arcsecond so that good resolution could be obtained. The resolution was limited by the accuracy with which the image could be positioned on the TV screen and was typically 1 arcsecond in wind-free conditions.

The other method of telescope positioning used was to send commands from the instrument computer to the telescope control computer. The communications link consisted of a serial line operating at 9600 baud between serial interfaces in the Camac crates of each computer. The interfaces used, Kinetic Systems Model 3340, use a first in first out (FIFO) buffer so that the computer is interrupted at a rate of once per message and not once per character. Offset commands, representing the distance to be moved from a reference position, are sent between the computers using a message format similar to that the observer would use at the telescope control console. The control computer sends a message acknowledgement after the message has been verified. A status word, representing the current pointing error, is also sent so that the instrument computer can start an integration when the telescope is sufficiently close to the desired position. This method of positioning was not used until accurate encoders (20 bit implementation giving a resolution of 1.23 arcseconds) were installed on the telescope. (Parts of this software were written by D.C.Richardson and K.Krisciunas).

The same serial link is used to transmit data from the telescope control computer to the instrumentation computer. These data include apparent coordinates and sidereal time and, together with the derived airmass, are printed with the data to allow subsequent checks and airmass corrections.

4. The Observations

4.1 General Procedure

The data were all obtained using the 3.8m United Kingdom Infrared Telescope (UKIRT) between March 1981 and August 1982 using the specially built two channel cryostat. Two versions of the cryostat were used, with the second offering the choice of larger apertures, improved beam profiles and better stability. The Mark I version was used until February 1982 with a 4 or 5 arcsecond aperture on the sky. The observations of July and August 1982 were made with the Mark II cryostat and a 10 arcsecond aperture.

The emphasis throughout was on obtaining accurate colour data at J-K and H-K, for which simultaneous measurements were necessary. Standard stars were used in the normal way to determine instrumental zero points, but checks on zero point shifts, due to instrumental or airmass effects, were made using the nucleus of the galaxy as the local standard. This eliminated the need for telescope slewing and allowed frequent checks to be made. This was especially necessary for the Mark I cryostat observations as the zero points tended to drift.

The galaxies observed all have bright and well-defined nuclei in the infrared. The signal was peaked up using the K filter, since, at this wavelength, varying extinction causes smaller brightness variations. In some cases, such as NGC 4565, this infrared nucleus was displaced from the optical nucleus due to the heavy extinction. The K peak served as the datum point from which all telescope position offsets were measured. Guiding was usually performed on the optical image of the nucleus in the TV monitor, but in a few cases a nearby star was used instead.

The configuration of the telescope and instrument is shown in Fig. 3.1. Chopping between source and sky was effected by the hydraulically driven chopping secondary mirror. This chopped beam was reflected into the cryostat of a dichroic beamsplitter which allowed the TV camera to see the same object as the infrared detectors. Frequent peaking up on the nucleus eliminated the effects of flexure and differential refraction; adjustments were never required in excess of one arcsecond.

The TV camera is mounted on a crosshead which can be moved accurately in two axes by stepper motors. This allowed small telescope offsets to be accurately controlled by moving the camera a known amount and then repositioning the telescope so that the image was always at a fixed point on the screen of the TV monitor. This method of telescope positioning was necessary for most of the observations, since the telescope encoders were only accurate to about 5 arcseconds over small distances. A new telescope encoder system, accurate to 1.2 arcseconds, was installed in 1982 and so for some of the later observations the telescope encoders were used directly.

The transducer signal from the chopping secondary was monitored in the control room so that the system could be set up to provide as good a chop as possible. It was found on some occasions that the chop waveform degraded with time and in these cases readjustments were made during the night. Before and after such adjsutments the galaxy nucleus was always measured. In spite of these adjustments a good square wave chop was often unattainable. The effects of poor chopper performance are analysed in Section 4.3.

Beamswitching techniques were used to correct for instrumental and sky offsets. For most of the observations, beamswitching was performed at every point after 20 seconds of signal integration. For some of the two dimensional maps, the beamswitching was performed only after the whole area had been mapped in one beam which took abut 20 or 30 minutes. This allowed the telescope to be used more efficiently, since less time was taken up by beamswitching.

The telescope focus was adjusted at the start of each night by measuring the beam profiles and choosing the focus setting that produced the flattest tops and steepest sides. This also corresponded to the setting at which the beams were best matched and usually, but not always, produced the maximum signal. The beam profiles in the K channel of the Mark I cryostat were dependent on the focus setting and so the telescope focus was readjusted if there was evidence that it had varied. Before and after such measurements the galaxy nucleus was measured.

4.2 Chopping and Beamswitching

4.2.1 Observations with Beamswitching

At infrared wavelengths the sky tends to be bright and is often brighter than the sources being measured. This is especially true at thermal wavelengths, i.e. around 10 μm, but can also be significant at near infrared wavelengths. Thermal emission and scattering at these wavelengths is insignificant and airglow emission due to the OH- radical is likely to be dominant (Suits 1978). This is strongest in the H band but is also present at J and K and varies with a timescale of several hours.

At H the spectral radiance is approximately 4 x 10^{-8}W/sr/cm^2/μm, which in a 10 arcsecond aperture is equivalent to the signal from a star of magnitude H = 8, assuming that an H magnitude of zero is equivalent to a flux density of 11.5 x 10^{-14}W/cm^2/μm (Allen and Cragg, 1983). To allow fainter objects to be measured, chopping techniques are used so that the detector sees alternately the source and a region of sky nearby. By subtracting the two measurements, usually electronically in a phase sensitive detector, the brightness of the source may be determined.

Because it is not possible to move the telescope rapidly, the chopping is normally performed by the secondary mirror, which at f/35 can be made light enough to be vibrated at 10 Hz in an approximate square wave. The secondary mirror is undersized so that the detector sees empty sky rather than the hot interior of the building, but this means that the whole primary mirror is not seen by the detector and different parts, possibly with different reflectivities, are seen in each beam. In addition the secondary mirror support presents slightly different images to the detector in the two beams. Thermal emission (at longer wavelengths) and reflection of sky brightness or dome lights may therefore cause a spurious signal to be detected, in addition to the signal from the source. This telescope offset must be removed by a subtraction process.

To remove these effects beamswitching is employed, in which the whole telescope is moved so that the source appears in the other beam. The signal due to the source is inverted in phase, whereas the instrumental signal remains constant and so subtracting the two measurements removes the instrumental signal. Linear variations in the sky background and d.c. offsets introduced after the modulation are also removed. In addition, the system may be operated at a frequency at which the system noise is low. Electronic devices such as JFETs and also the sky generate 1/f noise which is much reduced by modulating the signal at 10 Hz. The sources of offset and their methods of removal are summarised in Table 4.1.

	Chopping (10 Hz)	Beamswitching (20 seconds) or Sky Measurement (20 minutes)
Telescope	X	X
Sky brightness level	X	
" " gradient		X
Detector preamplifier	X	
Lock-in amplifier		X
Voltage-frequency conv.		X

Table 4.1 The sources of instrumental offset and their method of removal.

Beamswitching must be performed frequently enough to ensure that the sky and instrumental offsets do not vary significantly between measurements. Often the telescope is beamswitched after about 20 seconds of signal integration, but this is largely for the convenience of making statistical estimates of the signal to noise ratio. As will be demonstrated in Section 4.2.2, less frequent beamswitching, or even none at all, is possible with a subsequent increase in the observing efficiency.

4.2.2 Avoidance of Beamswitching

In a 10 arcsecond aperture, signal levels corresponding to a K magnitude of 12 (or 16.7 magnitudes per square arcsecond) are obtainable from a large area of a galaxy such as NGC 7331. Since the detector system has a 1 sigma 1 second magnitude of K=15, a signal to noise ratio of 40 can be obtained in a 5 second integration if the K magnitude is 12. In such a situation it is very inefficient to perform beamswitching around each point if each beamswitch takes about 10 seconds.

As mentioned in Section 4.2.1, the reason for beamswitching every 20 seconds in the near infrared is not necessarily based on noise considerations. Linear gradients in the sky background and instrumental offsets could possibly be determined much less frequently with a consequent increase in observing efficiency. The sky emission at the wavelengths of the J, H and K bands is believed to be dominated by OH- emission (Suits 1978) which is strongest in the H band. The emission originates in the ionosphere in regions of typical size 2500 km, which drift at a typical rate of 70 m/s, so taking 10 hours to pass over. Chopping removes the constant component of this emission, leaving the linear gradient. If the emitting source is at a height of 200 km, a 3 arcminute chop will subtend a region 7×10^{-5} times the linear extent of the source. Assuming a linear gradient over the extent of the source, such a chop would cause the detector to see a signal 1.4×10^{-4} times the peak signal in the source (estimated to be equivalent to H=8 in

Section 4.2.1). This would correspond to an H magnitude of 17.6, which is almost 100 times fainter than the bright regions of the galaxies under consideration. It should be possible therefore to avoid beamswitching altogether for bright sources.

The d.c. offsets in the electronics associated with the demodulated signal (Section 3.7) and the instrumental modulated signal still need to be removed. A convenient method is to observe a region of sky close to the object being observed. This means that minimal time is spent repositioning the telescope and more importantly ensures that stray reflections are similar. Also the gradient of the sky emission emission can be removed; although this is not as effective as with frequent beamswitching, this is not important because the sky emission is faint.

The technique employed, for the observations which did not include beamswitching, was to perform a sky measurement before and after the complete source had been mapped. The source was then observed in the other beam and another sky measurement was made. Ideally these two maps should have been equal and inverted with respect to each other, but this was found not to be the case. The imperfections in the chopper waveform, as described in Section 4.3, are believed to have been primarily responsible. In addition the sky measurements were found to vary, as described in Section 4.6.1, but could be adequately removed by interpolation in most cases.

4.3 Effect of Chopper Risetime

Ideally the chopping secondary mirror should execute a square wave so that the detector sees alternately two positions on the sky but not the region between them. This implies that the mirror must move very rapidly with a large acceleration at the start and end of the transitions. The system was designed to have a settling time of 20 ms with a 100 arcsecond chop, but the exposed environment of Mauna Kea meant that this was sometimes not met.

The motion was transferred from the drive mechanism to the mirror through a system of hydraulic pipes, which travelled the length of the telescope structure. They were exposed to the strong and cold winds, which cooled the hydraulic oil making it more viscous. The open loop gain of the control system servo was insufficient to allow the increased viscosity to be compenstaed for and so the chopper waveform degraded. The servo parameters could be periodically adjusted to provide the optimum waveform but this was often far from the desired square wave. Between adjustments the wavefrom degraded gradually.

The motion of the mirror, as observed by inspecting the transducer signal, could be represented by a square wave modified by a time constant τ . This is not completely realistic since it implies an infinite acceleration at the start of the transition, but provides an easily parameterised model for estimating the effect of chopping across an extended source. Using this model the position, x, of the mirror at time t is

$$x(t) = A \,.\, \exp(-t/\tau) \,/\, (\, 1+\exp(-T/\tau)\,)$$

where A is the low frequency amplitude of the chop, τ is the time constant and T is the interval between transitions. This formula may be rearranged to give the time t at which the chopper passes position x, i.e.

$$t(x) = -\,\tau .\, \ln(\, (1+\exp(-T/\tau)).\, x/A\,)$$

This formula was used to calculate the length of time the chopper spends at each position in the chopper range, which was assumed to be 180 arcseconds and divided into intervals of 0.1 arcseconds. Using an assumed brightness distribution, b(x),of the extended source, the contribution I(x) to the measured signal from each point in b(x) was calculated assuming that a circular aperture of 10 arcseconds was used. T was assumed to be 67 ms (7.5 Hz chop). The brightness distribution was taken from the observations of NGC 7331 made on 4 August 1982. Outside the region observed the brightness was assumed to decrease exponentially at a rate consistent with the outer two measured points.

The chosen point was 28 arcseconds SW of the nucleus. The chop was almost perpendicular to the major axis and so the detector saw a typical region of the major axis but not the nucleus while in transition. The calculations were made with the point in the left and the right apertures so that in one case the chop was across the major axis and in the other was towards fainter regions. Fig. 4.1 shows I(x) for these two cases superimposed on the source surface brightness. For a perfect square wave ($\tau = 0$) the two functions are equal, but as τ increases the functions differ significantly.

For a typical value of τ equal to 5 ms, the proportion of the signal contributed by the position at the end of the chop is only 50 percent when the chop is across the major axis, but this proportion is 87 percent when the chop is across the fainter regions. The proportions are calculated for other values of τ in Table 4.2. On one side of the major axis the left beam measurement will significantly overestimate the intensity, while on the other side the right beam measurement will be in error.

Time constant ms	Proportion of signal from measured position in: left beam	right beam	Proportion of time chopper is at end end position
0	100 %	100 %	100 %
1	83.9 %	97.2 %	95.7 %
2	72.2 %	94.5 %	91.4 %
5	50.8 %	86.6 %	78.4 %
10	33.9 %	74.0 %	56.9 %

Table 4.2 The effect of the chopping secondary time constant on the proportion of the signal originating in the measured position (assumed to be a point 28 arcseconds SW of the nucleus of NGC 7331, and on the time that the detector sees that part of the source within one beamwidth of end position. The beam size is assumed to be 10 arcseconds.

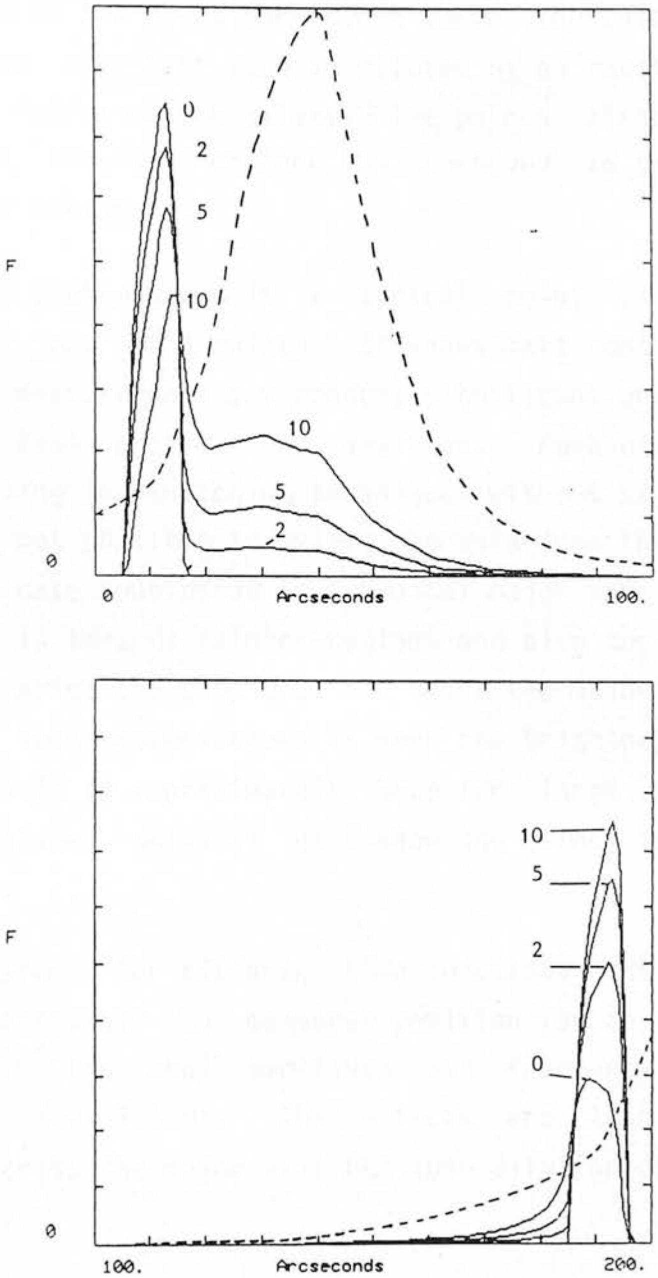


Fig. 4.1 The effective beam profiles when the chop is across the major axis of a galaxy (a) and away from the major axis (b). The chop frequency is 7.5 Hz and the chop rise time in ms is indicated on each profile. The galaxy brightness (dotted line) is a typical profile across the major axis of NGC 7331 (see text).

Similar considerations apply to the accuracy of the colour data. The chopper modulates both channels in an identical manner and so the colour measurements in the two beams may be considered equally accurate but representing different effective regions of the source. In the case of the point considered above, the colour of the point measured in the left beam is diluted by an equal amount of signal from other reions of the galaxy. The points although observed in both beams should therefore be combined so that these dilution effects are minimised.

The example quoted above is a typical point observed in a two dimensional map of a galaxy. It shows that combining the left and right beam measurements may produce significant uncertainty in the effective beam profile and position. Much of the data has been obtained using beamswitching techniques without sky measurements and so it is not possible to select the data from the most appropriate beam. The case considered is a typical major axis point for which the chop is towards fainter regions and also applies to minor axis points for which the chop does not cross the major axis. The worst case for such observations is when the brightness distribution is flat, as would be approximatelty true for large face-on galaxies. Highly inclined galaxies will show the effect much less than this worst case.

Table 4.1 shows, for different time constants, the proportion of the signal contributed by measured position for apertures of 4 and 10 arcseconds. The chop amplitude and frequency are again 180 arcseconds and 7.5 Hz. The effects are less severe than when chopping across the major axis but some dilution of the colours will still occur.

Variations in the chopper waveform did not affect the accuracy of the colours themselves but did alter the effective beam profile. If simultaneous measurements were not made at two wavelengths, the colours themselves could be seriously in error. The use of the two channel cryostat for these observations therefore avoided the worst effects of varying chopper waveforms.

Time constant ms	Proportion of signal from measured position 10" aperture	 4" aperture
0	100 %	100 %
1	93.0	89.9
2	86.8	80.9
5	71.8	60.2

Table 4.3 The effect of the chopping secondary time constant on the measured signal assuming a flat brightness distribution.

4.4 Standards

4.4.1 Filter Systems

A system of infrared standards was set up by Johnson (1966a) for the J, K and L passbands, which he also established. As pointed out by Allen and Cragg (1983) there are some inconsistensies in Johnson's work, viz. a change in the definition of the J passband and a change in the magnitudes assigned to alpha Lyrae which had previously used as the zero point of the magnitude system. Since Johnson's early work the H filter has become widely used, utilising a clean and wide atmospheric window.

The J atmospheric window is less well defined than the H and K windows and can be affected by a water absorption feature at about 1.1 μm (Manduca and Bell 1979). In addition there are several commonly used J filters with different passbands. The same filter may therefore produce different colour equations at sites of different altitudes. It is also the more likely to require airmass corrections than either H or K.

Because of the lack of any universal standard, workers have frequently transformed their data to the Johnson system using colour equations derived from a few stars. Until recently little success has been achieved in relating the different filter systems to each other. Recently however the California Institute of Technology (CIT) and the CTIO systems have been made equivalent, as have the Anglo Australian Observatory (AAO) and Mount Stromlo Observatory (MSO) systems. Colour equations have also been derived to allow transformations between the CIT/CTIO and AAO systems. The UKIRT filters are from the same batch as those used at AAO, but because of the different altitude and possible colour effects in beamsplitters and optics, colour equations have also recently been derived between the UKIRT and CTIO systems.

The MSO and AAO systems are related (Jones et al 1980) by

$$(J-H)_{AAO} = 1.07\ (J-H)_{MSO}$$

The CIT and CTIO systems became similar after June 1980. The J-K colours then became identical and the H-K colours differ by no more than 0.025 magnitudes (Elias et al 1982). The relationship between this system and the UKIRT system (denoted by subscipt u) is (Williams 1983)

$$(J-K)_{U} = 1.07\ (J-K)_{CIT}$$
$$(H-K)_{U} = (H-K)_{CIT}$$

For measurements on the CTIO system pior to June 1980 the transformation is

$$(J-K)_{U} = 1.03\ (J-K)_{CTIO}$$

The AAO and CIT systems have been related by Elias et al (1983) giving

$$(J-K)_{CIT} = 0.90\ (J-K)_{AAO}$$
$$(H-K)_{CIT} = 0.95\ (H-K)_{AAO}$$

implying that

$$(J-K)_{U} = 0.96\ (J-K)_{AAO}$$

The multi-aperture galaxy data of Aaronson (1977) are on the HCO system; these zero points differ to those of the CIT system but are close to the UKIRT system.

The H-K zero points agree to within 0.014 magnitudes and so can be considered identical on all the filter systems. However discrepancies in the H-K colours of galaxies between the work of Frogel et al (1978) and Griersmith et al (1982) have been observed. Jones and Hyland (1982) suggested that this effect could be due to small differences in the size and shape of the H beam profile relative to the K beam profile. Koorneef (1983b) suggested that it could be the result of small differences in the effective wavelengths of the H filter between the two systems. At H the stellar opacity is a strong function of wavelength and small differences in effective wavelength can have large effects in the measured colours. The AAO H-K colours agree very closely with those obtained by Koorneef and he considers the systems identical. However no late main sequence stars were included in the comparison and discrepancies could still exist. These problems indicate the difficulties involved in comparing infrared colours, especially of extended objects, obtained on different instrumental systems. The current observations are concerned primarily with the colour variations within single objects and so systematic errors are less important.

The difference between the UKIRT and AAO J-K may not be real since it is indirectly inferred on the basis of relatively few stars but may reflect the difference in altitude of the sites and properties of dichroic reflectors. Koorneef finds no significant difference between his and the AAO J-K colours, consistent with the conclusion of Jones and Hyland (1980) that the AAO system is identical to the Johnson as well as could be determined. The AAO, UKIRT and Koorneef colours will therefore be considered to be on the same system.

The standard stars used and their assumed magnitudes are listed in Table 4.4. The magnitudes are on the UKIRT system and were measured (Williams 1983) with the standard UKIRT single channel cryostats, which use filters from the same batch as those in the two channel cryostat and similar optical component materials.

Star	K	(J-K)	(H-K)	
BS 718	4.39	-0.005	0.00	
BS 1552	4.14	-0.11	-0.05	
BS 4550	4.385	0.56	0.07	
HD 203856	6.86	0.065	0.025	
HD 136754	7.135	0.02	0.01	
GL 748	6.305	0.79	0.23	
HD 201941	6.625	0.07	0.03	
BS 8541	4.25	0.05	0.025	(Elias et al 1982)
'BS 9538'	7.13	-0.01	0.00	(Elias private comm.)

Table 4.4 The assumed magnitudes and colours of the standard stars. The data are from Willimas (1983) except where indicated.

4.4.2 Use of Galactic Nuclei as Local Standards

The use of the nuclei rather than stars as local standards offered advantages due to their extended natures being similar to the points being measured. In the nuclear regions of galaxies the surface brightness can be described by the de Vaucouleurs (1948) law

$$\log\left(\frac{I}{I_e}\right) = -3.33\left[\left(\frac{r}{R_e}\right)^{1/4} - 1\right]$$

and so although it cannot be properly resolved, it is extended within the beam. Imperfections in the beam profiles or slight differences between them will therefore have less effect on the measured colours than for unresolved sources such as stars, in which all the radiation is concentrated into a small area of the aperture. For measuring extended sources an extended standard will therefore provide a more accurate determination of the colour zero point.

This effect has been calculated assuming that the profiles of the two channels are similar except for a narrow spike (Fig. 4.2). The beam profiles were otherwise assumed to be flat and of width 5 acseconds. The source profiles were convolved with a boxcar of width 2 arcseconds to represent the seeing and focussing errors and the instrumental response in each channel was calculated as this

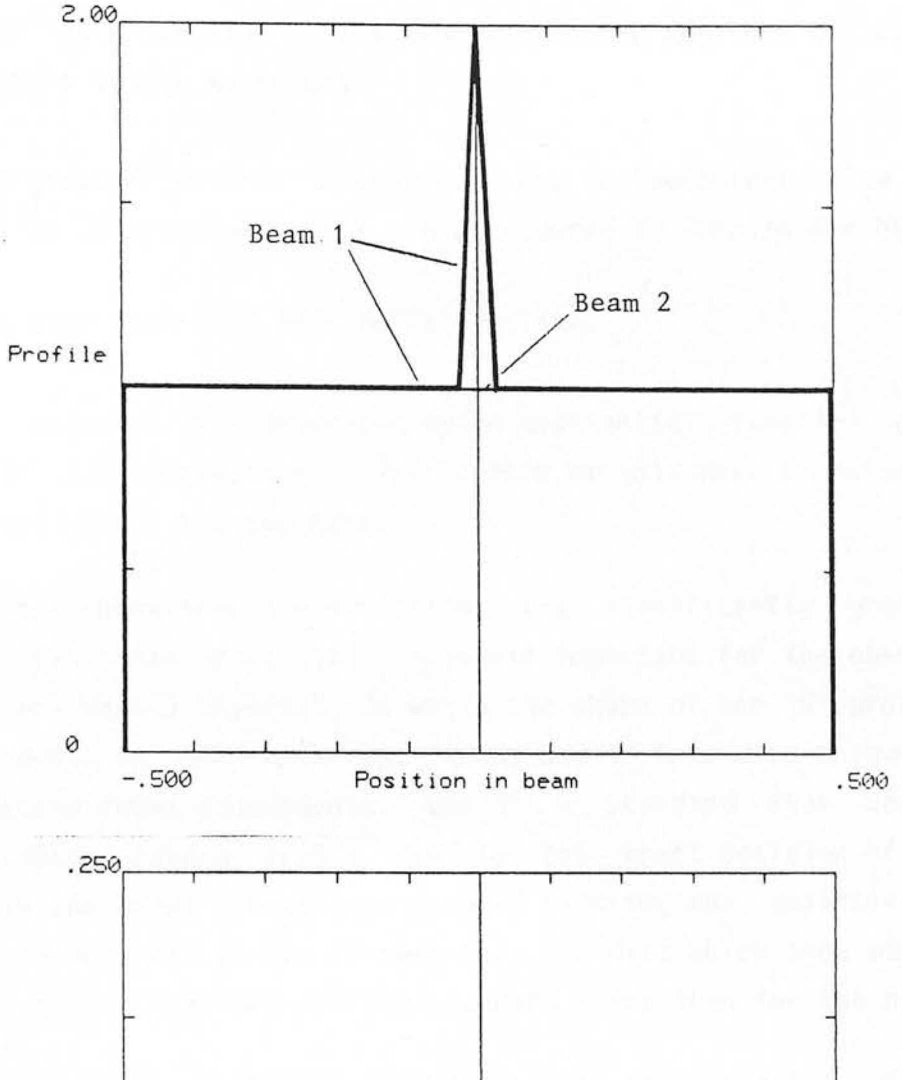


Fig. 4.2 The effect of mismatching beam profiles on the measured colours of a star, the nucleus of a galaxy and the disk of a a galaxy as the source is scanned across the aperture. The aperture is assumed to be 5 arcseconds, the seeing 1 arcsecond and the galaxy parameters correspond to NGC 7331 (Boroson).

profile was scanned across the aperture. The difference in apparent colour as a function of position within the aperture was calculated for three source functions.

(a) a galactic nucleus described by the de Vaucouleurs Law with R_e equal to 32 arcseconds (i.e. that deduced by Boroson for NGC 7331).

(b) a star described by a delta function.

(c) a galactic disk described by an exponential function of scale length 130 arcseconds. The centre of this disk is assumed to be scanned across the aperture.

Fig. 4.2 shows that the variations are significantly greater for case (a) than case (b). This was important for the observations with the Mark I cryostat, in which the shape of the K profile was dependent on the telescope focus and so tended to change between telescope focus adjustments. Use of a standard star would have introduced random errors due to the exact position of the star within the beam; these were reduced by using the galactic nucleus as the standard as can be seen from the data which show much larger zero point variations for the standard stars than for the nuclei.

The variations in the measured colours of the galactic disk, case (c), are much smaller than for either case (a) or case (b). Differences in beam profiles will therefore not affect the data if the zero points are derived from an equally extended object. Use of the galactic nucleus although not perfect is a significant improvement over a star. and so systematic errors in the colours will be determined only by the colour zero point errors of the nucleus.

For extended objects such as the regions of galaxies away from the nucleus the only significant effect of small differences in the beam profiles is a zero shift. This could cause systematic errors in the colours when compared with standard stars and is believed (Jones and Hyland 1982) to affect some colour data for galaxies and globular clusters in the literature. For surface photometry this effect is of little importance, when colour variations within single objects are being considered.

In practice, the zero points shifts of the galactic nuclei were smaller than those of the standard stars. For example, the measured J-K colour of the nucleus of NGC 4565 varied by only 0.05 magnitudes over 4 hours on 1 February 1982, when the measured zero points of the standard stars varied by 0.12 magnitudes. The use of the nuclei as standards therefore reduces the effect of zero point shifts. The data reduction procedure (Section 4.5.1) further reduces the random errors, although some systematic errors (up to 0.1 magnitudes) may remain.

4.4.3 Mark I Cryostat

The observations made of standard stars (Table 4.5) with the Mark I cryostat tended to have large errors due to the small apertures used (4 or 5 arcseconds) and the slightly peaky beam profiles. The statistical errors on the individual filter magnitudes were up to 0.07 magnitudes but the colour errors were not more than 0.02 magnitudes. The individual filter zero points shifted due to the chopper being readjusted, which was often necessary due to the low temperatures causing the hydraulic oil to become more viscous and so affecting the chopper response. These readjustments affected both channels by the same amount.

Gradual changes in telescope focus also caused zero point shifts, with the K channel affected much more than the reflected channel. The telescope focus was sometimes adjusted to compenstate. A standard star was usually observed before and after this procedure; if not the galaxy nucleus was observed to allow the change in zero points to be estimated.

As discussed in Section 4.5.1 these colour zero point shifts were effectively removed by the data reduction process, but a degree of uncertainty remains in assigning true colours to the data from the standard star observations. Comparisons of colours between objects may therefore be subject to systematic errors. Because the beam profiles were not flat topped and subject to variation it is not possible to convert the data to magnitudes per square arcsecond.

Date	Run	Star	Inst. Mag. J	Inst. Mag. K	Air-mass	Zero points J	Zero points K	Zero points J-K
31/8/81	3	BS 9538	-10.318	-10.176	1.01	-17.438	-17.306	-0.132
	14	BS 9538	-10.222	-10.054	1.10	-17.342	-17.184	-0.158
	15	BS 9538	-10.310	-10.167	1.11	-17.430	-17.297	-0.133
	28	BS 9538	-10.189	-10.129	1.31	-17.309	-17.259	-0.050
	29	BS 9538	-10.044	-10.021	1.32	-17.164	-17.131	-0.013
	38	BS 8541	-12.959	-12.844	1.15	-17.259	-17.094	-0.165
	48	BS 8541	-13.043	-12.927	1.12	-17.343	-17.177	-0.166
	49	BS 8541	-13.006	-12.864	1.21	-17.306	-17.114	-0.192
	93	BS 718	-13.009	-12.830	1.08	-17.394	-17.220	-0.174
29/1/82	4	BS 4550	-12.778	-13.269	1.63	-17.723	-17.654	-0.069
	19	BS 4550	-12.867	-13.413	1.19	-17.812	-17.798	-0.014
30/1/82	2	BS 1552	-13.069	-12.973	1.09	-17.099	-17.113	0.014
	50	BS 4550	-12.071	-12.468	1.27	-17.016	-16.853	-0.163
	51	BS 4550	-11.950	-12.339	1.27	-16.895	-16.724	-0.171
31/1/82	1	BS 1552	-13.104	-13.138	1.13	-17.134	-17.278	0.144
	30	BS 4550	-11.613	-12.190	1.37	-16.558	-16.575	0.017
	31	BS 4550	-12.342	-12.978	1.33	-17.287	-17.363	0.076
1/2/82	3	BS 1552	-13.289	-13.378	1.08	-17.319	-17.518	0.199
	29	BS 4550	-12.126	-12.769	1.55	-17.071	-17.154	0.083
19/2/82	2	BS 1552	-13.626	-13.597	1.07	-17.656	-17.737	0.081
	37	BS 4550	-12.736	-13.429	1.26	-17.681	-17.814	0.133
	47	BS 4550	-12.691	-13.399	1.11	-17.636	-17.784	0.148
	85	BS 4550	-12.758	-13.451	1.36	-17.703	-17.836	0.133

Date	Run	Star	Inst. Mag. H	Inst. Mag. K	Air-mass	Zero points H	Zero points K	Zero points H-K
31/1/82	91	BS 4550	-12.135	-12.494	1.10	-16.590	-16.879	0.289
1/2/82	1	BS 1552	-13.237	-13.409	1.09	-17.327	-17.549	0.222
	30	BS 4550	-12.518	-12.765	1.53	-16.973	-17.150	0.177

Table 4.5 Standard star observations and derived zero points for the Mark I cryostat.

Date	Run	Star	Inst. Mag. J	Inst. Mag. K	Air-mass	Zero points J	Zero points K	Zero points J-K
11/7/82	6	HD 203856	-11.383	-11.108	1.43	-18.308	-17.968	-0.340
	12	HD 203856	-11.416	-11.128	1.21	-18.341	-17.988	-0.353
12/7/82	2	HD 203856	-11.454	-11.155	1.15	-18.379	-18.015	-0.364
	17	HD 203856	-11.510	-11.192	1.12	-18.435	-18.052	-0.383
4/8/82	10	GL 748	-11.328	-11.649	1.05	-18.413	-17.954	-0.459

Date	Run	Star	Inst. Mag. H	Inst. Mag. K	Air-mass	Zero points H	Zero points K	Zero points H-K
3/8/82	6	HD 201941	-12.08	-11.36		-18.74	-17.99	-0.75
4/8/82	1	HD 136754	-11.586	-10.864	1.04	-18.731	-17.999	-0.732
	11	GL 748	-12.187	-11.654	1.05	-18.722	-17.959	-0.763

Table 4.6 Standard star observations and derived zero points for the data taken with the 10 arcsecond aperture of the Mark II cryostat.

4.4.4 Mark II Cryostat

The statistical errors in the measurements of the standard star data of Table 4.6 are less than 0.02 magnitudes for the individual filter magnitudes and 0.01 magnitudes or less for the colours. The colour zero points are much more stable than those obtained with the Mark I cryostat and are good to within about 0.02 magnitudes through each night. More extensive data were obtained for the nuclei of the galaxies to allow correction of possible airmass and instrumental shifts. No corrections were found to be necessary for instrumental shifts, and on only one occasion was an airmass correction demonstrating the much improved performance of the Mark II cryostat over the Mark I.

4.5 Reduction of Beamswitched Data

Most of the one dimensional scans and some of the two dimensional observations were made using the beamswitching technique, in which each point was measured until an adequate signal to noise ratio had been achieved.

Each position was measured individually and the telescope was beamswitched so that several data were taken with the source in each of the two beams, "left" and "right". The telescope was beamswitched every 20 seconds and the differences between the "right" and "left" measurements were calculated and printed for each pair. The running mean and standard error of the mean were also printed as the measurements were made and so the observation could be continued as long as necessary to achieve the desired accuracy. The average sky background, linear variations in the sky background and telescope and instrumental offsets were automatically removed by this technique and so no separate sky measurements were required.

The chop direction was selected to be perpendicular, or close to perpendicular, the major axes of the galaxies to allow the reference beams to lie on regions of the sky as faint as possible. The amplitude of the chop was made large enough to chop well away from the galaxy, so that no corrections to the data are necessary for emission in the reference beam. As a typical example, a 3 arcminute chop at 75 degrees to the major axis was used for NGC 7331 whose minor axis semi-diameter is 2.0 arcminutes at 25 B magnitudes per arcsecond^2. Even for points one arcminute along the minor axis, the reference beam closest to the galaxy was at twice the D_{25} distance so contributing negligible signal.

4.5.1 Zero points

Several observations of standard stars were generally made throughout the night, but greater user was made of the galactic nuclei as local standards as described in Section 4.4.2. The galactic nucleus was observed frequently enough to remove airmass effects and instrumental zero point shifts. The latter were significant in the obervations made with the Mark I cryostat, whose beam profiles were sensitive to the telescope focus. As this tended to change gradually through the night, the zero points shifted and at different rates for the two channels.

The profiles degraded more in structure than in size, becoming less flat topped than at optimum focus. Because of the cryostat's common focal plane aperture the profiles did not shift relative to each other and also did not become measureably narrower. In observations of extended sources, this had the effect mainly of altering the zero point in a gradual manner. In observations of stars, however, such degradations would have resulted in large random errors dependent on the positioning of the point source relative to variations in the beam pofile.

Each measurement of a galactic nucleus provided a colour zero point, say J-K, and also magnitude zero points relating to an individual filter. With the Mark I cryostat, significant variations were observed with both these types of zero points. The colour zero point varied monotonically whereas the magnitude zero points varied in a partly random manner. The explanation for this effect lies in the fact that guiding errors and windshake sometimes took the beam slightly away from the galactic nucleus, and when this happened the measurement was fainter by some random amount. The colour zero point was little affected by this, since the majority of the signal contributing to the colour still came from the nucleus of the galaxy. The colour zero points shifts were therefore attributable to the beam profile degradations.

The method of correction applied to the Mark I cryostat data was to force the colours of subsequent nuclear measurements to be equal to the colour of the first nuclear measurement, i.e. use the nucleus as a local standard. Each time the nucleus was measured colour zero points were calculated and these were interpolated for the intervening measurements of other points in the galaxy to give colours corrected for zero point drift.

The J (or H) magnitudes were left uncorrected and the K magnitudes were recalculated from the corrected colour and the uncorrected J magnitude. The K data were was chosen for correction rather than J or H, because the K channel suffered more from sensitivity to the telescope focus than the reflected channel. The correction procedure may be summarised as follows :

$$(J\text{-}K)_c = (J\text{-}K) - (J\text{-}K)_a + (J\text{-}K)_o + \left\{(J\text{-}K)_a - (J\text{-}K)_b\right\}\cdot\frac{t-Ta}{T_b-T_a}$$

where (J-K) is the colour of the point measured at time t, $(J\text{-}K)_a$ is the nuclear colour measured at time T_a and $(J\text{-}K)_b$ is the nuclear colour at time T_b (all instrumental magnitudes). $(J\text{-}K)_o$ is the true nuclear colour as determined by comparison with a standard. The corrected K magnitude is then calculated :

$$K_c = J - J_o - (J\text{-}K)_c$$

where J is the measured J instrumental magnitude, J_o is the true J magnitude of the nucleus as determined by comparison with a standard and K_c is the corrected K magnitude. The H-K data were corrected in the same manner.

This method can be demonstrated to give accurate results by comparing the K measurements obtained in separate observations with J and H. For NGC 7331 there are J-K and H-K data for the E/W direction measured on 26 August 1981. When the K magnitudes are compared the rms difference is 0.05 magnitudes and this can be attributed to small pointing inaccuracies since the brightness gradient is about 0.06 magnitudes per arcsecond. Using the Mark II cryostat, with which this correction was unnecessary, an rms difference of 0.07 magnitudes was obtained along the minor axis.

When the telescope focus was observed to have degraded by a significant amount, it was readjusted by optimising the beam profiles. Before and after such adjustments the nucleus was measured. However, because the zero points were linearly interpolated between nuclear measurements, the data reduction procedure did not depend on when the telescope focus was adjusted.

The data obtained with the Mark II cryostat did not require this type of correction due to the much better beam profiles and stability. The galactic nuclei were observed as local standards as in the previous observations, but no automatic correction procedure was necessary. The stability of the colour zero points was generally good to 0.02 magnitudes through the period of the observations, but on one occasion, 5 August 1982, significant shifts were observed in the J-K and the J zero points. These were corrected for as an airmass effect and are discused in Section 4.7.1

4.5.2 Colour and Brightness Errors

The errors of the individual wavelength magnitudes were estimated by assuming that the fluctuations obeyed a normal distribution. The standard error of the mean, σ_μ, was calculated from the standard deviation s of the measurements. This was divided by the mean, $\overline{x}$, to give a proportional error

$$\varepsilon_x = {}^{s}/(\overline{x}\sqrt{n}) \qquad (4.1)$$

This proportional error was recalculated and printed for each channel as the data accumulated, allowing judgements to be made on the quality of the data as they were obtained. This error is also very close to the erro expressed as a magnitude; for 1, 10 and 30 percent errors the equivalent magnitude errors are (-0.011,+0.011), (-0.103,+0.114) and (-0.285,+0.387) respectively.

When guiding errors, seeing fluctuations and windshake are the dominant sources of error, then the fluctuations in the measurements of the two channels are highly correlated, due to the common focal plane aperture of the cryostat, and the ratio of the the two measurements (X_1 and X_2) has much smaller variations than the measurements themselves. A valid estimate of the error in the colour may be obtained from the colour C of each observation

$$C = -2.5 \log \frac{X_1}{X_2} \tag{4.2}$$

The standard error of the mean of these colour measurements is then calculated from their standard deviation s_c

$$\varepsilon_c = s_c / \sqrt{n} \tag{4.3}$$

and the mean of the C values provides a good estimate of the colour. in this case is the estimated error of the colour already expressed as a magnitude.

If the measurements X_1 and X_2 are intrinsically noisy, then the errors in them are uncorrelated and the ratio X_1/X_2 provides a poorer estimate of the colour than

$$C = -2.5 \log \frac{X_1}{X_2} \tag{4.4}$$

$$\varepsilon_c = \sqrt{\varepsilon_1^2 + \varepsilon_2^2} \tag{4.5}$$

If the errors in X_1 and X_2 are correlated then calculating the colour by either method gives essentially the same result, and so equation 4.4 is used in all cases since it is also valid when the errors are uncorrelated. In bright regions applying equation 4.5 will overestimate the error of the colour. For example, if the nucleus is at the edge of the aperture then large errors, say 5 percent, due to slight guiding errors can be expected and equation 4.5 would then give a colour error of 7 percent. Equation 4.3 might on the other hand give an error of 1 percent; examples of such cases occur in the tabulated results. In this situation the error is really in the position and is independent of the brightness gradient.

Equation 4.3 would likewise over-estimate the error in faint regions of the galaxy and is invalid if any of the X_1 or X_2 measurements is zero or negative. For a large range of brightnesses equations 4.3 and 4.5 give similar estimates of the error, and equation 4.5 is used for all the data except for positions close to the nucleus where it is obvious, by and being much larger than expected for the values of X_1 and X_2, that equation 4.3 should be used.

4.6 Reduction of Rastered Data

4.6.1 Sky Background Level

The two dimensional maps of July and August 1982 were obtained by observing the complete source entirely in one beam, without beamswitching. A one dimensional scan was obtained in the same manner on 11 July 1982, this method being used to check the feasibility for two dimensional mapping. Although this method provided less immediate correction of offsets and backgrounds, it allowed more efficient use of the observing time.

Although beamswitching was not employed, the source was chopped in the usual manner to perform the subtraction of constant sky background and detector preamplifier d.c. offset. As described in Section 5.4 the data so obtained still contained contributions from the variation in sky background with position, the a.c. instrumental offset due to the detector seeing different parts of the telescope in the two beams, and the zero signal output of the lock-in amplifier and the voltage to frequency converter.

The sum of these components was determined by making measurements on a region of the sky near the source being measured and itself containing no significant sources of emission. Because the chop remained the same for the source and sky measurements, so long as the components did not change with time, they were removed by subtracting the sky from the source measurements. Any non-linear variations in the sky background would still be present, but would not be removed by any beamswitching technique. Such variations are always insignificant especially at these short wavelengths.

Sky measurements were made before and after every raster. These were always made at a fixed point close to the galaxy so that the properties of the sky could be assumed to be identical. The distance away from the galaxy was generally about 5 arcminutes. This was very much less than the distance the galaxy moved across the sky during a raster observation, which took at least 20 minutes so scanning 3 degrees of sky. The sky measurements consisted of about 40 individual 1 second measurements and provided data on the sky level to an accuracy twice that expected in a 10 second observation of the galaxy. The sky measurements also allowed estimates of the system noise to be made, so that errors could be assigned to the data.

The sky measurements are listed in Table 4.7 for the various nights on which the rastering technique was used. Those from 4 August 1982, when the variations were most pronounced, are also plotted in Fig. 4.3 against time and airmass. The sky measurements were found to be corrrelated well with time but not with airmass and so it is likely that the variations were not due to sky emission, but probably instrumental in origin.

Changes in the d.c. level of the preamplifier were not responsible because they were removed by the modulation and phase sensitive detection processes. Neither should changes in the detector preamplifier gain introduce significant shifts in the sky background level (Section 3.7.2). A change in the d.c. offset of the output from the phase sensitive detector or slight frequency shifts or gain

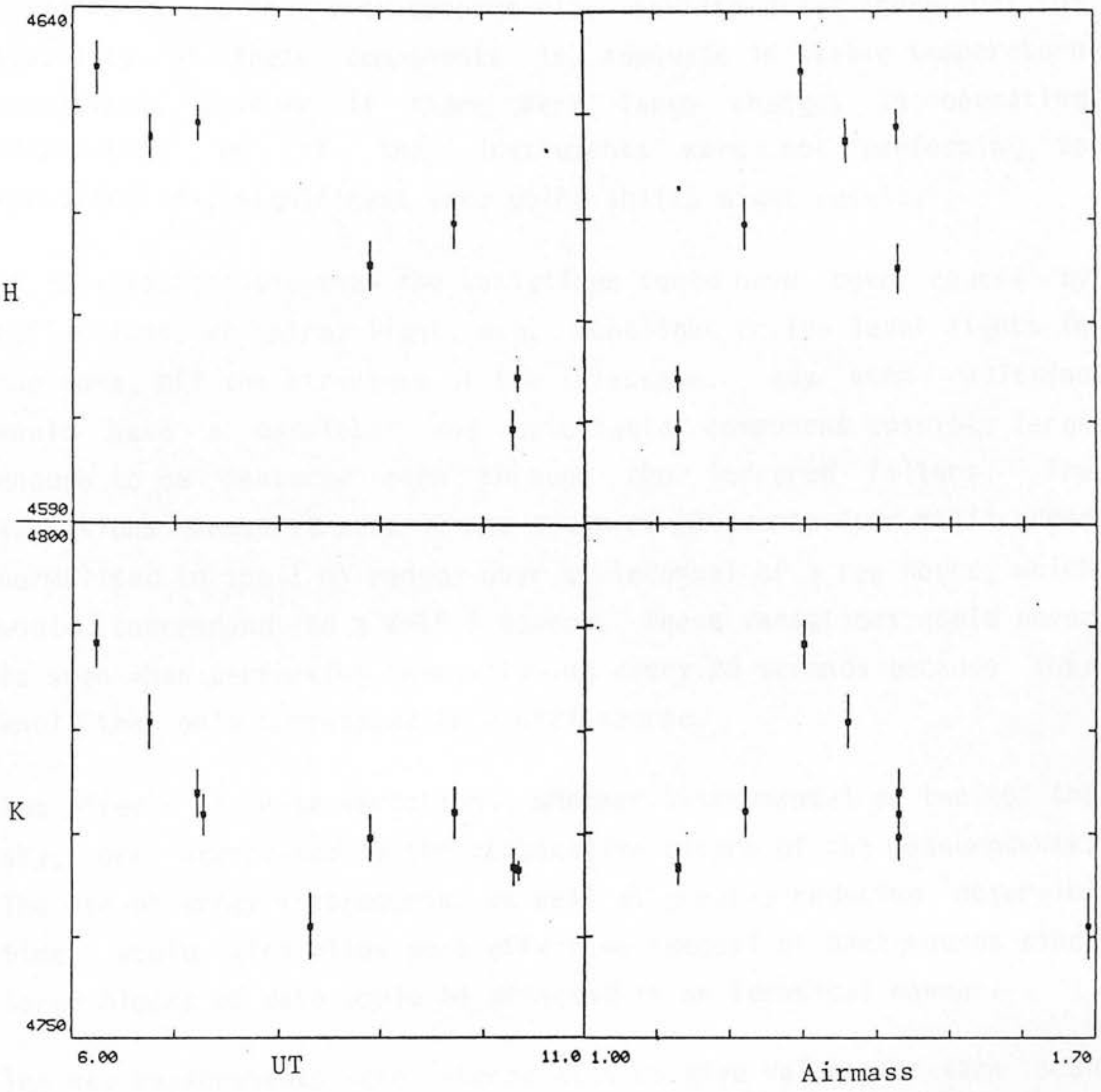


Fig. 4.3 The sky levels measured on 4 Aug. 1982 plotted against time and airmass. The error bars represent the estimated system noise in a 10 second integration.

changes in the detector preamplifier gain introduce significant shifts in the sky background level (Section 3.7.2). A change in the d.c. offset of the output from the phase sensitive detector or slight frequency shifts or gain changes in the voltage to frequency converter could have been responsible. Section 3.7.2 shows that the stability of these components is adequate in stable temperature conditions. However if there were large changes in operating temperature or if the instruments were not performing to specification, significant zero point shifts might result.

It is also possible than the variations could have been caused by reflections of stray light, e.g. moonlight or low level lights in the dome, off the structure of the telescope. Any such raditaion would have a modulated and detectable component possibly large enough to be measured even through the infrared filters. The variations measured were in the order of 10 counts (per millisecond normalised to the 1 mV range) over an interval of a few hours, which would correspond to a K=15.5 source. These variations would never be seen when performing beamswitching every 20 seconds because they would then only correspond to a K=22 source.

The effects of these variations, whether instrumental or due to the sky, are aggravated by the consecutive nature of the measurements. The use of array instruments, as well as greatly reducing observing time, would also allow more effective removal of backgrounds since large blocks of data would be affected in an identical manner.

The sky measurements were interpolated to give values for each scan line and were subtracted from it. They were assumed constant over each scan line since the variation within one line would have amounted to no more than one count, well down on the system noise level of about 5 counts. The sky levels were subtracted from the data to give two maps complemetary in sign.

Date	UT	Airmass	T	J		K		Near
11/7/82	11:11	1.24	5	4695.9	2.9	4754.0	4.1	NGC 7331
	12:05	1.11	5	4695.3	4.5	4754.4	3.5	"
	12:53	1.05	5	4695.1	4.3	4745.3	4.5	"
	13:46	1.03	5	4690.1	5.9	4744.2	10.9	"
	14:36	1.06	5	4689.3	5.7	4739.0	11.3	"
12/7/82	11:14	1.22	5	5150.6	2.7	4749.7	6.7	"
4/8/82	7:16	1.43	1	4710.0	11.3	4771.8	12.2	NGC 5907
	7:17	1.43	1	4707.9	10.6	4771.8	12.4	"
	8:18	1.68	1	4710.3	11.2	4759.6	16.1	"
	8:19	1.69	1	4712.1	8.4	4762.3	19.2	"

Date	UT	Airmass	T	H		K		Near
3/8/82	10:20	1.12	1	4634	10	4735	14	NGC 7331
	10:21	1.12	1	4631	10	4732	14	"
	10:50	1.07	1	4642	10	4734	12	"
	11:15	1.05	1	4644	12	4728	11	"
4/8/82	6:14	1.30	1	4633.7	16.8	4786.4	12.4	NGC 5907
	6:15	1.30	1	4634.6	12.4	4790.2	12.0	"
	6:45	1.36	1	4627.2	12.0	4780.7	14.7	"
	7:13	1.43	1	4628.7	10.0	4773.9	13.5	"
	8:52	1.44	1	4616.1	15.0	4770.3	9.9	NGC 7331
	8:53	1.44	1	4614.2	12.0	4768.9	14.3	"
	9:43	1.22	1	4619.4	14.3	4772.1	11.7	"
	10:18	1.13	1	4599.2	12.2	4766.8	11.0	"

Table 4.7 Sky measurements used for reduction of rastered data. For each measurement about 30 samples of T seconds duration were taken and the mean and standard deviation of these are given.

4.6.2 Combination of Data from the Two Beams

Ideally the two maps obtained with the galaxy in the "left" and "right" apertures should have been equal (apart from the polarity), but significant differences exist between them. This is illustated in Table 4.8 which shows the differences in the K magnitudes of two such maps, runs 2 and 5 from 11 July 1982. Adjacent points differ in brightness by about half a magnitude and so there are random errors due to pointing inaccuracies, but it is seen the errors are consistently positive on one side of the major axis and negative on the other. This is due to the detector seeing a significant amount of radiation from the galaxy as the chopper is in transition between its two end positions (discussed in Section 4.3). They cannot be due to pointing errors, which were checked by analysing a sample of the telescope's co-ordinates printed out at every point. The positions did not differ by more than 2 arcseconds in general and showed no systematic effects with position. This was a valid check since the positioning and printed co-ordinates are independently derived. The positioning was controlled by the TV crosshead and the printed co-ordinates were derived from the telescope encoders.

Y:	-2400	-1600	-800	0	800	1600	2400
X							
-4000	0.35	0.16	0.22	0.05	0.02	0.06	0.20
-3200	0.32	0.29	0.25	0.10	0.04	-0.02	0.06
-2400	0.33	0.20	0.12	0.04	-0.03	-0.07	-0.23
-1600	0.40	0.34	0.21	0.17	-0.07	-0.16	-0.22
-800	0.45	0.36	0.00	0.13	-0.17	-0.09	-0.18
0	0.53	0.43	0.04	0.04	-0.18	-0.31	-0.44
800	0.45	0.63	0.11	0.16	-0.01	-0.23	-0.64
1600	0.41	0.26	-0.01	-0.08	-0.08	-0.09	-0.32
2400	0.25	0.15	0.15	0.19	0.00	-0.10	-0.24
3200	0.14	0.08	0.05	0.05	-0.16	-0.18	-0.21
4000	0.10	0.14	0.02	0.02	-0.11	-0.25	-0.19

Table 4.8 The difference in magnitudes between the K signals measured in the "left" and "right" beams. The data is from runs 2 and 5 of 11 July 1982 and shows the change in sign around the major axis.

The data from the "left" and "right" rasters could simply be added, in which case the effect would be to smear out the effective beam profiles, but a better procedure is to combine the two rasters so that the measurement used is the one in which the chopper was moving away from the brighter regions of the galaxy. The effective beam profiles are then much better defined as shown in Section 4.3. On one side of the major axis therefore, one raster was used and the other raster was used on the other side of the major axis. The major axis points themselves were determined from the mean of the two rasters.

4.6.3 Estimate of errors

The sky level measurements provided an estimate of the system noise which would be expected to be the dominant source of error in the measurements of all but the brightest parts of the galaxies. The standard deviation of the sky measurements, typically 40 one second integrations, was calculated and divided by 2 or 2.8 to give the estimated error in 10 or 5 second integrations assuming that the noise was Gaussian. The noisiest sky measurement, taken close in time to the galaxy data, was used in this calculation to avoid under-estimating the error if the noise levels fluctuated.

The sky level measurements themselves were more accurately determined than the galaxy data due to the longer integration time of 40 seconds, and so the error in the sky level should not contribute significantly to the error in the difference of the galaxy and sky data. However if the sky level varies significantly with time, then there would be an extra contribution to the error of the difference. In general, apart from linear trends which are removed by the interpolation process, the time varying fluctuations appear to be less than the system noise.

In a few cases, there are significant deviations from a linear trend and it is not necessarily true that the deviations would be completey removed by the interpolation. The purpose of the two dimensional maps is largely to provide qualitative information on the distribution of the reddening and so any errors due to uncorrected variations in the sky level will not affect the conclusions. In addition the bright regions will not be affected by small variations in the sky level and so can be used quantitatively, although preference is given to the beamswitched data.

4.6.4 Presentation of Data

The one dimensional data are presented as profiles of surface brightness and colour plotted against distance in arcseconds from the nucleus. The magnitudes and colours are corrected as explained in Section 4.5 for instrumental zero points and airmass corrections where appropriate, but not for galactic extinction or redshift effects. These are always small and are discussed in Section 5.1. The magnitudes are also not normalised for the aperture size, due to the uncertainty in the effective aperture area of the Mark I cryostat observations.

Two colour diagrams, J-K against H-K, are also presented for some of the data. The diagrams are in this form rather than the more usual J-H against H-K form because the independent data are J-K and H-K. There can be an error due to the J-K and H-K data being measured at slightly different positions, which would be compounded if the J-H colours were plotted against H-K. However comparison of the K magnitudes obtained with the J-K and H-K measurements shows that the pointing errors are small, about 2 arcseconds or less, and so this error is not likely to be significant.

The two dimensional data are presented as contour plots, with the colour contours being overlaid as dotted lines on the magnitude contours. These maps are produced by the standard Starlink software (ASPIC, Hartley 1983), which applies bi-linear interpolation to the data. The magnitudes of the data obtained with the Mark II cryostat, NGC 7331 and NGC 5907, are normalised by the aperture size, which is assumed to be 10 arcseconds in diameter, to derive surface brightnesses. These should be accurate to 25 percent or better, taking into account a probable error of 10 percent in the estimate of the effective aperture diameter and the uncertainties in position produced by windshake.

The two dimensional data are also presented as colour photographs taken from the Starlink Args screen. These images, produced by linearly interpolating the data and applying the Starlink package TWOTONE, display both the intensity and the colour of the data. The intensities are reproduced on a logarithmic scale to increase the dynamic scale. The colours are produced by combining a blue image, obtained from the J or H data, with a red image obtained from the K data and so the red areas on the photographs represent areas in the galaxy which are red in J-K (or H-K). The colours are enhanced by factors of 2 for NGC 5907 and 4 for NGC 4565 and NGC 7331 to show the colour variations more clearly. This enhancement was carried out on the data after it had been transformed to the HSI plane (hue, intensity and saturation) where the saturation was increased before transformation back to the RGB (red, green, blue) plane for display on the Args.

All the data obtained are shown on the Args images but in the outer areas the colours are affected by noise. On the contour maps, the colour contours are drawn for those parts of the map where the K surface brightness is greater than 17.5 magnitudes per square arcsecond (corresponding to K = 12.8 in the 10 arcsecond beam) which is the level at which the colour errors approach 10 percent. The brightness contours are drawn at intervals of 0.5 magnitudes and the colour contours at intervals of 0.05 for H-K and 0.1 for J-K.

4.7 Features of the Data

4.7.1 Airmass Corrections

In general no airmass corrections have been applied or found to be significant. At near infrared wavelengths airmass corrections are generally small and non-repeatable from night to night. Extinction variations depend largely on the water vapour content of the atmosphere with the J band being most severely affected due to a water absorption feature at the short wavelength end of the band (La Rocca 1978). Objects and observing times were chosen so ranges of airmasses were usually small (less than 0.2) and standard stars were chosen of comparable airmasses.

The observations made with the Mark I cryostat suffered from instrumental zero point shifts which were removed by interpolating between measurements of the nuclei of the galaxies. This also removed the airmass effects so that airmass corrections were neither measurable nor necessary. Any systematic error due to airmass could be expected to be small because of the limited airmass range.

The Mark II cryostat observations did not need to be corrected in this way. The zero point shifts observed were generally small (±0.02 magnitudes) and comparable to the statistical uncertainty. These shifts were often in the correct sense to be attributed to airmass but because of their small size no corrections were made.

On the night of 4 August, however, NGC 7331 was observed over a large range of airmass (1.61 to 1.04) and systematic J-K zero point shifts, using the nucleus as a standard, were observed. The nucleus was also observed 5 times during the H-K measurements, which covered a smaller airmass range (1.03 to 1.13). The zero point shifts are plotted in Fig 4.4 against airmass, showing a strong correlation (Table 5.8) for the J-K and J zero points. There was no systematic variation with time. The correlations for the J and J-K zero point shifts, but not the K, are significant using the t test at the 10 percent level.

The J-K colours were therefore corrected for airmass, but not the H-K or K data. This was the only night on which a significant correction was found to be necessary, due presumably to the large airmass range covered and high water vapour content in the atmosphere affecting the J band. The size of the effect, 0.16 mag's/airmass, is larger than would generally be expected at the lower altitude Kitt Peak site, but near infrared airmass corrections depend very much on the prevailing conditions (Manduca and Bell 1979)

	Mag./airmass	Zero point for airmass = 1	Zero point for airmass = 0	Correlation coefficient
J	0.208	-18.548	-18.756	0.985
K	0.048	-18.034	-18.082	0.689
J-K	0.160	-0.513	-0.673	0.993

Table 5.8 Calculation of airmass corrections using the measurements of the nucleus of NGC 7331 on 5 Aug. 1982. The two points at airmass 1.13 were combined.

4.7.2 Self-consistency of Data

The K magnitude was measured for all points both with the J filter and the H filter. When both J and H measurements were made, therefore, the two sets of K data provide a check on the data. Factors which could cause these data to differ, by amounts in excess of the system noise, include

(a) pointing errors due to windshake and guiding errors

(b) background variations for the non-beamswitched data

(c) differences in beam profiles due to focus effects in the Mark I cryostat

(d) differences in effective beam profiles due to chopper variations, especially for off-axis data taken on different nights.

The differences between the K magnitudes from the pairs (J-H and J-K) of various data set have been calculated. The rms values of the differences are listed in Table 5.9. Excluded from the statistics are those data for which the estimated error in either K magnitude is greater than 5 percent. Also excluded are those points which graze the nucleus, i.e. are not more than one beamwidth away. These points tend to have high values for the differences since the brightness gradients are very large.

Date	Scan	rms difference
26/8/81	NGC 7331 NS	0.06
4/8/82	NGC 7331 minor axis	0.06
4/8/82	NGC 5907 map	0.14
11/7/82, 3/8/82	NGC 7331 small map	0.16
11/7/82, 4/8/82	NGC 7331 large map	0.15
5/8/82	NGC 7331 major axis	0.06

Table 5.9 The rms difference in K magnitudes between pairs of data sets measured with J and with H. The points included are those for which the statistical errror is less than 5 percent; points grazing the nucleus are excluded.

The scan data have relatively low values for the the differences in K magnitudes and are consistent with small pointing errors (next Section). The small value for differences in the N/S scan of NGC 7331, measured with the Mark I cryostat, shows that the interpolation procedure for the correction of zero point drifts (Section 4.5.1).

The map data have significantly larger rms differences than the scan data. The NGC 7331 data were obtained on differents nights and so differences in the effective beam profile may be responsible. NGC 5907, being very edge-on, has brightness gradients in the order of 0.2 mag's/arcsec and so the differences are consistent with small pointing errors. Since quantitative infromation for NGC 7331 is drawn primarily from the one dimensional scans, the slight lack of consistency in the data from the maps is not important.

4.7.3 Pointing Errors

The pointing errors associated with the data may be obtained directly from the printed positions for the data of July and August 1982, when accurate telescope position encoders were in operation.

This cannot be done for the earlier data because the telescope position encoders in use suffered from cyclic errors. These repeated every 128 encoder bits (about 2 arcminutes) and even after computer correction could reach 5 arcseconds.

The positions for the July and August 1982 data were found to agree to about 3 arcseconds rms. The encoders were implemented to give a resolution of 1.2 arcseconds and so the measured errors were only slightly in excess of this. When the positioning was effected by use of offsetting the TV camera the encoders provided an independent check on the positioning; the small value of the pointing error indicated that using the encoders directly should not produce significant pointing errors with a 10 arcsecond aperture.

The pointing errors can also be obtained for all the data by comparing the K measurements taken in the J-K and H-K observations. By comparing the differences between these with measurements with the brightness gradient, the pointing error in the direction of the maximum gradient may be estimated. If the brightness gradient in the orthogonal direction is smaller, then multiplying this value by root 2 gives an estimate of the total pointing error in all directions since there is no preferred direction for the error. If the gradient in the orthogonal direction is comparable, as along the major axis of a highly inclined galaxy, then this factor does not need to be applied.

This gives the difference in position of the effective apertures and so takes into account variations in the chopper performance, which could make the effective positions different even if the telescope were pointing in exactly the same position for the two measurements. The points adjacent to the nucleus will be more sensitive to small errors and so are not used in these calculations. When this procedure is carried out it is found that the effective pointing errors are less than 2 arcseconds for the 10 arcsecond aperture, and about 1 to 1.5 arcseconds for the 5 arcsecond aperture, when greater care was necessary in the telescope positioning and guiding.

These small but inevitable inconsistencies between the J-K and H-K data will introduce scatter into the two colour diagrams (Chapter 6). The construction of a three channel system would eliminate this effect in the same way that the two channel system eliminates scatter in the individual colours.

4.8 Technical Conclusions

(a) Accurate colour data were obtained with both the Mark I and Mark II cryostats. The Mark II system provided good stability in both colour and intensity measurements; colour repeatability was ±1 percent and intensity stability was ±3 percent.

(b) The simultaneity of measurement in two channels made possible the accurate measurement of colours in regions of large brightness gradients (up to 0.2 mag's/arcsec), even in the presence of windshake and chopper variations.

(c) Imperfections and variations with time of the chopper waveform caused changes in the effective beam profile when measuring extended objects. The quality of data could be improved either by improving the chopper waveform or by gating out the transitional region of the chop cycle.

(d) The omission of beamswitching when mapping in the near infrared should not introduce additional errors. Poor stability of the electronics in the data acquisition chain, or possibly stray light, did however cause background variations. In the worst case, these were equivalent to a change in sky level (180 arcseconds chop) over an hour corresponding to a source of K magnitude 15.5. The variations tended to be linear with time, but not with airmass. Some random fluctuations of rather smaller size were also observed.

(e) The use of larger detectors in the Mark II cryostat enabled flat beam profiles to be obtained at larger apertures. Although the larger detectors were of very high impedance, excess noise was introduced due to their larger capacitiance amplifying the JFET voltage noise. Temperature control of the JFETs may help reduce this.

(f) Variations in the chopper performance and small pointing differences can cause consecutive measurements, especially if obtained on different nights, to refer to slightly different positions. A three channel system, similar to the 2 channel system used here, would allow more accurate data to be obtained for use on two colour diagrams.

(g) Mapping, by measuring points consecutively, can be subject to instrumental background shifts and effective beam profile variations. The use of arrays could help alleviate these problems as well as allow data to be obtained more efficiently.

5. The Data

5.1 The Galaxies Observed

The galaxies selected for observation are normal undisturbed spiral galaxies of large angular size and generally highly inclined. So that foreground stars did not present serious problems, objects near the galactic plane were not observed. Galaxies of size too large to allow chopping onto empty sky were also rejected. The chosen galaxies and their adopted parameters are listed in Table 5.1.

Galaxy	Type	R	D	i	PA
		'	Mpc	°	°
NGC 4216	SAB(s)b:	8.3[2]	20[8]	83[5]	20[9]
NGC 4565	SA(s)b I	16.2[2]	12[4]	86[5]	135[3]
NGC 5907	Sc	12.3[2]	15[2]	87[5]	155[1]
NGC 7331	Sb(rs) I-II	10.7[2]	20[2]	72[6]	167[7]
NGC 7814	SA(s)ab	6.3[2]	23[2]	90[9]	135[9]

Table 5.1 The galaxies chosen for observation. R is the angular diameter of the major axis, D is the adopted distance, i is the inclination of the galaxy to the plane of the sky and PA is the position of the major axis (E of N). The sources are

1) van der Kruit and Searle (1981a)
2) de Vaulouleurs et al (1976), H = 55 km/s/Mpc
3) Huchtmeier et al (1980). Also 135 (Danver 1942), 136 (Nilson 1973).
4) Jensen and Thuan (1982)
5) de Vaucouleurs (1958)
6) Boroson (1981)
7) Adopted value. Also 167 (Bosma 1981a) and 171 (Boroson 1981).
8) Sandage and Tamann (1975a)
9) Hamabe and Okamura (1981)
10) Adopted value. Also 136 (Hamabe and Okamura 1981).

Corrections to observed colours of galaxies are generally required to correct for galactic absorption and recessional velocity. The galactic absorption can be derived from the galactic latitude b (Sandage 1973). The visual extinction A_V derived from b, is tabulated for these galaxies by Sandage and Tamann (1981); the extinction is taken to be zero for $|b| > 50$. It only provides approximate estimates for the extinction to individual objects but indicates likely sizes for the corrections. The colour corrections E(J-K) and E(H-K), shown in Table 5.2, are derived from A_V using the colour ratios of (Koorneef 1983b) i.e.

$$E(J-K)/A_V = 0.17$$
$$E(H-K)/A_V = 0.06$$

K corrections allow for the shift in pass bands relative to the spectrum on account of the redshift z. The corrections to the colours (Table 5.2) are calculated using the relationships used by Griersmith et al (1982) i.e.

$$K(J-K)/z = 2.5$$
$$K(H-K)/z = 2.0$$

A larger value for K(H-K)/z (3.5) is given by Persson et al (1979) but is not significant since the resultant infrared colour corrections are small.

Galaxy	z	K(J-K)	K(H-K)	A(V)	E(J-K)	E(H-K)
NGC 4216	.000	0.00	0.00	0.00	0.00	0.00
NGC 4565	.005	0.01	0.01	0.00	0.00	0.00
NGC 5907	.002	0.01	0.00	0.00	0.00	0.00
NGC 7331	.003	0.01	0.01	0.24	0.04	0.01
NGC 7814	.003	0.01	0.01	0.05	0.01	0.00

Table 5.2 The colour corrections for the galaxies due to redshift and galactic absorption. The redshifts z and the galactic absorption A(V) are from Sandage and Tamann (1981). The corrections for redshift (K) and for absorption (E) are derived as described in the text.

It is seen from Table 5.2 that the only correction which could be considered significant is that to the J-K colour of NGC 7331 due to galactic absorption. NGC 7331 has the smallest galactic latitude (|b| = 21) of the galaxies observed. Since the correction is relatively small and it is primarily the colour variations within the galaxies which are of interest, no correction will be applied to the observed colours.

The details of the one dimensional scans are given in Table 5.3, which also gives details of the map of NGC 4565 made in the beamswitching mode. Table 5.4 lists the maps made in the rastering mode. The data themselves are presented in graphical form in the following sections and are tabulated in Appendix A. Fig. 5.1 shows the visual appearance of the galaxies in reproductions from the B plates of the Palomar Observatory Sky Survey. Also shown are the positions observed in the infrared.

5.2 NGC 7331

5.2.1 The Data

NGC 7331 was observed with both versions of the cryostat. The earlier data (Fig. 5.2), with an aperture of 5 arcseconds, consist of scans in the N/S direction (chopping 50 arcseconds E/W) and in the E/W direction (chopping 90 arcseconds N/S). A star in the reference beam was avoided by adding a 40 arcsecond E/W displacement to the chop for some points in the E/W scan.

Scans along the major and minor axes and two dimensional maps were made with the 10 arcsecond aperture of the Mark II cryostat, chopping about 180 arcseconds E/W. No differences in pointing between the two methods of positioning (TV crosshead and telescope encoders) were noticed. No standard star was observed on 5 August since the colours of the nucleus had been well determined the previous night. Airmass corrections (Section 4.7.1) were found to be necessary on this night alone due to the atmospheric conditions and the large range of airmasses involved. A star close to the major axis about 85 arcseconds NW of the nucleus affected the 80 and 90 arcsecond points which have been omitted from the data.

The major axis was observed twice, sampled every 10 arcseconds to the SE and NW (Fig. 5.3) and every 5 arcseconds to the SE

Date	Object	Aperture	Spacing	Filters	Airmass	Length and direction
26/8/81	NGC 7331	5	4.8	J-K, H-K	1.03 to 1.54	67 S, 57 N
"	"	5	4.8	J-K	"	38 E, 38 W
31/8/81	NGC 7814	5	5.1	J-K	1.01 to 1.09	30 NW, 40 SE (major axis)
"	"	5	5.1	J-K	1.08 to 1.30	27 NE, 20 SW (minor axis)
29/1/82	NGC 4216	4	3.6	J-K	1.01 to 1.13	39 N, 39 S, 18 E, 25 W
30/1/82	NGC 4565	4	3.0	J-K	1.06 to 1.34	55 SE (major axis)
31/1/82	"	4	3.0	J-K	1.01 to 1.43	42 NW (major axis)
"	"	4	3.0	J-K	"	21 NE, 18 SW (minor axis)
1/2/82	"	4	4.3	J-K,H-K	1.26 to 1.45	30 NE, 26 SW (PP, 9" SE)
"	"	4	4.3	J-K	1.01 to 1.45	51 x 26 map
12/7/82	NGC 7331	10	4.8	J-K	1.03 to 1.23	128 SE (major axis)
4/8/82	NGC 7331	10	5.0	J-K, H-K	1.03 to 1.32	60 NE, 60 SW (minor axis)
5/8/82	NGC 7331	10	10	J-K, H-K	1.03 to 1.61	120 NW, 130 SE (major axis)

Table 5.3 One dimensional scan data. The apertures, spacing between points and scan lengths are given in arcseconds. The scan directions are relative to the axes as defined in Table 5.1. Where no axes are indicated, the scans were made due north, south east or west. The Mark II cryostat was used for the 10 arcsecond aperture observations; the smaller apertures were with the Mark I cryostat.

Date	Object	Run	Area	Beam	Filter	Spacing	Pos	UT	Airmass	Sec's	Noise	
											J/H	K
11/7/82	NGC 7331	2	48 x 29	R	J-K	4.8	TV	11:42 to 12:02	1.16 to 1.12	5	4.5	4.5
"	"	5	48 x 29	L	J-K	4.8	TV	12:08 to 12:28	1.11 to 1.08	5	"	"
"	"	8	120 x 80	R	J-K	10	TV	12:58 to 13:42	1.05 to 1.03	10	4.2	8.0
"	"	10	120 x 80	L	J-K	10	TV	13:50 to 14:30	1.04 to 1.05	10	"	"
3/8/82	"	9	48 x 29	R	H-K	4.8	TV	10:24 to 10:44	1.13 to 1.10	10	4.0	5.0
"	"	11	48 x 29	L	H-K	4.8	TV	10:53 to 11:14	1.09 to 1.07	10	"	"
4/8/82	NGC 5907	3	120 x 40	L	H-K	10	TV	6:19 to 6:43	1.31 to 1.36	10	5.3	4.6
"	"	4	120 x 40	R	H-K	10	Enc	6:49 to 7:09	1.37 to 1.42	10	"	"
"	"	8	120 x 40	R	J-K	10	Enc	7:36 to 7:56	1.50 to 1.56	10	3.6	6.1
"	"	9	120 x 40	L	J-K	10	Enc	7:57 to 8:12	1.58 to 1.63	10	"	"
"	NGC 7331	13	120 x 80	R	H-K	10	Enc	9:01 to 9:43	1.37 to 1.22	10	4.4	4.5
"	"	14	120 x 80	L	H-K	10	Enc	9:47 to 10:16	1.21 to 1.15	10	"	"

Table 5.4 The maps made in rastering mode (without beamswitching at each point). The area mapped and the spacing between points are given in arcseconds. "Pos" refers to the method of positioning used, either using the TV or the telescope encoders. "Secs" is the integration time per point in seconds. The noise in each channel was estimated from sky readings before and after the measurements and is instrumental counts/ms normalised to 1 mV sensitivity.

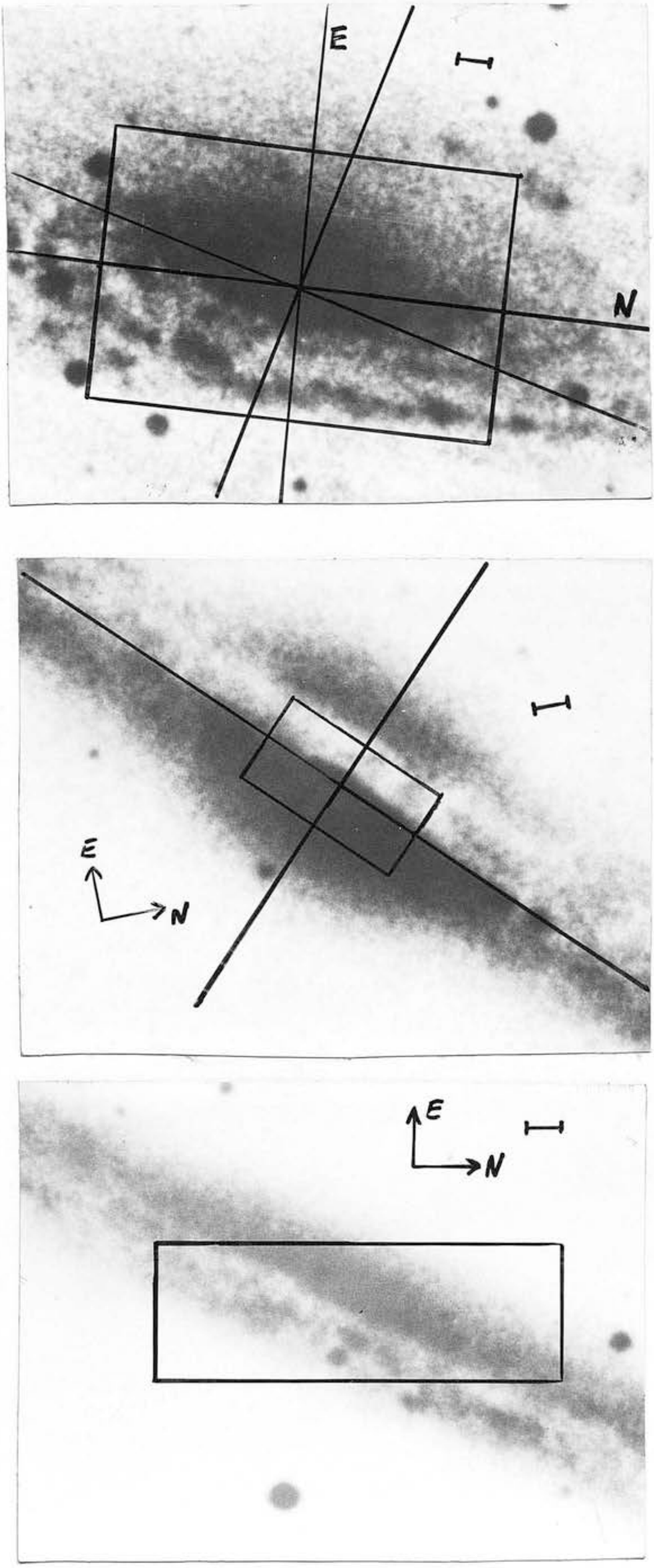
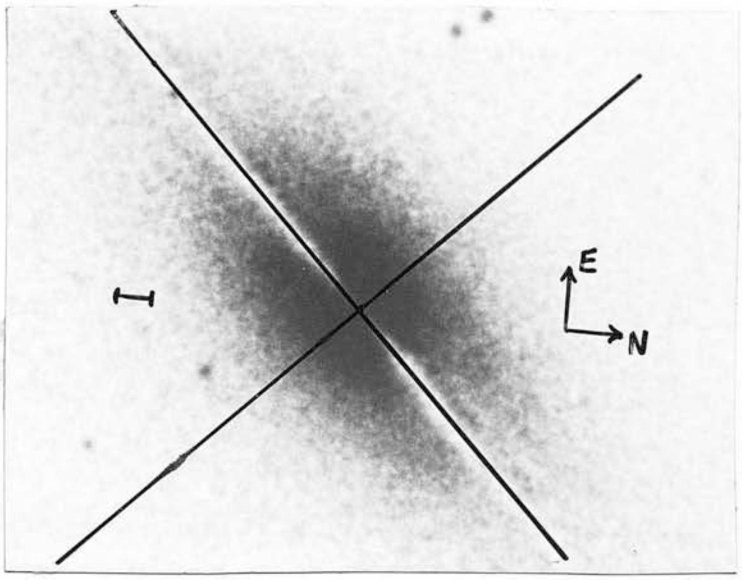


Fig. 5.1 (a)-(c) Reproductions from the POSS blue plates showing the observed positions in (a) NGC 7331, (b) NGC 4565 and (c) NGC 5907. Further details are given overleaf.



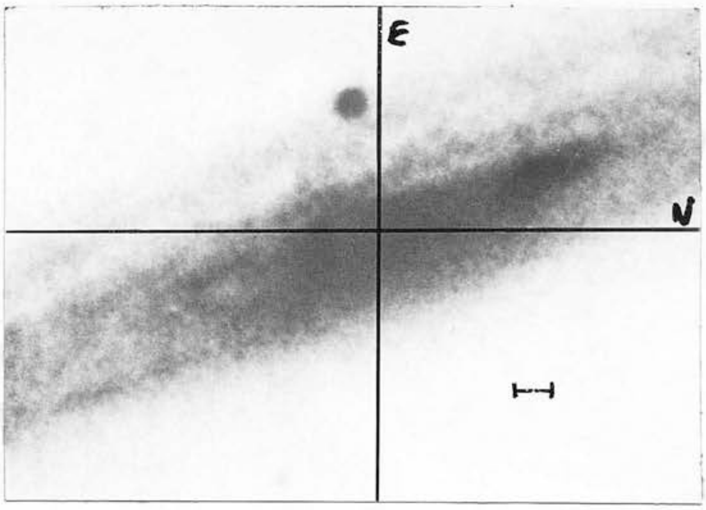


Fig. 5.1 (d)-(e) Reproductions from the Palomar Observatory Sky Survey blue plates showing the observed positions in (d) NGC 7814 and (e) NGC 4216. The scale is 2 arcseconds/mm.

The lines represent the scan directions and the tick marks represent intervals of 10 arcsecond. The areas enclosed by boxes represent the areas mapped in two dimensions. All positions were measured relative to nucleus of the galaxies as peaked up in the K filter.

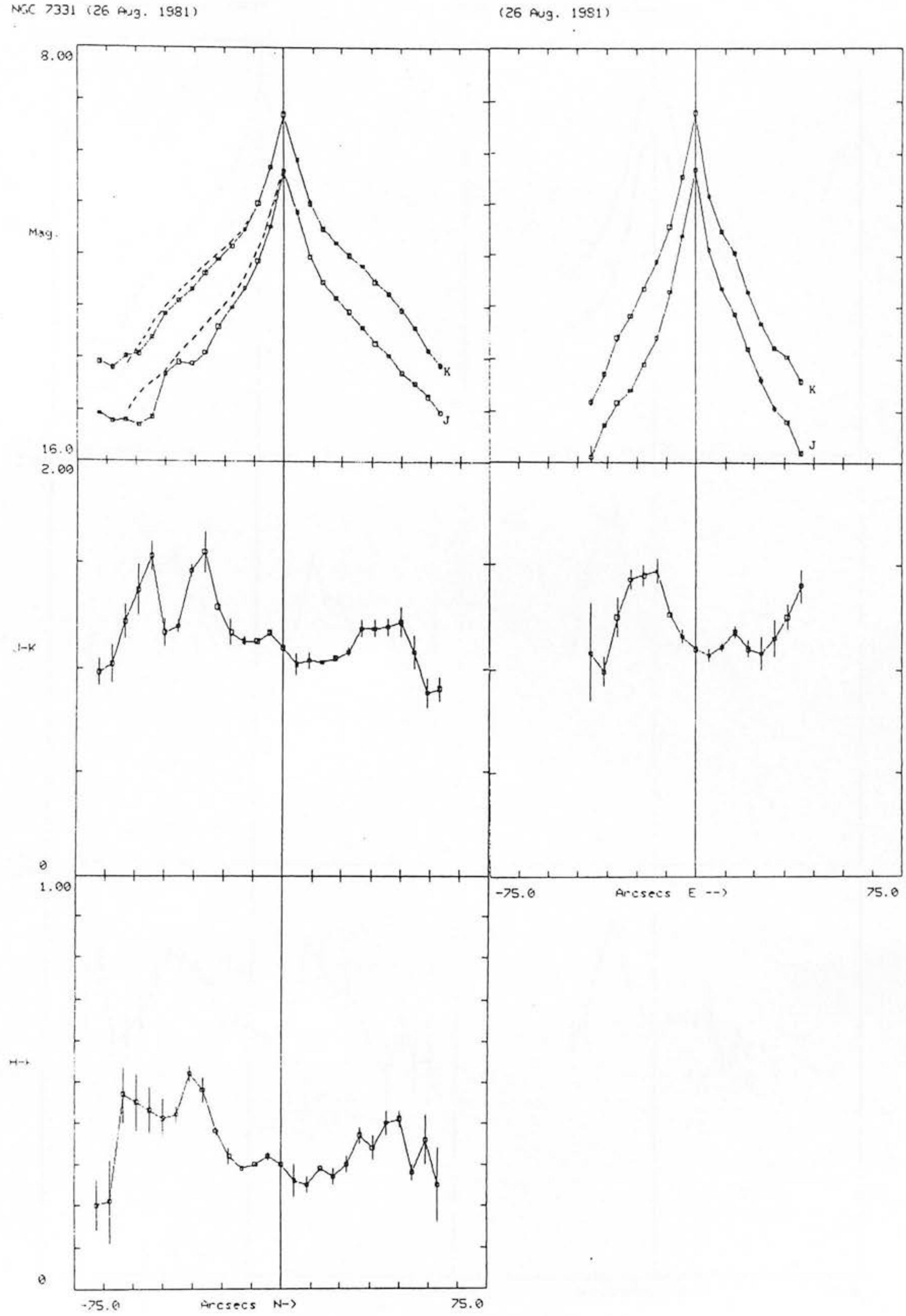


Fig. 5.2 The N/S and E/W profiles of NGC 7331 measured in a 5 arc-aperture. The dotted line on the profile to the south is the north profile and indicates that the reddening in J-K and H-K is caused by reductions in the J and H light.

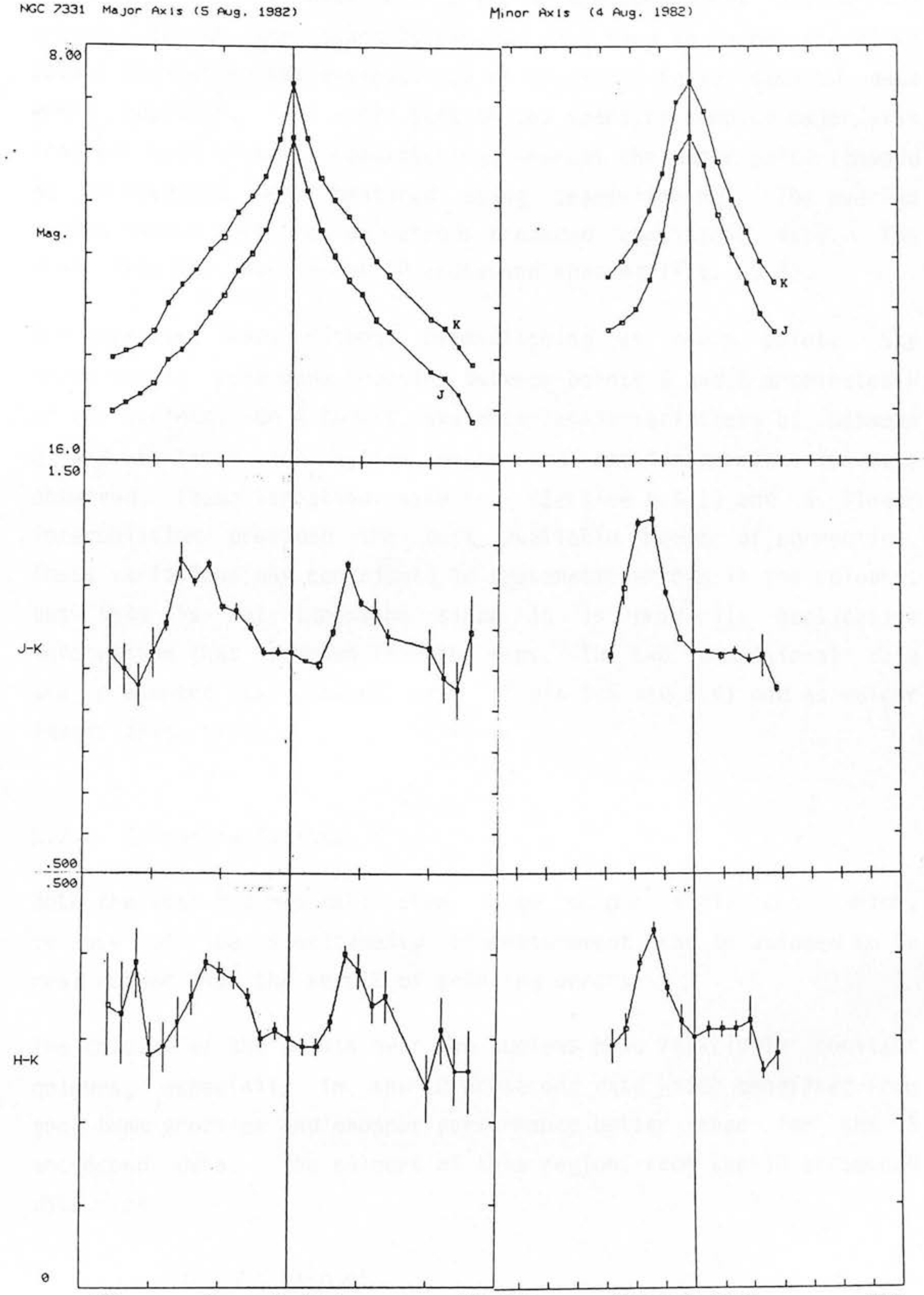


Fig. 5.3 The major and minor axis profiles of NGC 7331 measured in a 10 arcsecond aperture.Omitted from the major axis profile are two points possibly contaminated by starlight.

(Fig. 5.4). The two data sets to the SE are consistent within the errors, though the sparsely sampled data tend to favour the bluer points indicating the desirability of obtaining fully sampled data where possible. The inner part of the sparsley sampled major axis scan was made without beamswitching, whereas the outer parts (beyond 62 arcseconds) were measured using beamswitching. The overlap region showed that the two methods produced consistent data. The minor axis was observed at 10 arcsecond spacing (Fig. 5.3).

The maps were made without beamswitching at every point. Sky measurements were made chopping between points 5 and 8 arcminutes W of the nucleus. On 4 August, sky measurement variations of between 2 and 5 times the system noise and not linear with time were observed. These variations were real (Section 4.6.1) and a linear interpolation provided the best available means of correction. These variations may contribute to systematic errors in the colours, but this is not important since it is primarily qualitative information that is drawn from the maps. The two dimensional data are presented as contour maps (Fig's 5.5 and 5.6) and as colour images (Fig. 5.7).

5.2.2 Colour Variations

Both the scan and map data show large colour variations, which, because of the simultaneity of measurement, can be assumed to be real rather than the result of pointing errors.

The colours of the points near the nucleus have relatively constant colours, especially in the 10 arcsecond data which benefited from good beam profiles and chopper performance better than for the 5 arcsecond data. The colours of this region, from the 10 arcsecond data, are

$$J-K = 1.04\pm0.01$$
$$H-K = 0.31\pm0.01$$

The errors are the statistical errors. As discussed in Section

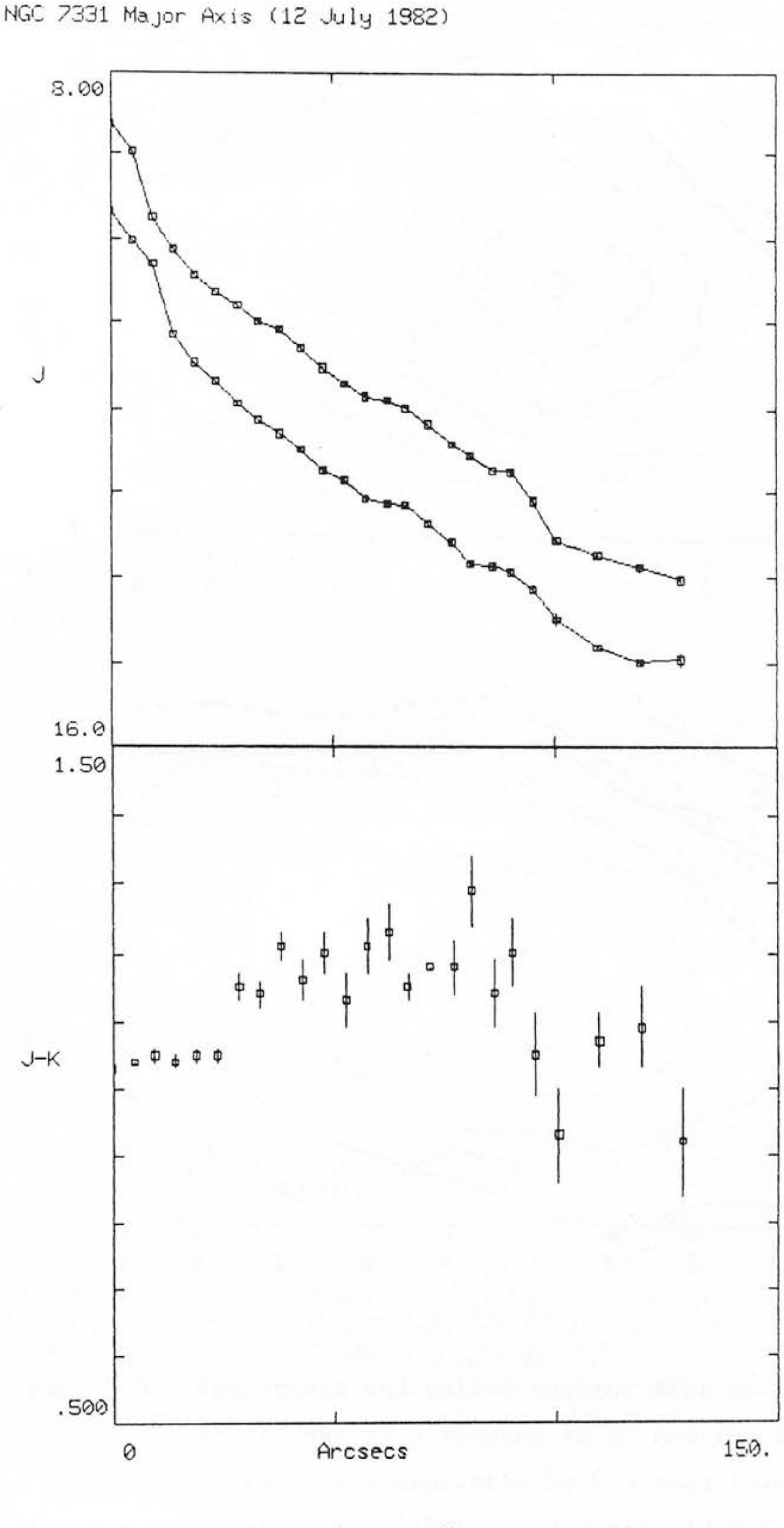


Fig. 5.4 The major axis profile of NGC 7331 measured in a 10 arcsecond aperture with a point spacing of 5 arcseconds. The scan was made to the southeast of the nucleus.

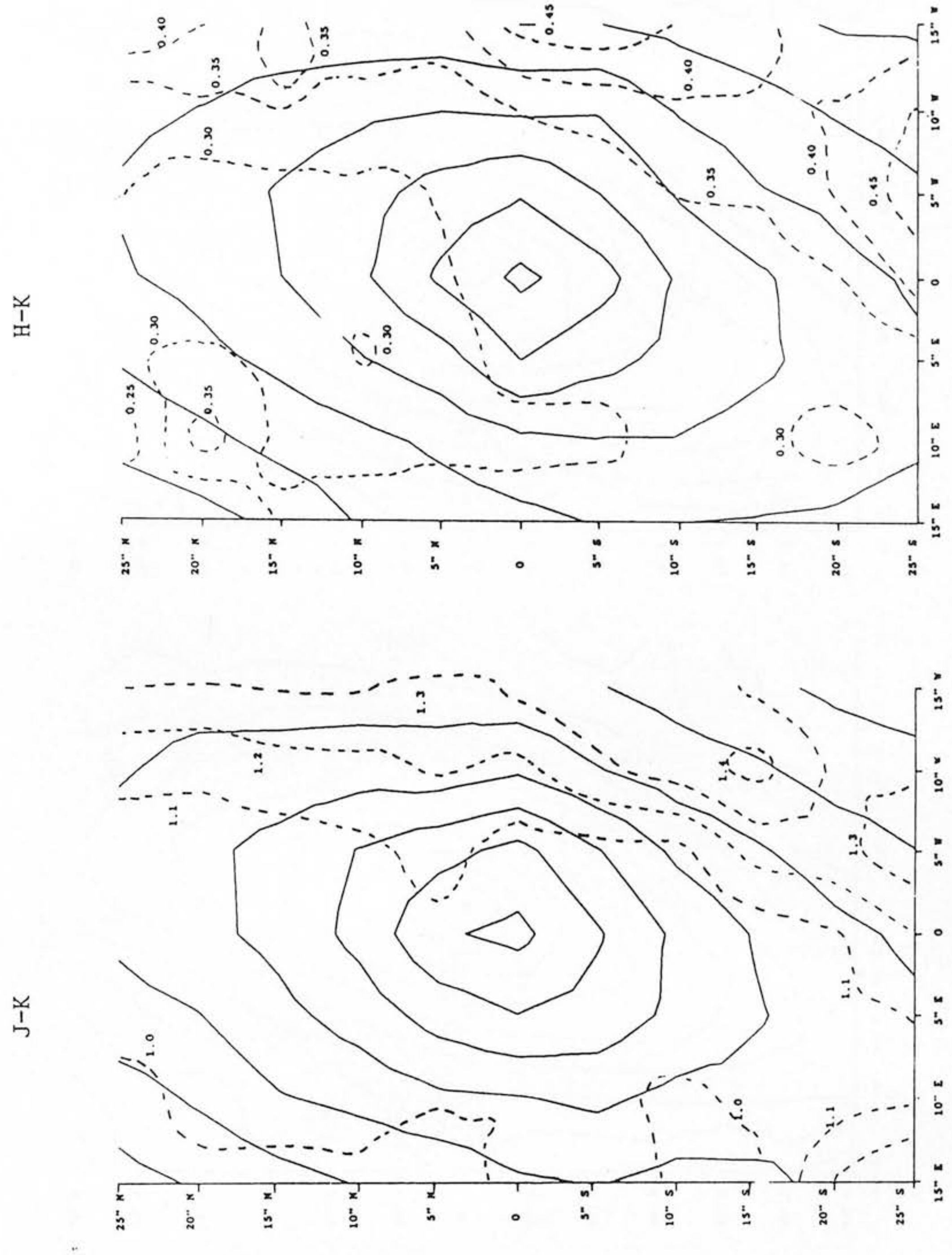


Fig. 5.5 Brightness and colour contour maps of NGC 7331 (small area). The grid spacing is 5" and the brightness contours are separated by 0.5 magnitudes.

H-K (top): brightest contour is 13.5 K mag's arcsecond^{-2}; colour contours are H-K = 0.25, 0.35, 0.40 and 0.45.

J-K (bottom): brightest contour is 13.5 K mag's arcsecond^{-2}; the colour contours are J-K = 1.0, 1.1, 1.2, 1.3 and 1.4.

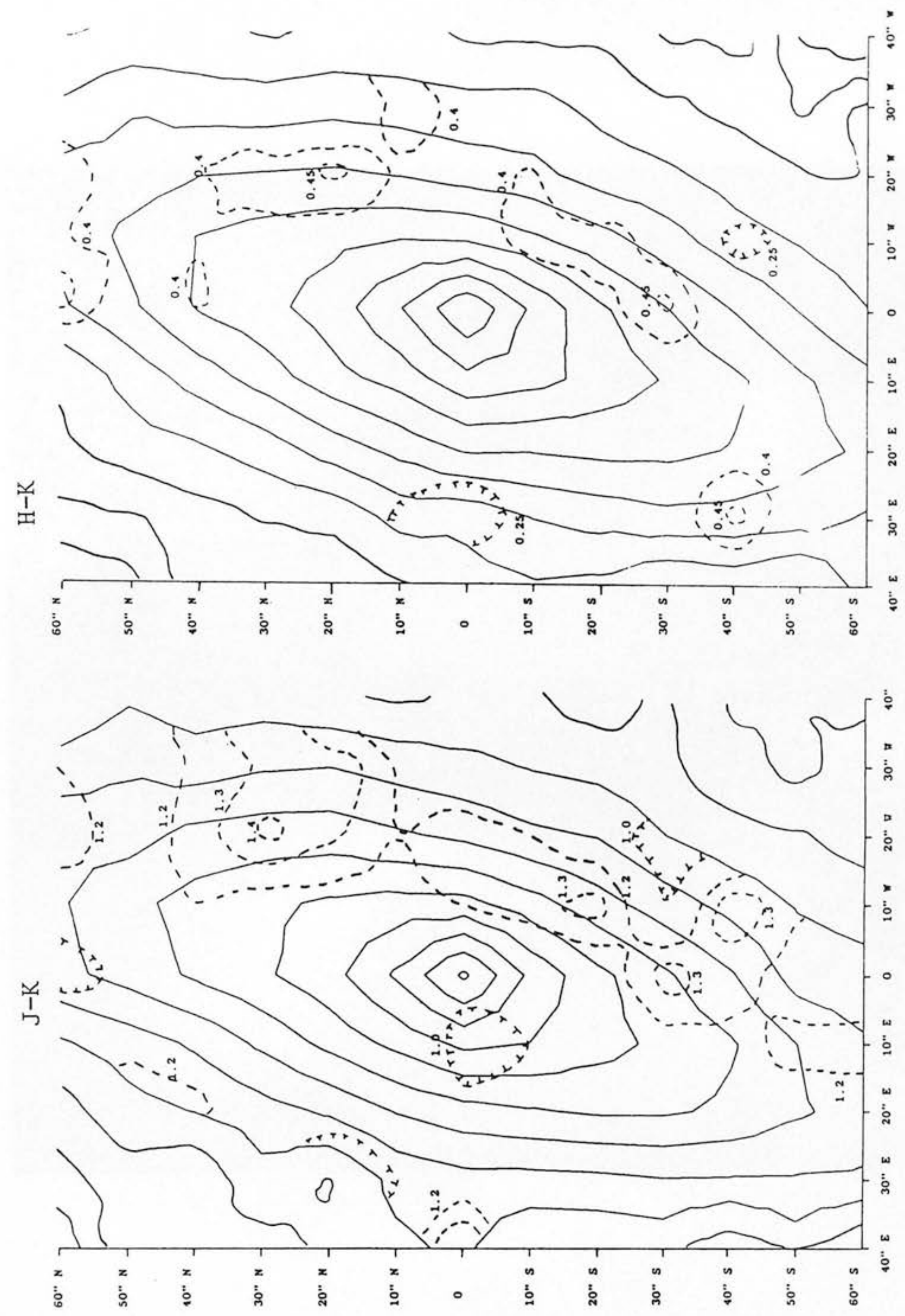


Fig. 5.6 Brightness and colour contour maps of NGC 7331 (large area). The grid spacing is 10" and the brightness contours are separated by 0.5 magnitudes.
H-K (top): brightest contour is 13.5 K mag's arcsecond^{-2}; colour contours are H-K=0.25, 0.40 and 0.45.
J-K (bottom): brightest contour is 13.0 K mag's arcsecond^{-2}; colour contours are J-K=1.0, 1.2, 1.3 and 1.4.

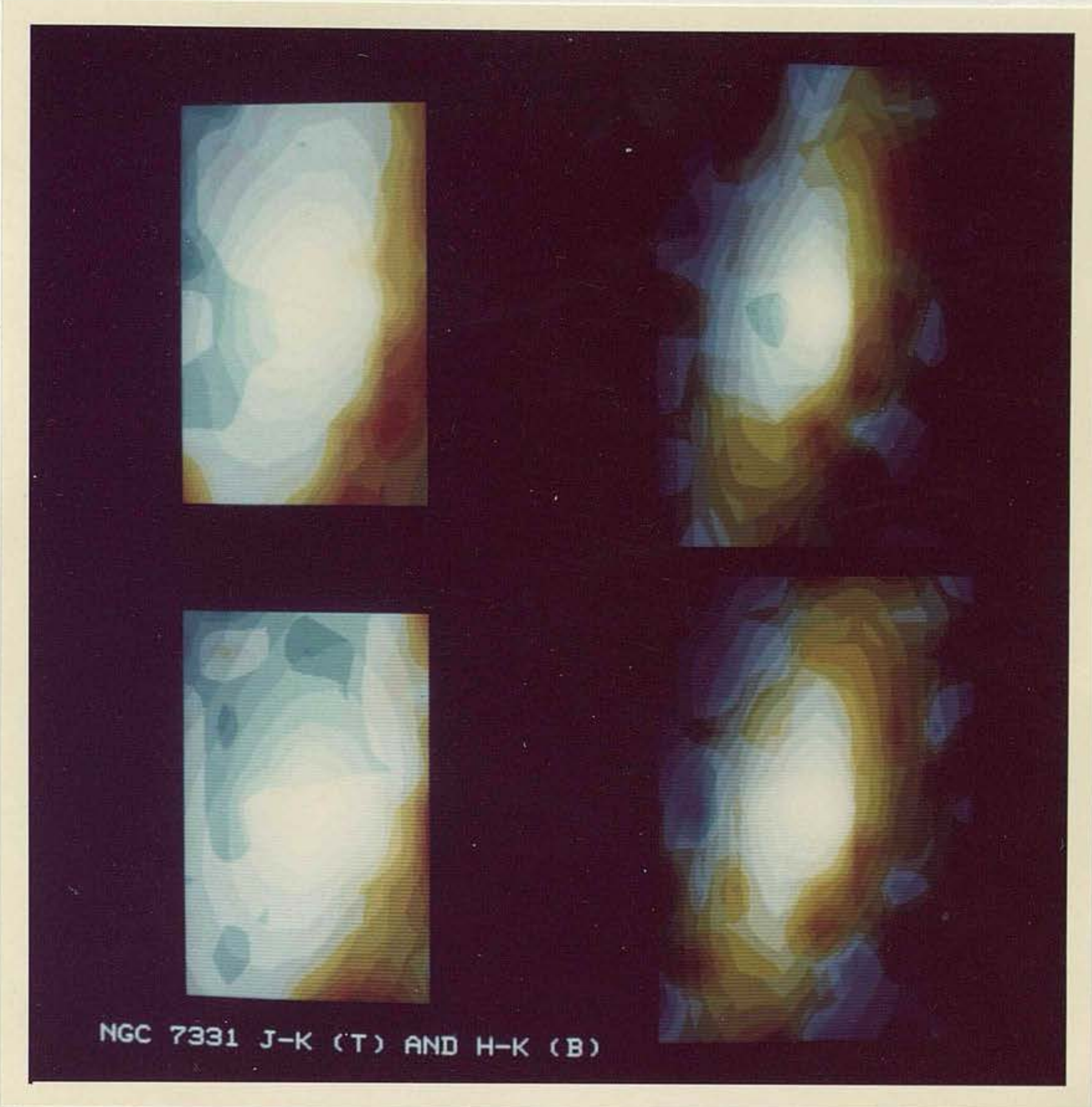


Fig. 5.7 Starlink images of the two dimensional NGC 7331 data. The images on the left represent the 50" x 30" inner region, shown as contour maps in Fig. 5.5. The imaages on the right correspond to the 120" x 80" region (Fig. 5.6). In both cases the J-K map is on the top with the H-K map on the bottom. The brightnesses are on a logarithmic scale and the colours are enhanced with red colours on the photographs corresponding to red regions in J-K and H-K.

4.4.1 systematic errors may exist in comparisons with other data.

Comparison with the two colour diagram of Fig. 2.2 indicates that these colours are typical of bulges of spiral galaxies. Comparisons between the infrared data and the corresponding positions on the visual photograph (Fig. 5.1) show that the red regions correspond very closely with the visually obscured regions. It is clear therefore that the dominant source of reddening is obscuration due to dust. This conclusion can be confirmed by inspecting some of the data more closely.

The J N/S profile is plotted in Fig. 5.2 and it is apparent that the profile to the south is much less smooth than that to the north. The dotted line drawn on the south side represents the profile of the north side and demonstrates that the undulations represent a J deficit rather than a K excess. At these points there is some evidence that the K signal is also smaller on the south side but by a lesser amount than the J signal. The regions which are deficient at J are therefore also the red regions on the J-K profile. A similar but smaller effect is also seen in the H-K data. The deficit becomes greater with shortening wavelength further suggesting that the cause is extinction by dust.

The location of the data points on the two colour diagram is considered in Chapter 6. However as further confirmation of the extinction interpretation for the colours, selected points from the N/S scan are plotted on the two colour diagram of Fig. 5.8. The points chosen are the nucleus, the two reddest points and two representative bulge points 10 arcseconds either side of the nucleus. The arrow drawn from the nucleus point represents the extinction vector for an A_V of 2(Koorneef 1983b). The points lie very close to this vector suggesting that extinction, having similar properties to those observed in the Milky Way, is responsible for the colour variations.

Chapter 2 shows that extinction cannot easily be distinguished on the two colour diagram from some other forms of reddening. The black body locus (Fig. 2.2) lies in approximately the same

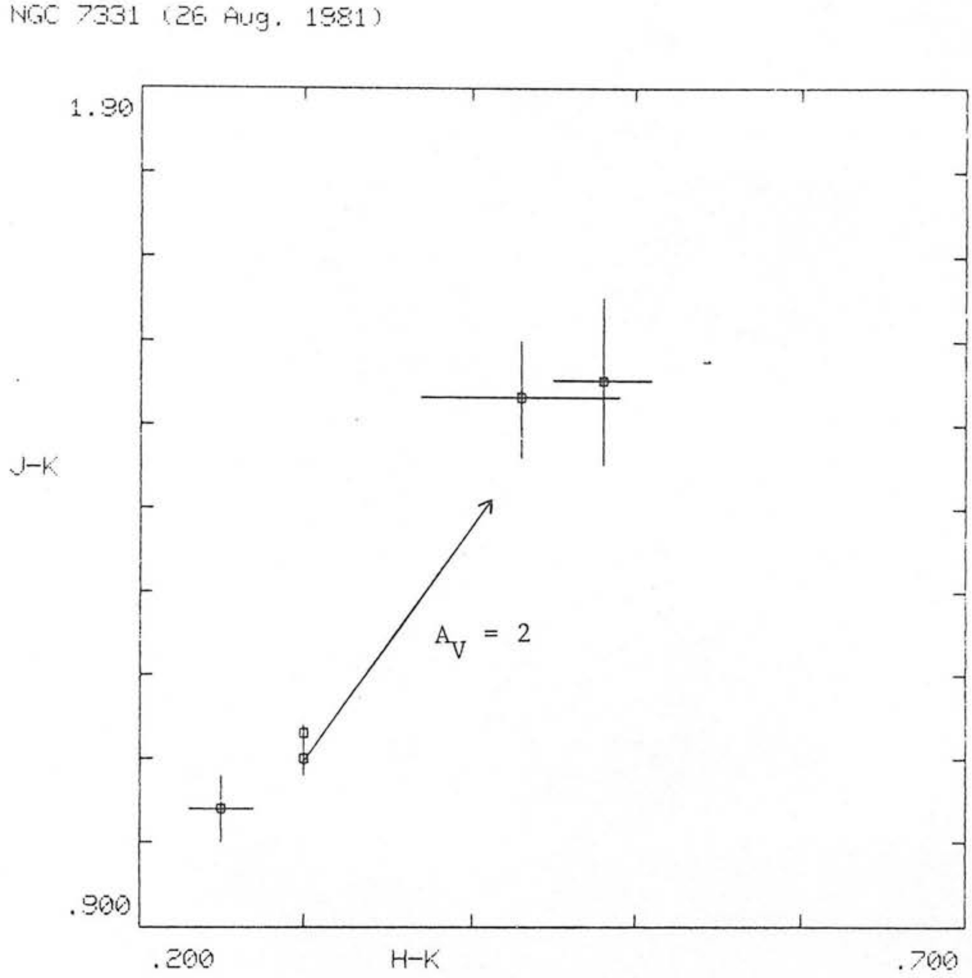


Fig. 5.8 The two colour diagram for selected points from the N/S profile of NGC 7331. The points are the nucleus, two points 10 arcseconds either side of the nucleus and the two reddest points. A reddening vector (Koorneef 1982b) is drawn from the point representing the nucleus.

direction as the stellar sequence. Similar colour variations could conceivably be caused by variations in the stellar population such as increased numbers of carbon stars or late type supergiants. NGC 7331, by nature of its Hubble type, should be similar to the Milky Way, in which the proportion of carbon stars is believed everywhere to be small (Section 2.5). Although large concentrations of carbon stars are known to occur in the Magellanic Clouds and nearby dwarf galaxies (e.g. Cohen et al 1981), it is not reasonable to expect that spiral galaxies should resemble these types of galaxies rather than the Milky Way.

If either carbon stars or late type supergiants were responsible for the red colours, then the K signal should be enhanced rather than the J signal diminished. The 5 arcsecond data show very clearly that the J signal is diminished; the 10 arcsecond data do not show this so clearly as may be expected due to the beam dilution effects. There is no evidence for an excess in either the J or K data. Although an enhanced population of late type supergiants may cause red colours over small regions, they could not be responsible for the red colours over a region extending from 3 to 9 kpc, without very active star formation whose effects should still be seen as very extensive HII regions.

Other possible causes for the reddening, such as hot dust or gaseous emission, can also be discounted. As discussed in Section 2.6 neither is likley to provide enough luminosity to be detectable. The combination of such emission with the bulge light is also unable to reproduce the colours observed when plotted on the two colour diagram.

While it is possible to reproduce the colours by an appropriate, if unlikely, mixture of constituents, the conclusion must be drawn that the dominant source of reddening in the data is extinction caused by dust. The properties of the dust itself are considered in Chapter 6 and the distribution of the dust, with implications for the molecular gas distribution, are considered in Chapter 7.

5.2.3 Comparison with Optical Data

NGC 7331 is considered to be a normal spiral galaxy (e.g. Kormendy and Norman 1979) and so may be expected to exhibit luminosity profiles representing disk and bulge components. Kormendy (1977) has shown that an iterative technique is required to fit theoretical profiles to the data. The method fits disk profiles to the disk dominated regions and bulge profiles to the bulge dominated regions, first subtracting from each the contribution of the less dominant component. Such techniques could be applied to the present data but this was not considered worthwhile because the number of free parameters (two for each component and the position angle) is quite large compared to the number of independent data points. In addition the profiles show irregularities from the smooth theoretical profiles.

These irregularities can be lessened by azimuthal averaging, so that the brightness measurements in all radial directions from the nucleus can be used. Boroson (1981) has applied this technique, using photographic data, to 26 spiral galaxies including NGC 7331. Ellipses are fitted to the isophotes and the average brightness around these ellipses are used as input data for the model. In the case of NGC 7331, the extensive extinction on the near side of the galaxy may be expected to affect both the fitting of the ellipses and the average B brightness around the ellipses. The effect of the extinction on the fitting of the ellipses would be to underestimate the ratio of the lengths of the minor and major axes and so overestimate the inclination. The larger width of the bulge compared to the disk would have the opposite effect when the galaxy is highly inclined. The inclination derived by Boroson (72 degrees) is very close to that derived using HI velocity data (75 degrees, Bosma 1981). This close correspondence may indicate that the extinction does not affect the aximuthally averaged B data but it could also be somewhat fortuitous.

The thickness of the bulge may also affect the fitting of the bulge profile. However Boroson indicates that this is not significant outside the central 20 arcseconds. In spite of these possible problems, Boroson obtains a very close fit of the two component model to the data. The deviations from the computed profile are less than 0.1 magnitudes for most of the points as far out as 260 arcseconds from the nucleus. A limited comparison can be made between this azimuthally averaged profile and the infrared major axis profile.

The K major and minor axis profiles are shown in Fig's 5.3 and 5.4. The N/S and E/W profiles (Fig. 5.2) are generally very similar to the profiles of the axes. The major and minor axis scans will be considered in more detail, because they were obtained with the much more stable Mark II cryostat and because the larger beam size used averages out small scale irregularities and small pointing errors. The major axis K profile shows variations from a smooth profile, but when the data from the two sides of the scan are averaged they form a generally smooth profile (Fig. 5.9), in which the error bars represent the differences between the two sides. The statistical error associated with each point is considerably smaller.

Boroson does not present a major axis B profile, but the variations in K brightness seem generally to follow the brightness as indicated in the B photograph (Fig. 5.1, also Sandage and Tamann 1981). The maximum difference in K magnitude between the two sides of the major axis scan is 0.4 magnitudes but generally much less. It has been suggested that spiral features affect not only the young population but also the old underlying disk population. Zwicky (1955) observed "red arms" in NGC 5194 and the phenomenon has been studied more generally by Schweizer (1976). He found density variations in the old disk population of Sb and Sc galaxies of up to +30 percent. It is possible that the variations in the K profiles represent the same phenomenon. Simultaneous K and B measurements would allow the contrast between arm and interarm regions to be compared more accurately between the young and old disk populations, although the extinction could be expected to affect the B data in such a highly inclined galaxy.

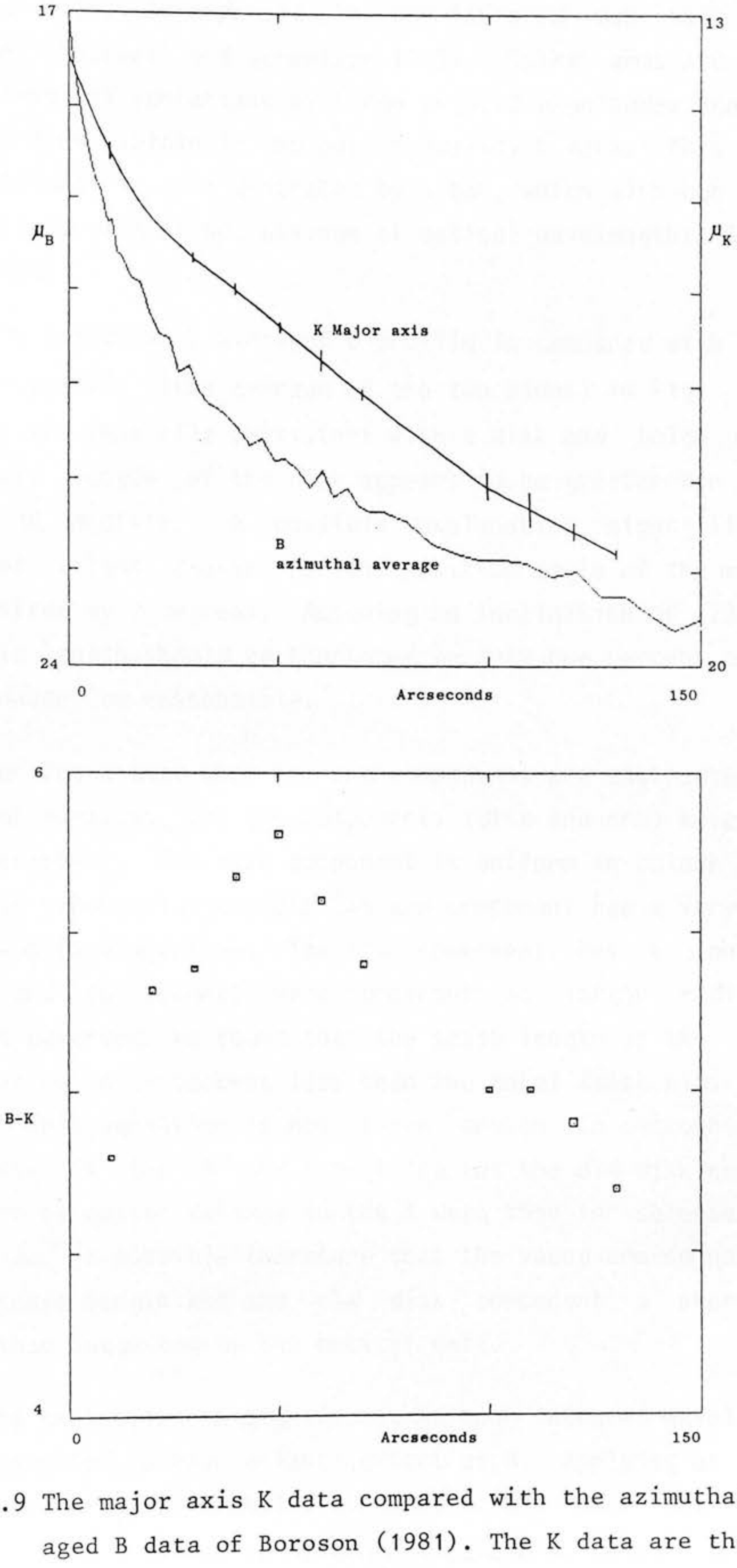


Fig. 5.9 The major axis K data compared with the azimuthally averaged B data of Boroson (1981). The K data are the average of the NW and SE scans (Fig. 5.3) and the error bars represent the difference between the two sides.

A similar effect is observed in an infrared map made at H of NGC 1566 (Hackwell and Schweizer 1983). Spiral arms are seen at H, where azimuthal variations as large as 0.78 magnitudes are observed, and are also visible in the poorer quality K data. This appears to be due to massive arms generated by a bar, which although evident in the infrared data is not obvious at optical wavelengths (Sandage and Tamann 1981).

Boroson's azimuthally averaged B profile is compared with the major axis K profile (the average of the two sides) in Fig. 5.9. Both profiles are generally consistent with a disk and bulge model but the scale length of the disk appears to be greater for the B than for the K profile. A possible explanation might lie in the different values chosen for the position angle of the major axis; these differ by 7 degrees. Assuming an inclination of 72 degrees, the scale length should be shortened by only one percent and so this effect cannot be responsible.

Schweizer found that when the arm components are subtracted from the disks of galaxies, the two components (disk and arm) have different characteristics. The disk component is uniform in colour and obeys a smooth exponential whereas the arm component has a very composite and non-uniform spectrum. The arm component has a longer scale length and so becomes more dominant at larger radii. In the galaxies observed, he found that the scale length of the underlying disk was up to 22 percent less than the total (disk plus arm) scale length. This variation is not large enough to account for the difference in the B and K profiles but the old disk component is likely to be better defined in the K data than in selected optical data. It is possible therefore that the young arm component has a longer scale length and the old disk component a shorter scale length than suggested by the optical data.

Since the extinction is significant at near infrared wavelengths, it can be expected to have a large effect at B. Applying an extinction correction to the azimuthally averaged B data does not allow the intrinsic B luminosity to be deduced reliably, because of the large optical depths. Access to the major axis B data would be advantageous, but because of small scale variations in extinction simultaneous B and K measurements would be required properly to address the problem. Face-on galaxies would suffer less from extinction effects, but these would still be significant for a galaxy similar to NGC 7331 since the equivalent face-on excess J-K colour is 0.04 magnitudes, corresponding to an extinction of about 0.3 magnitudes at B.

The K magnitudes were measured in a 10 arcsecond aperture and so can be converted to mag's/arcsec^2 by adding 4.7. The resulting B-K colours are shown in Fig. 5.9. The colours in the inner and outer parts of the profile are similar to the colours derived by Griersmith (1982) for early type spiral galaxies but are much redder at intermediate radii. The overall shapes of the J-K and B-K profiles are similar suggesting that extinction affects the azimuthally averaged B profile. In the inner region the difference between the J-K and B-K profile shapes could be due to the errors introduced by the thickness of the bulge affecting the fitting of elliptical isophotes to the B data.

If Boroson's averaged profile is interpreted simply as a major axis profile and subject to the effects of extinction, then a correction can be applied based on the J-K colour excess. In Section 7.1 it is shown that the J-K excess appears to be constant, within the errors, at 0.14 magnitudes between 30 and 90 arcseconds and zero elsewhere. The equivalent B extinction is then 1.06 magnitudes (Koorneef 1983) equivalent to an optical depth of 0.98. Assuming that the extinction is centrally located within the luminosity, then the observed and corrected signals (I and I') are related by

$$I = I'(0.5 + 0.5\, e(-\tau))$$

and so in this case the B luminosity wil suffer 0.4 magnitudes of extinction. This will be an underestimate if the dust is not evenly distributed (Section 6.2.4).

The excess B-K colours observed are comparable in size with this calculated value of 0.41 magnitudes. The maximum possible colour introduced by dust centrally embedded in the luminosity is 0.75 magnitudes, less than the value of about 1 magnitude observed at 50 arcseconds. This is not however surprising since the B and K data sets are not completely equivalent.

The K profiles therefore may represent the luminosity distribution more accurately than the B profiles. They imply a shorter scale length for the disk component than that derived by Boroson or alternatively an extra component could be introduced to explain the profiles. There is a slight suggestion in the K data (Fig. 5.9) that such a component exists in the middle part of the profile. This could be a lens or thick disk but could also be the result of a central hole in the stellar disk. More K data extending to larger radii would help resolve the problem.

5.3 NGC 4565

5.3.1 The Data

The 4 arcsecond aperture of the Mark I cryostat was used for all the measurements of NGC 4565. The TV crosshead was used for positioning and beamswitching was applied about every point. The chop was about 120 arcseconds E/W.

The major (Fig. 5.10) and minor (Fig. 5.11) axes were observed in J-K on 30 and 31 January 1982. A two dimensional map was made on 1 February by making scans parallel to the minor axis; the minor axis itself was not remeasured. Subsequent inspection of the data showed that although the minor axis profile was similar in shape to the parallel profiles, the colours were significanly bluer after normalisation to the nucleus. This is believed to be the result of inferior chopper performance on the 31 January relative to 1 February causing smearing of the effective beam profile. It was not therefore possible to combine the data from the two nights and in the two dimensional map (Fig's 5.12 and 5.13) the minor axis is

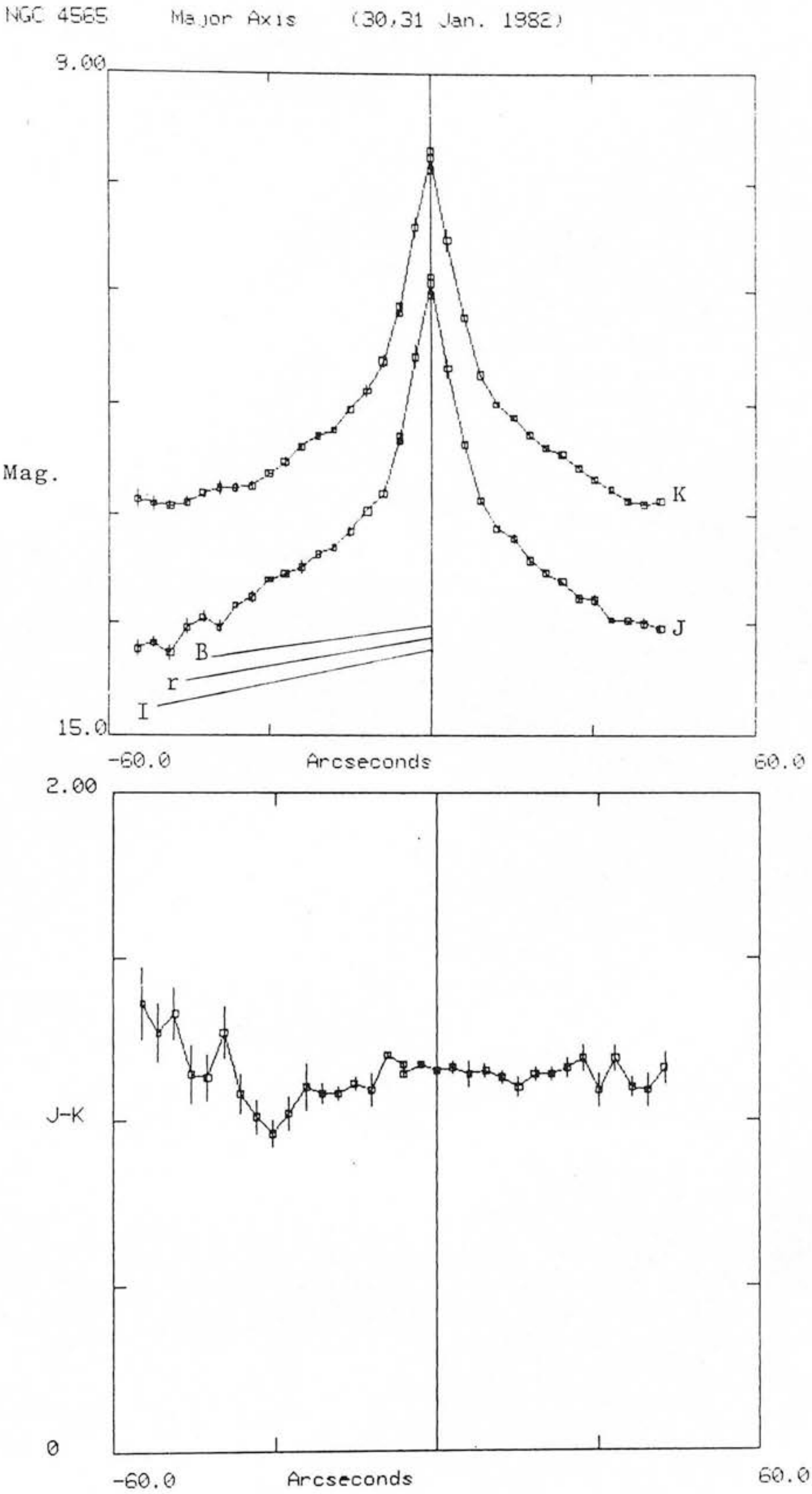


Fig. 5.10 The major axis profile of NGC 4565 measured in a 4 arcsecond aperture. The three lines in the upper frame represent the disk scale heights in B, r and I as determined by Jensen at Thuan (1982) at larger radii.

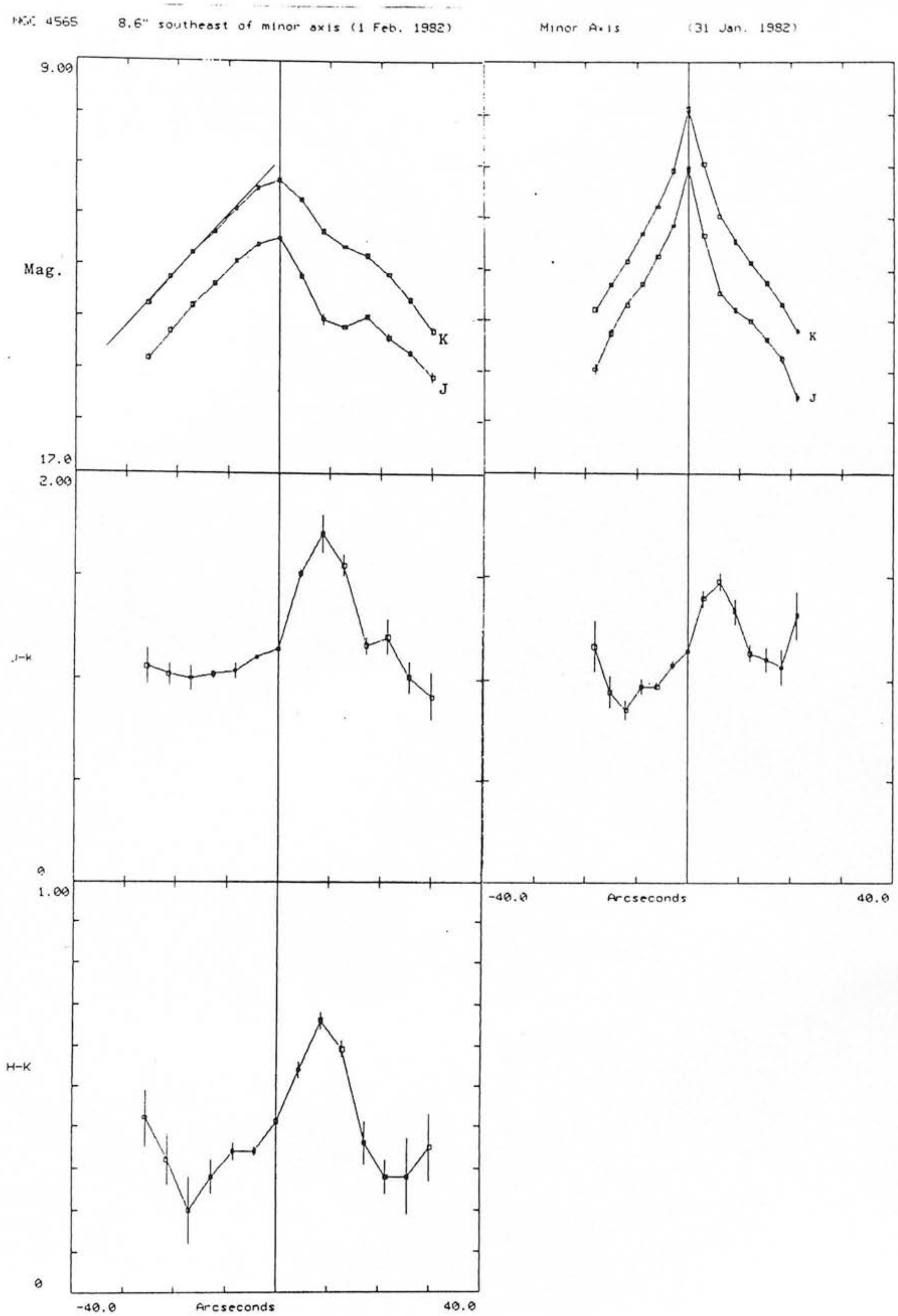


Fig. 5.11 The profiles of a scan parallel to the minor axis, 9 arcseconds SE of the nucleus (left) and of the minor axis (right). A 4 arcsecond aperture was used. The straight line in the upper left frame has slope 0.107 mag's/arcsecond as deduced from the optical data (Jensen and Thuan 1982).

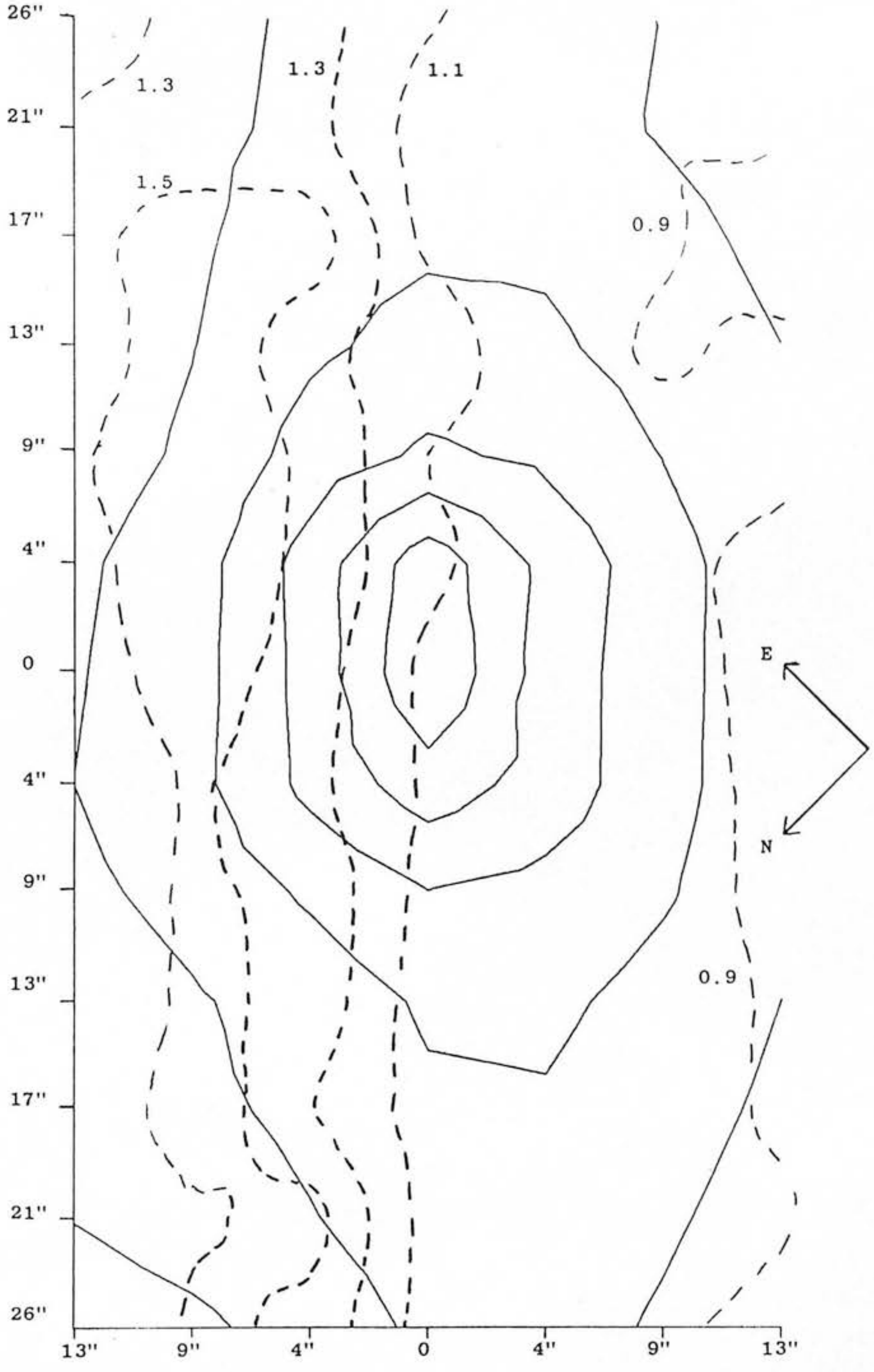


Fig. 5.12 Brightness and J-K colour contour map of NGC 4565. The brightness contours are spaced by 0.5 magnitudes and the brightest is K = 10 in the nominal 4 arcsecond beam. The colour contours are J-K = 0.9, 1.1, 1.3 and 1.5. The grid spacing was 4.3 arcseconds nominal. The scan through the minor axis (horizontal direction on the diagram) is an interpolation between the adjacent scans (see text).

a linear interpolation between the adjacent scans.

A perpendicular profile (PP) 9 arcseonds SE of the minor axis was observed in J-K and H-K on 1 February 1982 to provide z profiles and two colour data without the effects of the large brightness gradient near the nucleus. The H-K data were obtained in two sets either side of the major axis. A zero point shift in H-K was observed affecting the nucleus by 0.11 magnitudes and the PP major axis point by 0.05 magnitudes. This is believed to be the result of improper filter positioning causing an internal vignette of the H detector. The PP data were normalsied to the PP major axis point but a large systematic H-K error may be present. The relative H-K colours in the scan are believed to be correct; any significant difference in effective beam position between the two channels would cause the colours to appear systematically bluer on one side of the major axis than the other. A shift in the positions parallel to the major axis would not cause such an effect but would introduce scatter into the data.

5.3.2 Colour Variations

The colours of the nucleus are

J-K = 1.15
H-K = 0.40

These data were obtained with the Mark I cryostat and so can be expected to be subject to quite large systematic errors. This is especially true for the H-K colour in which a large zero point shift was observed (Section 5.3.1). The J-K colour lies in the region common to other spiral bulges (Fig. 2.2). The H-K colour is considerably redder than would be expected but this effect is considered instrumental rather than real.

The two dimensional data, represented as a Starlink colour image in Fig. 5.14 and as a contour plot in Fig. 5.12, show very strong colour variations. NGC 7331 (Section 5.1.5) showed colour variations which were attributed to the effects of extinction by dust. The arguments used for this deduction are illustrated even more clearly by the NGC 4565 data. Comparison of the infrared and visual images (Fig's 5.14 and 5.1) shows that the red infrared colours correspond to the dark dust lane.

The strong colour variations are also seen in the minor axis scan and the parallel profile (9 arcseconds SE of the minor axis) of Fig. 5.11. They correspond to dips in the profiles, which are stronger at J than at K, with the H dip being intermediate, and so suggest very strongly that the cause is extinction.

The red colour seen at the ends of the minor axis scan and the PP H-K scan do not have any obvious physical explanation. They correspond to positions in the bulge which do not differ in visual appearance to positions closer to the major axis. While metallicity variations are known to occur in bulges, the size and sense of the observed effect make this interpretation unlikely (Section 2.7). A more likely explanation is an instrumental effect. These red colours occur where the bulge is becoming faint rapidly and so a large proportion of the signal could be provided by the nearby and bright dust lane region due to the imperfect performance of the chopper (Section 4.3). In any case the red outer points individually are barely statistically significant. There is no evidence therefore, in the minor axis data for any cause of reddening other than extinction.

The major axis profile does not show any systematic colour variations except for a tendency to become redder towards the SE. The major axis and the dust lane are very close, only about one beamwidth apart near the nucleus, and tend to become closer at larger radii. Also dust features can be seen in the photograph, closer to the major axis than the pronounced dust lane, and so it can be expected that colour variations will occur, due to the irregular nature of the dust distribution.

For all the data, therefore, the colour variations can be interpreted solely as the result of extinction by dust. The red colours at the ends of the minor axis and PP scans are believed to be instrumental in origin. The properties of the extinction are discussed in Chapter 6.

5.3.3 Comparison with Optical Data

Surface photometry of NGC 4565 has been concerned primarily with profiles perpendicular to the major axis. Minor axis profiles have been obtained by van Houten (1961), Frankson and Schild (1976), Kormendy and Bruzual (1978), van der Kruit (1979), Davis et al (1980) and Hamabe et al (1980) at various wavelengths and using various techniques. Jensen and Thuan (1982) also obtained various scans perpendicular to the major axis by photographic and photoelectric techniques through several filters. They found that their data were consistent with those obtained previously but of higher quality.

The profiles of Jensen and Thuan show dips corresponding to the dust lane. Excluding these data they fit stellar distributions to the perpendicular profiles and find that the luminosity between z = 20 and z = 60 arcseconds (where z is the distance from the major axis) is dominated by a thick disk of the type originally found in S0 galaxies (Burstein 1979). The thick disk has bulge-like colours and so might be expected to contribute more to the infrared profile than the thin disk composed of young objects. However in the region mapped, the perpendicular profiles are likely to be bulge dominated.

The major axis of NGC 4565 has been less extensively observed than profiles perpendicular to it. Jensen and Thuan present low resolution (30 arcsecond) data extending to large radii and show that the major axis is asymmetric and has a well defined cut-off at a distance of about 6 or 7 arcminutes. They fitted exponential functions to the disk surface brightness measurements at distances greater than 3 arcminutes and found that the scale length decreased with increasing wavelength. The scale lengths they deduced for B, r and I were 163, 127 and 110 arcseconds respectively and are indicated by the vectors in Fig. 5.10, which shows the infrared profiles.

The J and K profiles also suggest decreasing scale lengths with increasing wavelength. The scale lengths themselves cannot be reliably deduced from the major axis data because the K profile shows that the bulge contributes over most of the region covered, which is inside that used by Jensen and Thuan to determine their scale lengths. The outer part of the J profile appears more linear than the K profile but it is possible that extinction may also be more significant in this region. The effects of extinction cannot easily be removed because it is likely (Section 6.8) that the dust is optically thick. To determine the infrared scale lengths accurately requires measurements to be made at greater radii avoiding were possible those regions obscured by dust.

Jensen and Thuan postulate that the different scale lengths are the result of varying populations within the disk and the effects of the dust lane. The colour changes are in the same sense as observed in other spiral galaxies by Schweizer (1976), who showed that the spiral arms become more important relative to the old underlying disk as the radius increases. The same effect may also be present in NGC 4565 but is difficult to test because of the effects of dust, even at infrared wavelengths.

The infrared minor axis profile is likely to be affected by windshake due to the proximity of the points to the nucleus. The PP of Fig. 5.11 was obtained on a night when the observing conditions were better and all the points are at least two beamwidths from the nucleus. It is nevertheless very close to the nucleus when compared to the size of the galaxy and so can be compared to the optical minor axis profile. Frankston and Schild (1976) showed that the minor axis profile could be represented better by an exponential than by any of the customary functions (Hubble law or $r^{1/4}$ law) for fitting elliptical galaxy and spiral bulge profiles. Their findings were confirmed by Jensen and Thuan, who showed that the inner and outer parts could both be fitted by exponential functions. For the inner 50 arcseconds of the minor axis profile they showed that the B surface brightness was described by

$$\mu_B = 0.106\ z + 18.61$$

A line with this slope is shown on the PP of Fig. 5.11 and it is seen to fit both the J and K data very closely. The minor axis profile of NGC 7331 (Fig. 5.3) is seen to be incompatible with an exponential function. Another very edge-on bulge, NGC 891, was found by van der Kruit and Searle (1981b) to be well described by an $r^{1/4}$ law. The bulge of NGC 4565 is therefore unusual both at optical and infrared wavelengths.

Kormendy and Illingworth (1982) showed that the rotation of the bulge in NGC 4565 was also unusual in that the rotational velocity did not decrease significantly with increasing distance from the galactic plane. For other spiral galaxies they did find such a decrease, consistent with dynamical calculations for a disk and spherical bulge. They postulated that the bulge was not spheroidal but represented a bar seen end-on. The box shaped, rather than elliptical, isophotes of the bulge (Jensen and Thuan 1982) are also consistent with this interpretation. The infrared data, to which the old bulge stars contribute significantly, confirm that the inner part of the bulge is well represented by an exponential profile.

These infrared data therefore show that the structure of NGC 4565 as seen at infrared wavelengths is very similar to that deduced from shorter wavelengths. The bulge is not spheroidal but could be a bar seen end-on, as suggested by the kinematics. The disk scale lengths show evidence for decreasing with increasing wavelength, consistent with the hypothesis that spiral arms become more important at greater radii. However even at these infrared wavelengths dust may affect the major axis data significantly. Measurements at greater distances along the major axis would allow scale lengths of the old disk population to be determined. Measurements perpendicular to the major axis at greater radii would allow the properties of thick disks to be investigated.

5.4 NGC 5907

5.4.1 The Data

The 10 arcsecond aperture of the Mark II cryostat was used to observe NGC 5907 in J-K and H-K. An area 120 by 40 arcseconds was mapped with a grid spacing of 10 arcseconds (Fig's 5.13 and 5.14). The galaxy is very narrow perpendicular to the major axis and much of the mapped area was too faint to provide accurate data. The chopper characteristics were as chosen for NGC 7331, i.e. 180 arcseconds E/W.

The first of the four rasters (Table 5.4) was made using the TV crosshead whereas the others used the telescope encoders for positioning. The positions printed for each point agreed to better than 1.5 arcseconds rms between the rasters. The last half of the last scan of the crosshead raster was improperly positioned due to a fault driving the crosshead, but was not repeated since the region was too faint to provide accurate data. These points are omitted from the maps.

5.4.2 Colour Variations

The optical appearance of NGC 5907 shows it to be dominated by a dust lane which passes very close to the nucleus.The nuclear colours were measured to be

$$J-K = 1.50 \pm 0.02$$

$$H-K = 0.46 \pm 0.01$$

and so are considerably redder than other spiral bulges. Compared to NGC 7331 the excess colours are

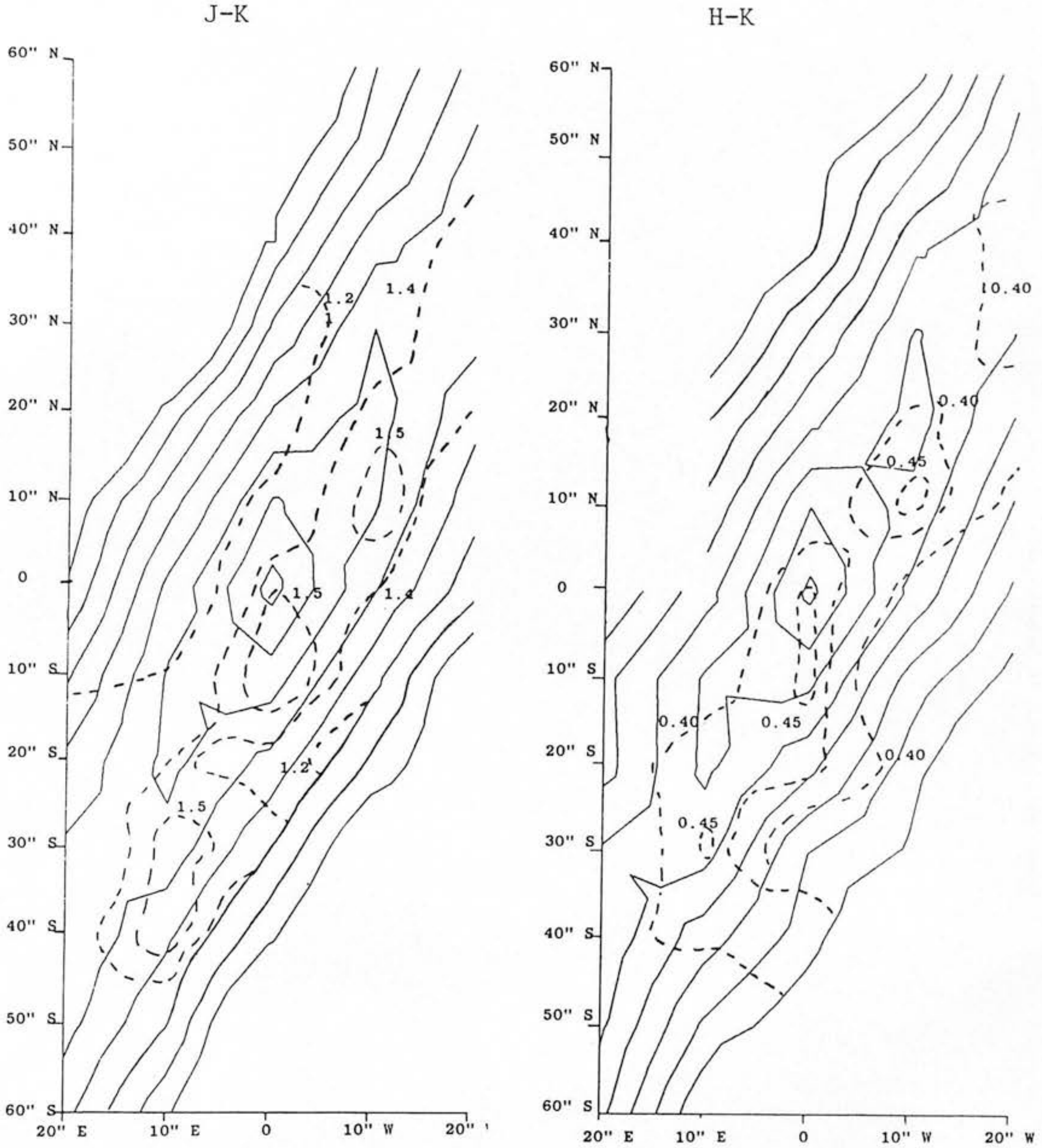
$$E(J-K) = 0.46 \pm 0.02$$

$$E(H-K) = 0.15 \pm 0.02$$

implying

$$E(J-H)/E(H-K) = 2.1 \pm 0.4$$

This colour excess ratio and the optical appearance indicate that the cause of the reddening is extinction due to dust. Although the colour excess ratio is very close to that obtained by Jones and Hyland (1980) for extinction reddening in galactic dark clouds, this


J-K

H-K

Fig. 5.13 Brightness and colour contour maps of NGC 5907. The brightness contours are spaced by 0.5 magnitudes with the brightest being 14.5 K magnitudes arcsecond^{-2}. The J-K contours (left map) are 1.2, 1.4 and 1.5. The H-K contours (right map) are 0.40 and 0.45. The grid spacing was 10 arcseconds.

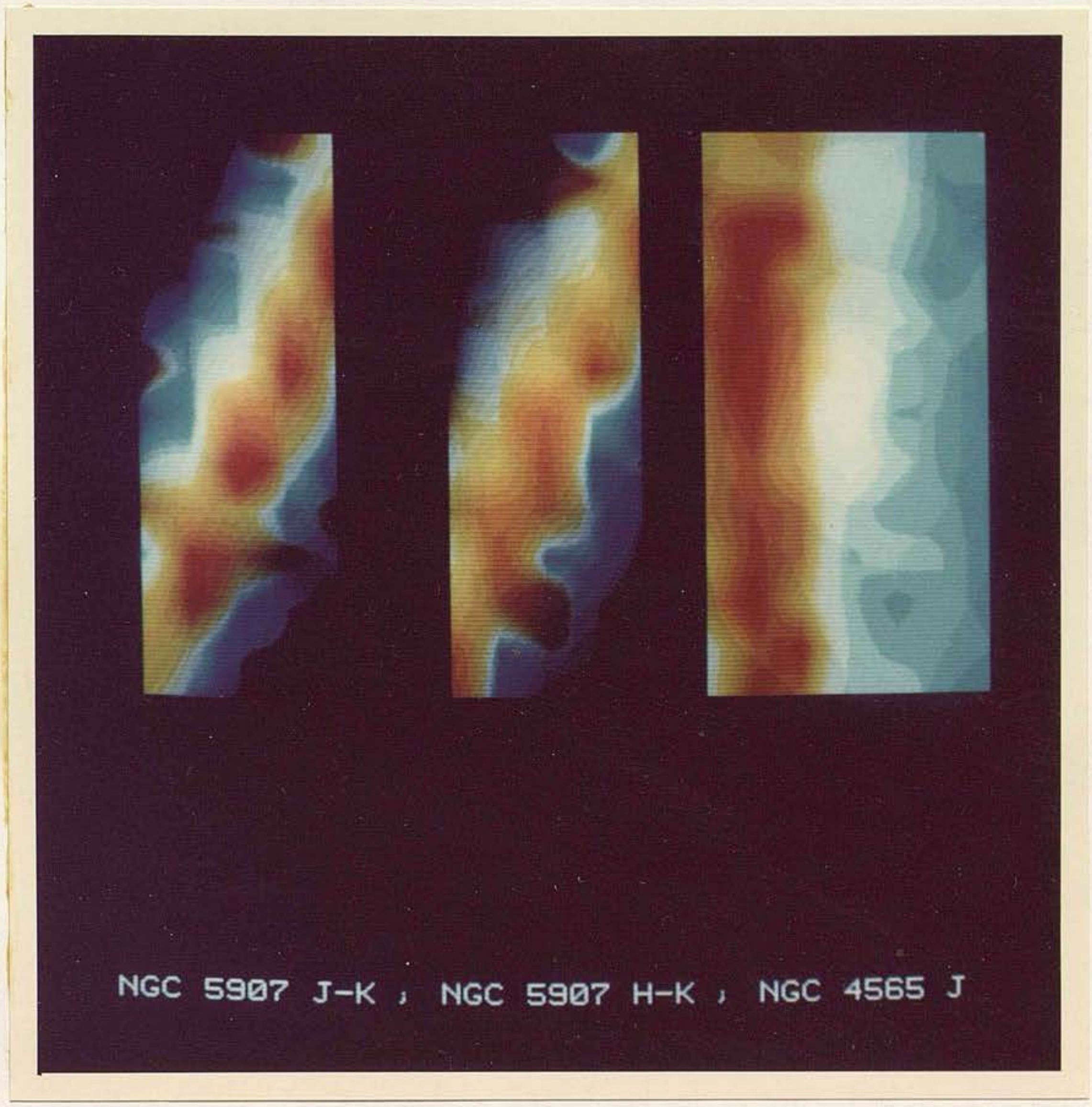


Fig. 5.14 Starlink images of the two dimensional NGC 5907 and NGC 4565 data. The images are, from left to right, NGC 5907 in J-K, NGC 5907 in H-K and NGC 4565 in J-K, corresponding to the contour maps of Fig's 5.12 and 5.13. The middle horizontal scan through NGC 4565 is interpolated between the two adjacent scans (see text). The brightness and colour of the maps are as described in the caption for Fig. 5.7 and in the text.

correspondence may be fortuitous if the dust is optically thick or if positioning errors affect the data. The latter effect appears not to be significant since the K magnitudes from the two sets of data differ by only 0.04 magnitudes, in a region in which the brightness gradient is 0.2 mag's/arcsec. The reddening is further discussd in Chapter 6.

5.4.3 Brightness Distribution

The brightness gradients in NGC 5907 are very large, typically 0.2 mag's/arcsec. The data were obtained in a grid with 10 arcsecond spacing and so the major and minor axis profiles can only be obtained by interpolation. Since the measured points on each side of the axes will be fainter than the points on the axes, such an interpolation would require the fitting of curves, the peaks of which would indicate the brightness on the axes. This then really constitutes an extrapolation procedure which was not considered justified.

In addition small pointing errors can produce large brightness errors. Comparing the K magnitudes of the J-K and H-K maps shows that some points differ by up to 0.4 magnitudes in K. This is consistent with small pointing errors in the order of 2 arcseconds, but would make any extrapolation procedure very uncertain.

The minor axis profile could be obtained without these extrapolation uncertainties, since the brightness gradient parallel to the major axis is small. However the galaxy is so narrow compared to the grid spacing that no minor axis profile was obtained. No detailed comparison with optical data, such as that of van der Kruit and Searle (1982a) has therefore been attempted.

The infrared map (Fig. 5.14) can however be compared with the optical photograph (Fig. 5.1) in qualitative terms. In the infrared, the galaxy appears as a disk with a very small bulge; the dark lane is not apparent due to the much reduced optical depth at K compared with B. Higher resolution data of high positional accuracy would allow profiles to be obtained perpendicular to the plane. The spheroidal appearance of the bulge and the absence of a thick disk, deduced from optical data (van de Kruit and Searle 1981), could be tested in the light of the old disk and bulge populations. Low resolution data at larger distances along the major axis would allow mass to light ratios to be studied in the region where there is an optical cut-off but where the disk continues in neutral hydrogen (Sancisi 1976).

5.5 NGC 7814

NGC 7814 was observed in the 5 arcsecond aperture of the Mark I cryostat in J-K on 31 August 1981. The major and minor axes (Fig. 5.15) were measured, chopping about 150 arcseconds E/W. The galaxy is very edge-on, de Vaucouleurs (1958) deriving a value of 90 degrees for the inclination. The major axis scan is therefore along the prominent dust lane and the minor axis meets the dust lane at the nucleus.

Surface photometry has been performed by Hamabe and Okamura (1981) in the B band along the major and minor axes. The minor axis B profile is brighter than the inner part of the major axis due to the extinction of the edge-on dust lane. The unextincted major axis profile will be at least as bright as the minor axis profile and so the maximum of the minor and major axis profiles gives a lower limit to the major axis luminosity. The minor axis B profile matches well with the major axis K profile up to 20 arcseconds beyond which the B minor axis profile declines more rapidly. This is consistent with the B and K scale lengths being similar.

The major axis B profile shows more irregularities than the K profile due to the greater extinction. The extinction is also more clearly defined in the B data due to the 1.1 arcsecond aperture used (Hamabe et al 1979) rather than the 4 arcsecond infrared aperture.

NGC 7814 was identified by Griersmith et al (1982) as being anomalously red in J-K and H-K; this they suggested was due to the

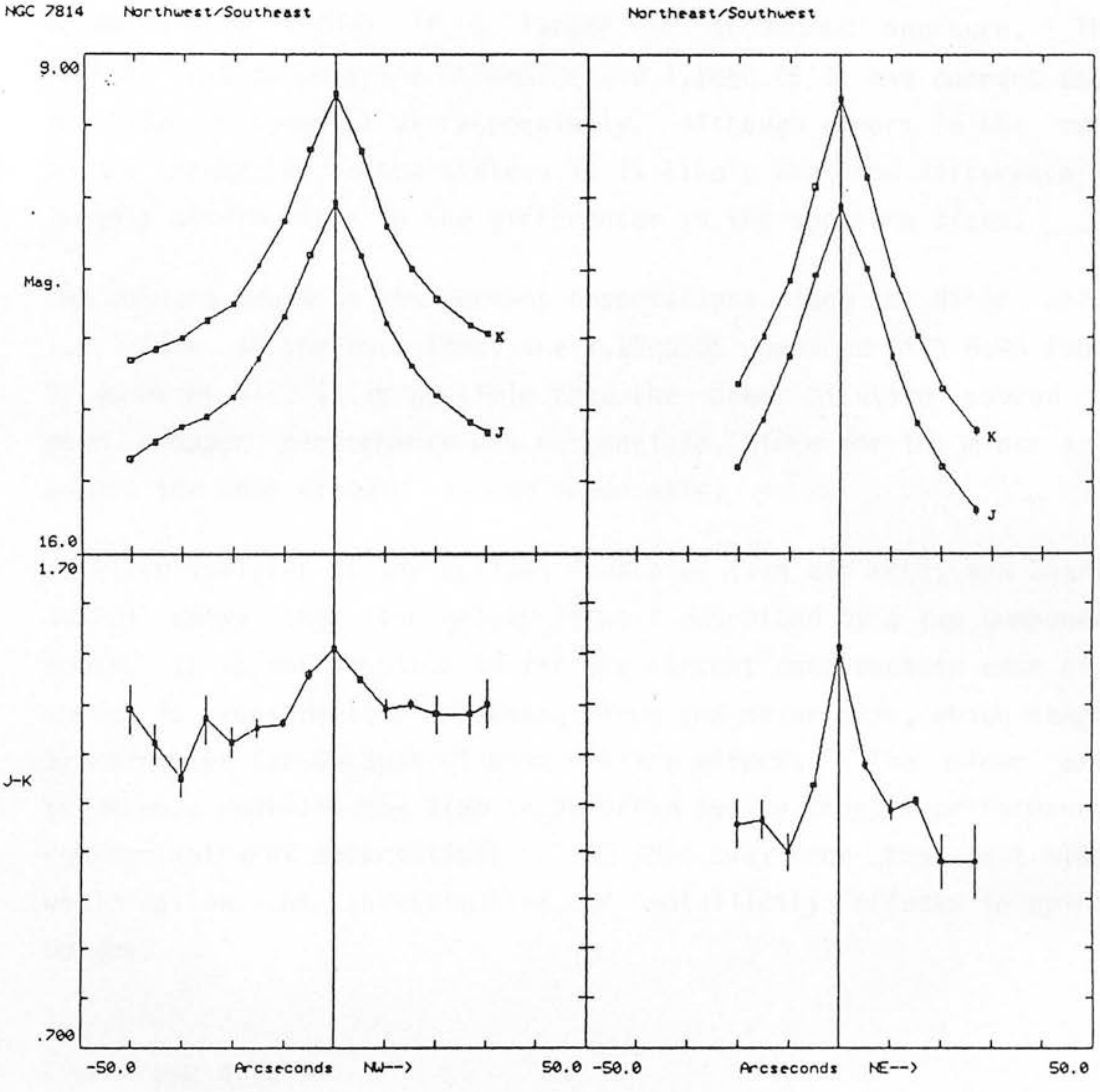


Fig. 5.15 NW/SE and NE/SW profiles of NGC 7814 in a 5 arcsecond aperture.

dust lane. These data and those of Jones et al (1983) show that significant infrared reddening is observed. Although the dust lane is very narrow (it is not resolved in the 5 arcsecond spaced profile) and so covers only a small proportion of the surface area, it does obscure the nucleus and major axis, contributing significant reddening to the integrated colours. The excess J-K colour at the nucleus is about 0.38 magnitudes, equivalent to about 3 magnitudes at B. This is somewhat larger than the excess colour of 0.32 found by Jones et al (1983) in a larger 12 arcsecond aperture. The nuclear J-K colours are 1.50±0.02 and 1.28±0.05 in the current data and those of Jones et al respectively. Although errors in the zero points could be responsible, it is likely that the difference is largely attributable to the differences in the aperture sizes.

The colours found in the current observations along the minor axis, i.e. out of the dust lane, are 1.12±0.05 compared with 0.96 found by Jones et al. It is possible that the beam dilution caused by poor chopper performance was responsible, since for the minor axis points the chop crossed the red major axis.

Detailed analysis of the optical isophotes (van der Kruit and Searle 1982b) shows that the galaxy is well described by a two component model. It is not possible to fit the current data because even at K there is considerable reddening along the major axis, which cannot be corrected for because of beam filling effects. The minor axis brightenss profiles may also be affected by the chopper performance. Further infrared observations in NGC 7814 away from the dust lane would allow the investigation of metallicity effects in spiral bulges.

5.6 NGC 4216

The 4 arcsecond aperture of the Mark I cryostat was used to measure the N/S and E/W profiles of NGC 4216 in J-K, chopping about 120 arcsecond E/W. The E/W profile is asymmetric due to a centring error of about 2 arcseconds but this does not affect the interpretation of the data.

The most significant colour variations are those seen in the E/W scan (Fig. 5.16) between 14 and 21 arcseconds E of the nucleus. Comparison with the optical photograph (Fig. 5.1) shows that this is the region corresponding to a strong dust lane. The conclusion is therefore drawn again that extinction by dust is responsible for the reddening. Both the E/W and the N/S colour profiles show a gradient which can be interpreted as the effect of the changing geometry as the near side of the galaxy is approached. The proportion of the dust in front of the extinction becomes greater and so the colours become redder. Estimates of the dust density are made in Section 7.6.

The brightness profiles in both directions are generally smooth and bulge dominated and disk dominated regions are clearly visible. Although no fitting of theoretical profiles to the data has been attempted, it is clear that the z profile of the bulge is likly to be better fitted by an $r^{1/4}$ law rather than an exponential as for NGC 4565. Watanabe et al (1982) have obtained profiles along the major and minor axes. These are difficult to compare with the N/S and E/W data because of the 20 degree difference in position angle of the measurements. The large inclination makes simple transformations inaccurate without a three dimensional model of the luminosity distribution.

The K light is seen to fall off more rapidly in the E/W direction which is closer to the major axis than the N/S direction. Both profiles decline at a rate intermediate between those of the major and minor axes observed in B, consistent with the B and K scale lengths being similar.

Optical surface photometry has also been performed by Simkin (1967 and 1975). These data show colour variations along the minor axis caused by population changes as well as by extinction. The infrared data suggest that effects other than extinction may affect the N/S colour profile but the colour variations are not much above the noise level. Higher signal to noise data would allow population effects to be studied, particularly if two dimensional data of high

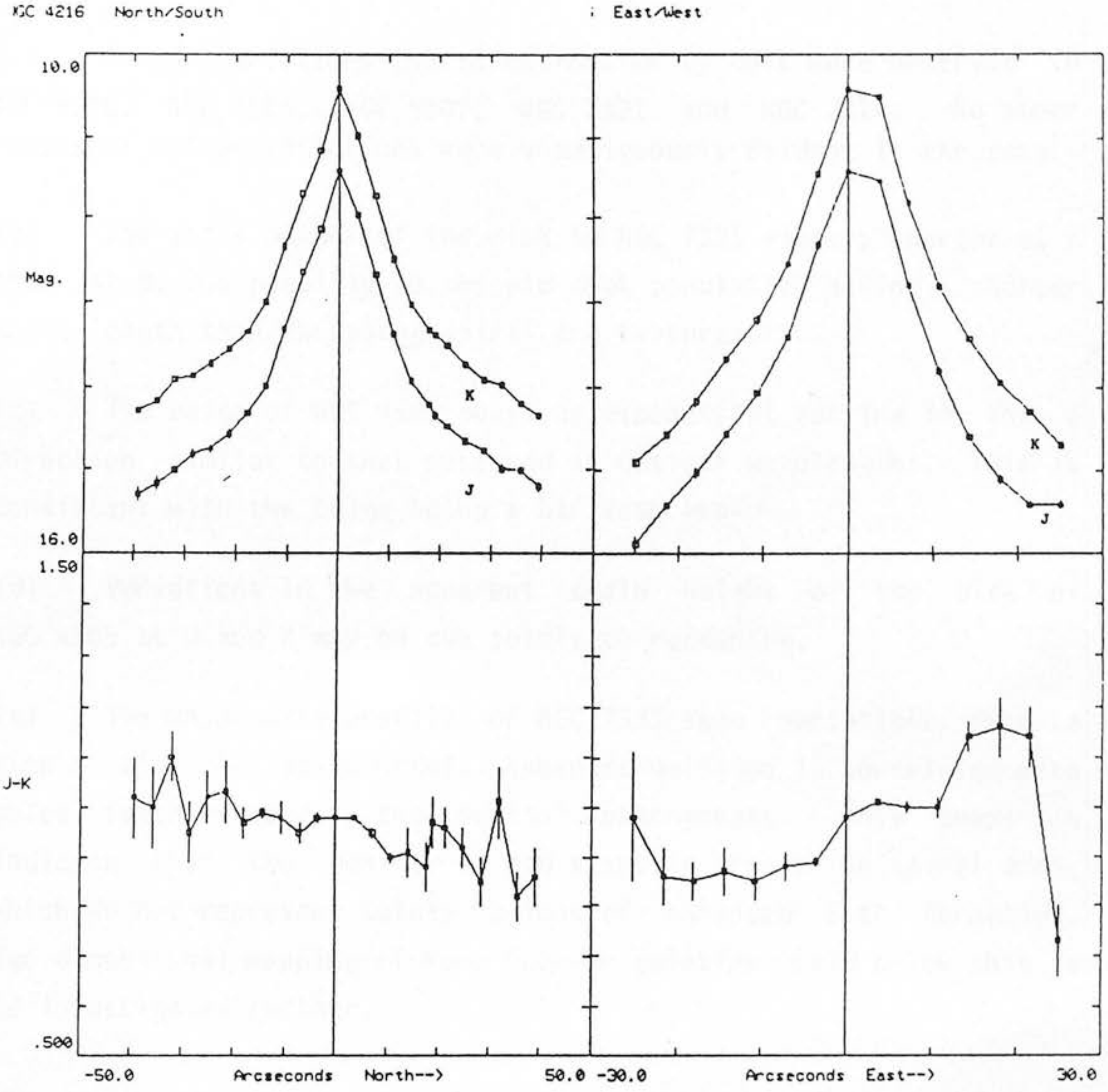


Fig. 5.16 N/S and E/W profiles of NGC 4216 in a 4 arcsecond aperture.

positional accuracy and resolution (e.g. from an array) were available for comparison with the optical appearance.

5.7 Summary of the Data

(a) Colour variations due to extinction by dust were observed in NGC 4216, NGC 4565, NGC 5907, NGC 7331 and NGC 7814. No other causes of colour variations were unambiguously evident in the data.

(b) The scale height of the disk in NGC 7331 appears shorter at K than at B, due possibly to the old disk population having a shorter scale length than the young spiral arm features.

(c) The bulge of NGC 4565 obeys an exponential decline in the z direction similar to that observed at optical wavelengths. This is consistent with the bulge being a bar seen end-on.

(d) Variations in the apparent scale height of the disk of NGC 4565 at J and K may be due solely to reddening.

(e) The major axis profiles of NGC 7331 show variations from a simple disk plus bulge model. Enhanced emission is correlated with spiral features seen in the optical photographs. This seems to indicate that the denisty of old stars is greater in spiral arms, which do not represent solely regions of enhanced star formation. Two dimensional mapping of more face-on galaxies would allow this to be investigated further.

6. Interstellar Dust

6.1 Origin of the Reddening Curve

The first indications that the interstellar medium was not completely transparent were that stars of the same spectral type tended to become redder with increasing faintness and hence distance. Without this wavelength dependence of extinction it would have been very difficult to deduce the existence of interstellar dust. Although it has always been difficult to measure the amount of extinction suffered at any wavelength, it has long been possible to determine colour excesses of sources. If the intrinsic colours of a star are known, then the colour excess represents the difference between observed and intrincic colours. By observing a single source at many wavelenghts the extinctiom curve may be plotted. This conventionally plots against the wavelength λ, the excess colour of measurements at λ and the V band, written as $E(\lambda-V)$. In order to compare extinction curves of sources suffering different amounts of extinction, $E(\lambda-V)$ is normalised by $E(B-V)$, the excess colour of measurements made at B and V. In addition, for reasons explained below, this quantity $E(\lambda-V)/E(B-V)$ is normally plotted against $1/\lambda$ (Fig. 6.1a).

Extinction curves of this type were derived at optical wavelengths by e.g. Whitford (1958) and the striking feature of these observations was that the reddening was approximately linear with reciprocal wavelength, although other features were present particularly a knee or change in the proportionality constant at a wavelength of approximately 4000 Å.

The theory of how dust grains affect the transfer of radiation, the Mie theory, has been extensively investigated (e.g. van de Hulst, 1949) making assumptions about the grain shape and size distribution. The simplest assumption is that the grains are spherical and of uniform size. Energy is removed from the beam by both absorption and scattering and the grains may be defined as having cross sections C_a and C_s for these two cases. The extinction cross section C_s is the sum of C_a and C_s for the

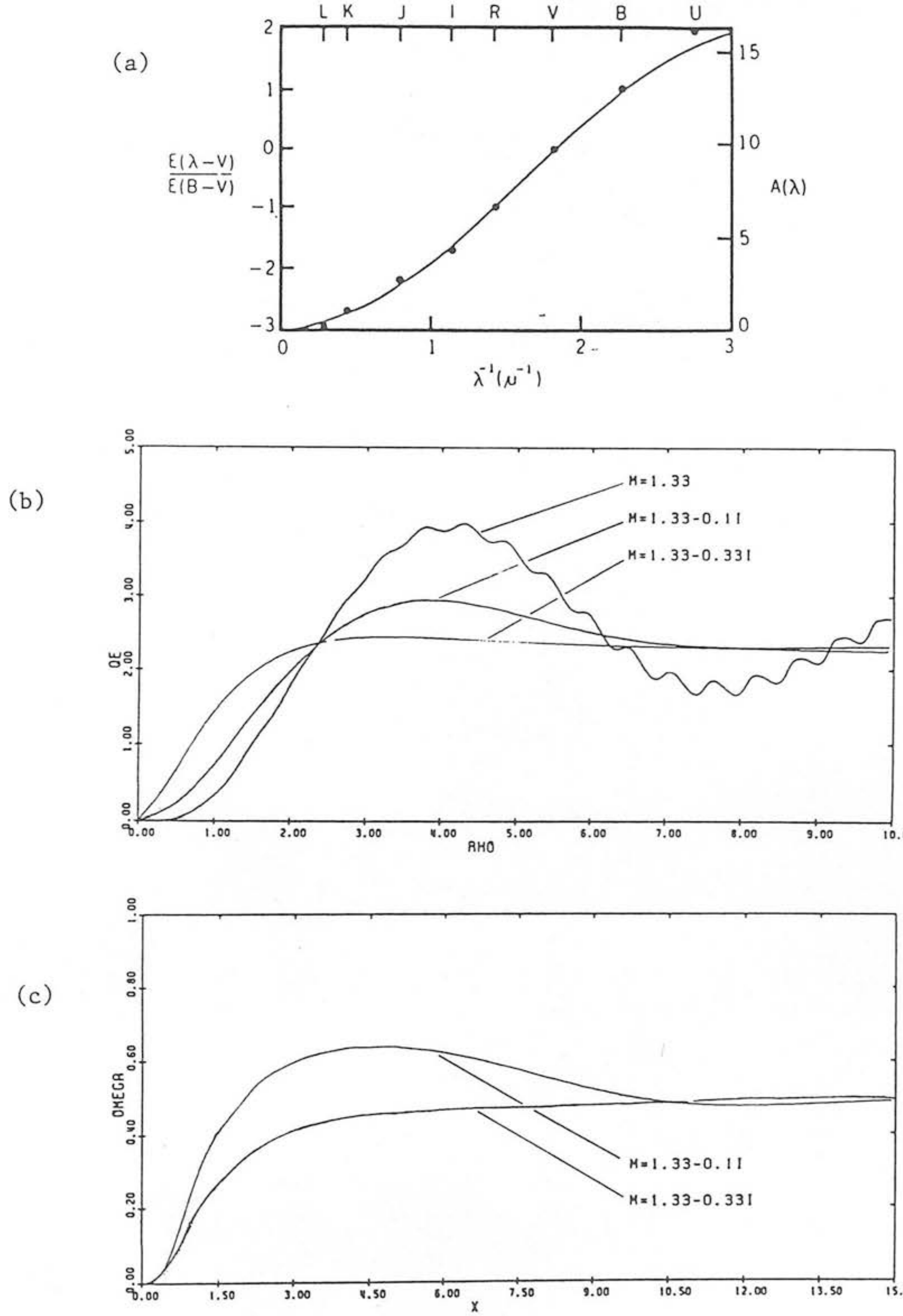


Fig. 6.1 Observed and computed dust properties. (a) shows a typical reddening curve. (b) shows the extinction efficiency as a function of the size parameter ρ, for different values of imaginary refractive index. (c) shows the albedo as a function of the size parameter x for different values of imaginary refractive index, typical of 'dirty ice'. The diagrams are from Martin (1978).

case of physically separate luminosity and extinction when the scattered radiation is removed from the beam. This is not the case for circumstellar dust shells, for example, where as much radiation is scattered into the beam as is removed.

The absorption and scattering cross sections are related to the physical cross section by efficeiencies Q_a and Q_s respectively, and an extinction efficiency Q_e is similarly defined.

$$C_s = Q_s . \pi a^2 \quad , \text{ where } a \text{ is the radius of the particle}$$
$$C_a = Q_a . \pi a^2$$
$$C_e = Q_e . \pi a^2$$

The refractive index of the particles affects the efficiencies and is determined by the composition of the grain. The complex index of refraction is defined as

$$m = n - in'$$

where n' represents the absorption. The efficiencies Q_a, Q_s and Q_e may be expressed as functions of the size parameter x to allow comparisons to be made easily between different particle sizes and wavelengths. This parameter is defined (e.g. Martin 1978) as

$$x = 2\pi a/\lambda$$

The efficiencies are also often plotted as a function of ρ, defined as

$$\rho = 2x(n-1)$$

to show the effect of variation of refractive index. When Q_e is plotted against ρ for perfect dielectric spheres (Fig. 6.1b), it is seen that the relationship is approximately linear when $1< \rho<3$. This linear region is observed to occur in the optical part of the spectrum, implying characteristic particle sizes of 0.1 µm for likely refractive indices of between 1.3 and 1.5. If the characteristic particle size is increased then this linear region will occur at longer wavelengths. The same effect will occur as the refractive index is increased.

The extinction efficiency includes the contributions from absorption and scattering. Pure dielectrics, with no absorption and hence real refractive indices, if placed between the observer and the luminosity source, will cause extinction simply by scattering radiation out of the beam. Only an insignificant amount of radiation will be scattered back into the beam, if the source and extinction are physically well separated, and so the extinction may be described by an optical depth which varies with wavelength as Q_s. If instead the non-absorbing grains are situated around the source in a spherically symmetric arrangement, all of which is within the observing beam, then there is no extinction. Other geometries can scatter more radiation into the beam than out of it, producing reflection nebulae which are rare but not unknown in the infrared (Elias, 1978a).

Real grains are likely to have significant absorption, as indicated by the imaginary part of the refractive index, and this is often more important than scattering, especially at longer wavelengths, so simplifying the radiative transfer problem. The albedo of the grains may be defined as Q_s/Q_e and is a function of x (or wavelength) even when the refractive index is constant with wavelength. For large values of x, i.e. grains large compared to the wavelength, the albedo tends towards the value 0.5 because the only energy removed is that which is diffracted. For small x, which may be applicable to the infrared, the albedo tends towards zero. Scattering will therefore be much less significant in the infrared than in the optical but cannot be disregarded entirely since infrared polarisation (e.g. Dyck and Jones, 1978) and reflection nebulae (Elias, 1978a) have been observed. In the event of scattering being significant, its effect would be to decrease the reddening since the albedo will be larger for the shorter wavelengths.

Grain composition, resulting in various refractive indices, cannot be deduced from the reddening curve unless there are deviations from it caused by identifiable solid state resonances such as the 2200 Å carbon feature or the 9.7 μm silicate feature. The effect of a larger refractive index is essentially indistinguishable from that of smaller grains. The effect of increasing the imaginary component of the refractive index is to move the linear part of the curve to longer wavelengths (Fig. 6.1b), a similar effect to that of larger grains, and to decrease the effects of scattering (Fig. 6.1c).

The assumption that the grains are spherical is not correct as is shown by the existence of polarisation (e.g. Dyck and Jones, 1978), which is zero for spherical grains. However when calculations are performed on oblate spheroids (Martin, 1978, P 31) it is found that the extinction efficiency depends effectively on the smaller dimension. If the grains are non-spherical then the mass of the dust rather than the number of grains will be underestimated if sphericity is assumed.

Although it is impossible to predict the nature of the extinction curve, a semi-theoretical curve derived by van de Hulst (1949) fits the optical data remarkably well. This curve (van de Hulst's No. 15) is calculated from theory assuming spherical grains with a size distribution determined by the condition that grain nuclei form and grow at a constant rate and are subject to evaporation by collision. The wavelength dependent refractive index is that of ice with a mixture of other components which add absorption. The curve describes the optical extinction well although it does not explain the 2200 Å feature and cannot necessarily be extrapolated into the infrared. It nevertheless provides a useful basis for comparison.

Observations of extinction in the ultraviolet (e.g. Nandy et al 1975) show that the extinction curve continues to rise and by comparison with Fig. 6.1b it is evident that a single type of grain cannot explain both the UV and optical characteristics. A population of smaller grains is therefore proposed to explain the extinction in the UV, where larger grains will exhibit neutral extinction. In the optical these smaller grains will have little effect due to the effect of the larger grains. When extending the extinction curve to other wavelengths, such as the infrared, it must always be considered possible that a different population of grains is responsible for the observed properties.

The above discussion has been concerned primarily with the colour variation produced by extinction, because it is this that is most easily observed. To determine dust mass density or to correct brightness measurements for extinction, the amount of the extinction itself must be determined. If the extinction curve is known, then relating the total extinction at one wavelength to the reddening at a pair of wavelengths allows the extinction to be determined at all measured wavelengths. This ratio of total to selective extinction is normally expressed as

$$R = A_V/E(B-V)$$

This quantity allows comparison of reddening in different regions and estimates to be made of the dust densities responsible. It is not easy to measure observationally and is determined by one of two methods. The first is to determine the extinction at some wavelength by assuming the distance and luminosity of the extincted source. The second method is to extend the extinction curve as far as posssible towards longer wavelengths, where, in the limit, $E(\lambda -V)/E(B-V)$ equals R.

Previous estimates of infrared colour excess ratios and ratios of total to selective extinction, both in the Milky and in other galaxies, are discussed in the following sections. Estimates are made of these parameters for other galaxies from the current data.

6.2 Extinction Geometry Effects

6.2.1 Background

Because the extinction due to dust is wavelength dependent, if a source of luminosity is obscured its position will change on the two colour diagram. This excess colour is the difference in magnitudes between the observed colour and the intrinsic colour of the source and is denoted e.g. E(J-K) or E(H-K). The excess colour will be calculated below and plotted on the two colour diagram for a few idealised geometries of extincting material and luminosity.

Assuming that the albedo is low, i.e. that scattering may be neglected, the optical depth τ_λ may be defined in terms of the absorption coefficient κ_λ

$$\tau_\lambda = \kappa_\lambda \int \rho(x)\, dx$$

where $\rho(x)$ is the density per unit volume of the dust and x is the distance along the line of sight between the observer and the source of luminosity. The intensity I_λ that the observer sees is

$$I_\lambda = I_0 \exp(-\tau_\lambda)$$

where I_0 is the intensity which would be observed in the absence of extinction. This may also be conveniently expressed as the extinction A_λ in magnitudes

$$A_\lambda = -2.5 \log \frac{I_\lambda}{I_o} = 1.09\,\tau_\lambda$$

The excess colour E is then

$$E(\lambda_1 - \lambda_2) = 1.09(\tau_1 - \tau_2)$$

Since the dust column density $\rho(x)\, dx$ is independent of wavelength the ratio of the excess colour at the two wavelengths is constant, i.e.

$$\frac{E(\lambda_1 - \lambda_2)}{E(\lambda_2 - \lambda_3)} = 1.09\,\frac{\kappa_1 - \kappa_2}{\kappa_2 - \kappa_3} = 1.09\,\frac{\tau_1 - \tau_2}{\tau_2 - \tau_3}$$

and so, for varying amounts of dust along the line of sight to the luminosity source, a vector is defined of slope $\frac{\kappa_1 - \kappa_2}{\kappa_2 - \kappa_3}$. The distance along this vector, at which the excess colours occur, is dependent on the dust column density and may be used to measure it.

This assumes that the luminosity is all located behind the extincting material as is true when individual stars are being measured. In galaxies however, the luminosity source and the extincting material are generally mixed and the effects of some simple geometries will be considered to illustrate what may be observed.

6.2.2 Homogenous Luminosity and Dust Distribution

In this case (Fig. 6.2a) the luminosity is assumed to be distributed in the same way as the extincting material so that $\ell(\lambda)$, the luminosity per unit area per unit optical depth is constant. The intensity $I(\tau)$ seen by the observer originating in the region down to optical depth τ is

$$\begin{aligned} I(\tau) &= \int_0^{\tau} \ell . \exp(-\tau') \, d\tau' \\ &= \ell(1-\exp(-\tau)) \qquad (6.1) \end{aligned}$$

If I_0 is defined as the intensity from the same volume in the absence of extinction, then

$$I_0 = \ell \, \tau$$

Equation (6.1) can then be written

$$I(\tau) = I_0 \, (1-\exp(-\tau)) \, /\tau$$

For small τ this reduces to

$$I(\tau) = I_0 \, (1-\tau/2)$$

and for large τ to

$$I(\tau) = I_0/\tau \quad (=\ell)$$

The excess colour at wavelengths 1 and 2 is therefore

$$E(\lambda_1-\lambda_2) = 1.09 \ln \frac{1-\exp(-\tau_2) \; \tau_1}{1-\exp(-\tau_1) \; \tau_2} \qquad (6.2)$$

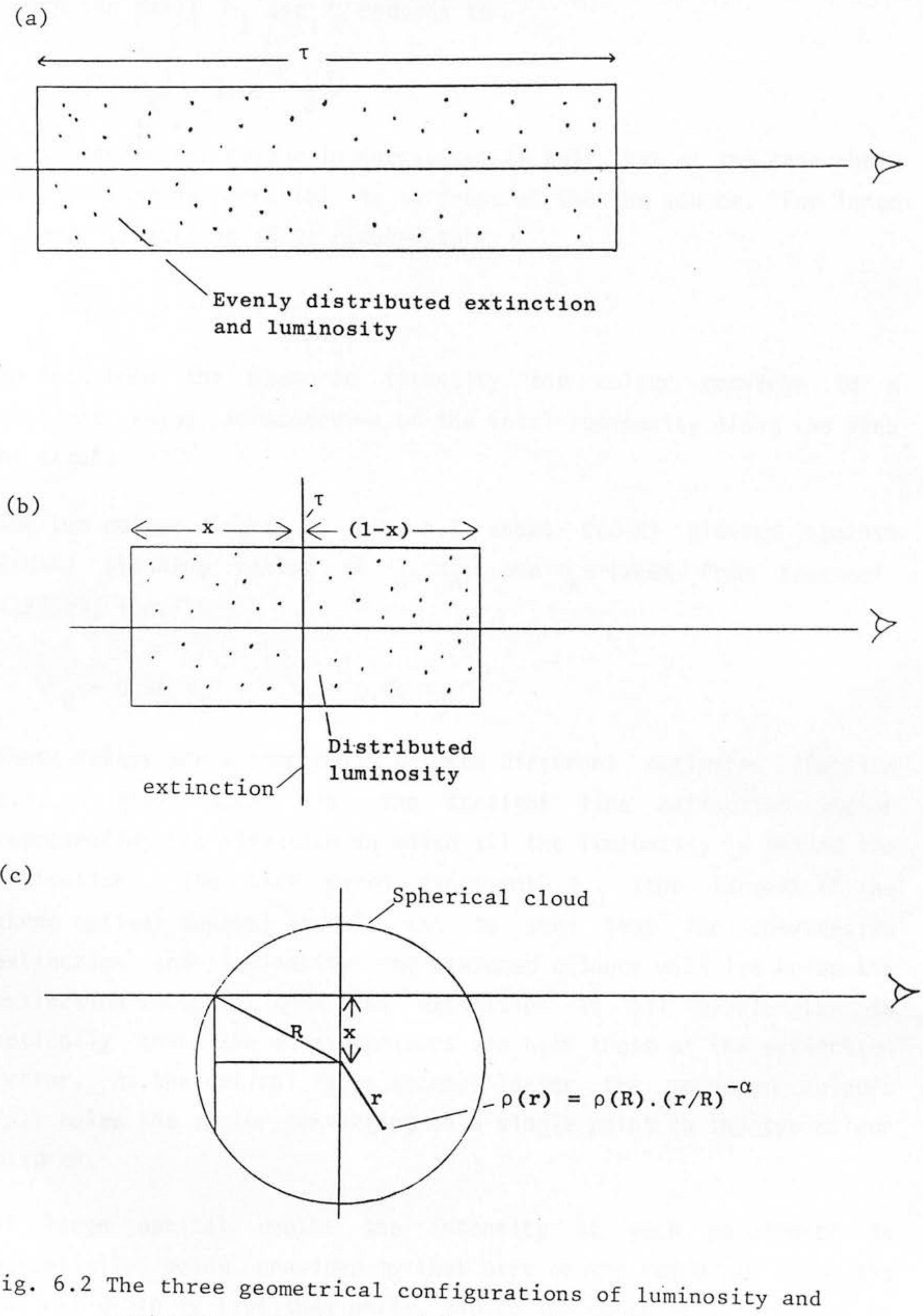


Fig. 6.2 The three geometrical configurations of luminosity and extincting material used in the calculations.

which for small τ_1 and τ_2 reduces to

$$E(\lambda_1 - \lambda_2) = 1.09 \, \frac{\tau_1 \quad \tau_2}{2}$$

i.e. the excess colour in magnitudes is half that of the case where the extincting material is in front of the the source. For large τ_1 and τ_2 equation (6.2) reduces to

$$E(\lambda_1 - \lambda_2) = 1.09 \, \ln\left(\frac{\tau_1}{\tau_2}\right)$$

so that both the measured intensity and colour converge to a constant value irrespective of the total luminosity along the line of sight.

The two colour diagram of Fig. 6.3 shows E(J-K) plotted against E(H-K) assuming ratios of τ_J, τ_H, and τ_K taken from Koorneef (1983b), i.e.

$$\tau_H = 0.58 \, \tau_J \qquad \tau_K = 0.34 \, \tau_J$$

These values are a compromise between different estimates (Section 6.4). Also shown is the straight line extinction vector representing the situation in which all the luminosity is behind the extinction. The tick marks represent τ_J (the largest of the three optical depths) and it can be seen that for coextensive extinction and luminosity the measured colours will lie below the extinction vector. When the extinction at all wavelengths is optically thin the excess colours are half those of the extinction vector. As the optical depth becomes larger the measured colours fall below the vector converging on a single point in the two colour diagram.

At large optical depths the intensity at each wavelength is essentially being provided by that part of the region in which the optical depth is less than unity, and so the contributing region is larger for longer wavelengths. Care must always be exercised therefore in interpreting colours of composite systems where large optical depths are suspected, since differences in the distribution

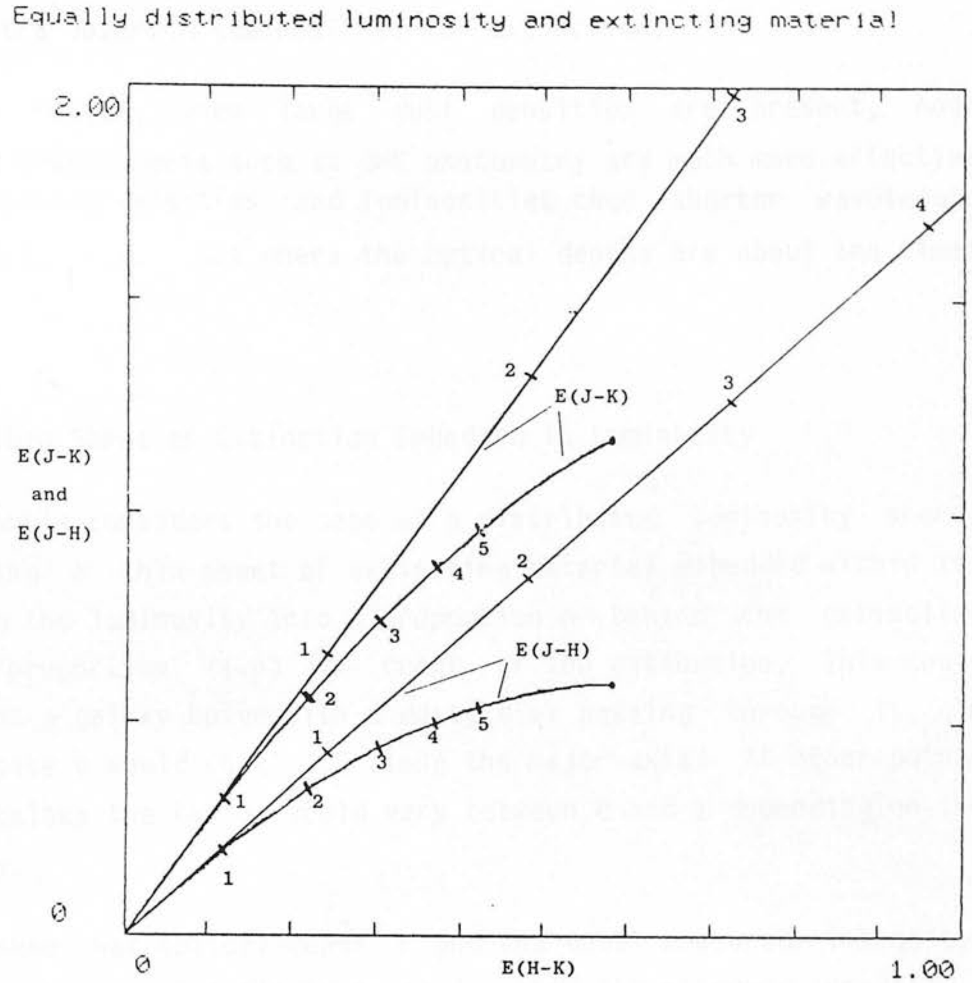


Fig. 6.3 The two colour diagram for equally distributed luminosity and extincting material. The loci represent the colours of the light as τ_J increases as indicated by the tick marks. Also shown are the normal reddening vectors which represent the observed colours when the luminosity is behind the extincting material. The extinction ratios are from Koorneef (1983b).

of extinction and luminosity or variations in the intrinsic colour of the luminosity will cause large fluctuations in the positions on the two colour diagram. In addition such colours cannot be used to correct the observed luminosities for extinction.

For this reason, when large dust densities are present, near infrared measurements such as JHK photometry are much more effective in determining densities and luminosities than shorter wavelength photometry, e.g. UBV where the optical depths are about ten times larger.

6.2.3 Thin Sheet of Extinction Embedded in Luminosity

This example considers the case of a distributed luminosity source containing a thin sheet of extincting material embedded within it, dividing the luminosity into a proportion p behind the extinction and a proportion (1-p) in front of the extinction. This could represent a galaxy bulge with a dusty disk passing through it, in which case p would equal 0.5 along the major axis. At other points in the galaxy the factor would vary between 0 and 1 depending on the geometry.

If the sheet has optical depth τ and the total measured intensity, in the absence of extinction, is I_0, then the observed intensity I is given by

$$I = I_0\ (1-p) + I_0\ p\ \exp(-\tau)$$

The excess colour at wavelengths 1 and 2 is therefore

$$E(\lambda_1-\lambda_2) = -2.5\ \log\left(\frac{1 - p + p\ \exp(-\tau_1)}{1 - p + p\ \exp(-\tau_2)}\right)$$

For small τ_1 and τ_2 this reduces to

$$E(\lambda_1-\lambda_2) = 1.09(\tau_1-\tau_2)p$$

and for large τ_1 and τ_2 to

$$E(\lambda_1-\lambda_2) = 0$$

Thus for small optical depths the colour is as for the case of the luminosity behind the extincting material, but reduced by the factor p. On a two colour digram these points will lie along the normal extinction vector but their displacement will be reduced by the factor p. For larger optical depths the points will lie below the vector eventually returning to the origin when the extincting sheet is effectively opaque at all three wavelengths. The two colour plot for J-K and H-K is shown in Fig. 6.4a for different values of p. It is seen that any point below the vector may be reached with suitable values of τ and p. For values of τ_J up to unity, the measured colours will lie close to the straight line vector for all values of p. For larger optical depths this is only the case for values of p close to unity, i.e. when the luminosity is mostly behind the extincting material. (Fig. 6.4a shows as a dotted line the locus as p increases for $\tau_J = 2$.)

Both these examples show that in composite stellar systems, such as galaxies, the extinction vector defines the upper boundary of points that may actually be measured. This is true in all possible geometries since, as the amount of extincting material increases, the shortest wavelength will be the first at which the assumption of optical thinness ceases to be valid. The measured points will therefore fall below the vector, i.e. become bluer in J-K. As the amount of extincting material increases still further the second wavelength becomes optically thick and the measured points will be displaced to the left, i.e. become bluer in H-K eventually returning to the origin representing the case where the extinction is opaque and the observed intensity comes only from the unextincted portion of the luminosity.

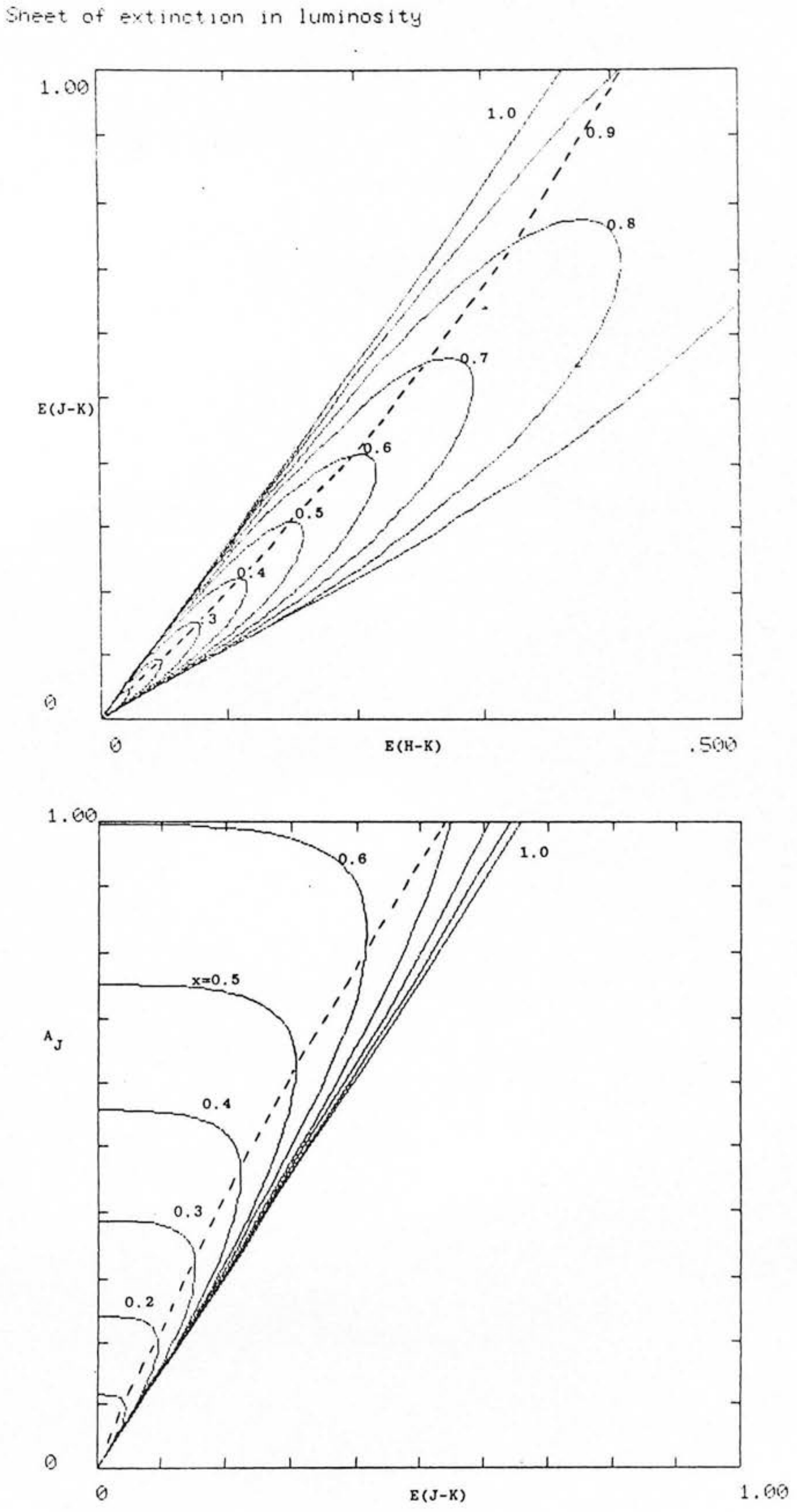


Fig. 6.4 Loci on the two colour diagram (a) and the extinction colour diagram (b) for different values of p, the proportion of the luminosity behind the extinction. Each locus shows the effect of increasing optical depth. The dotted line is the locus of varying p for constant optical depth, $\tau_J=2$.

6.2.4 Effect of Clumping

The previous two examples assume that the dust is evenly distributed and is not concentrated into clumps. It is known that interstellar dust is clumpy by the observed variations of extinction in the galactic plane. Discrete clouds also show internal clumping and dense molecular clouds exist in complexes with lower density material separating the molecular cores (Section 7.5). If the extinction is used to derive column densities from observations with beams which do not resolve the density variations and an even distribution is assumed, then the amount of extincting material will be underestimated.

An estimate of the amount of material producing extinction can be made by performing calculations on model clouds of simple structure. The case considered here is a spherically symmetric cloud centred in the beam. It is assumed that the density declines away from the centre as a power law of index α and that outside the radius corresponding to the size of the beam the density is zero. This last condition assures that the mass of the cloud is finite for all values of α less than 3.

The density at the centre of the cloud and hence the optical depth along the line of sight in the centre of the beam will be infinite for positive values of α but the total amount of material will be finite for values of α less than 3. The extinction is calculated for a fixed amount of material as a function of α, which takes values between 0 and 3. The extinction is normalised to that which would be produced if the material were distributed in a thin disk perpendicular to the line of sight. For α equal to zero the extinction is then 0.75, not unity since the material is distributed in a sphere.

The luminosity will be assumed to lie completely behind the cloud and be evenly distributed. The radius of the cloud is R (see Fig. 6.5) and the density at this point is ρ_R. The density at points inside the cloud $\rho(r)$ is then given by

$$\rho(r) = \rho_R \left(\frac{r}{R}\right)^{-\alpha}$$

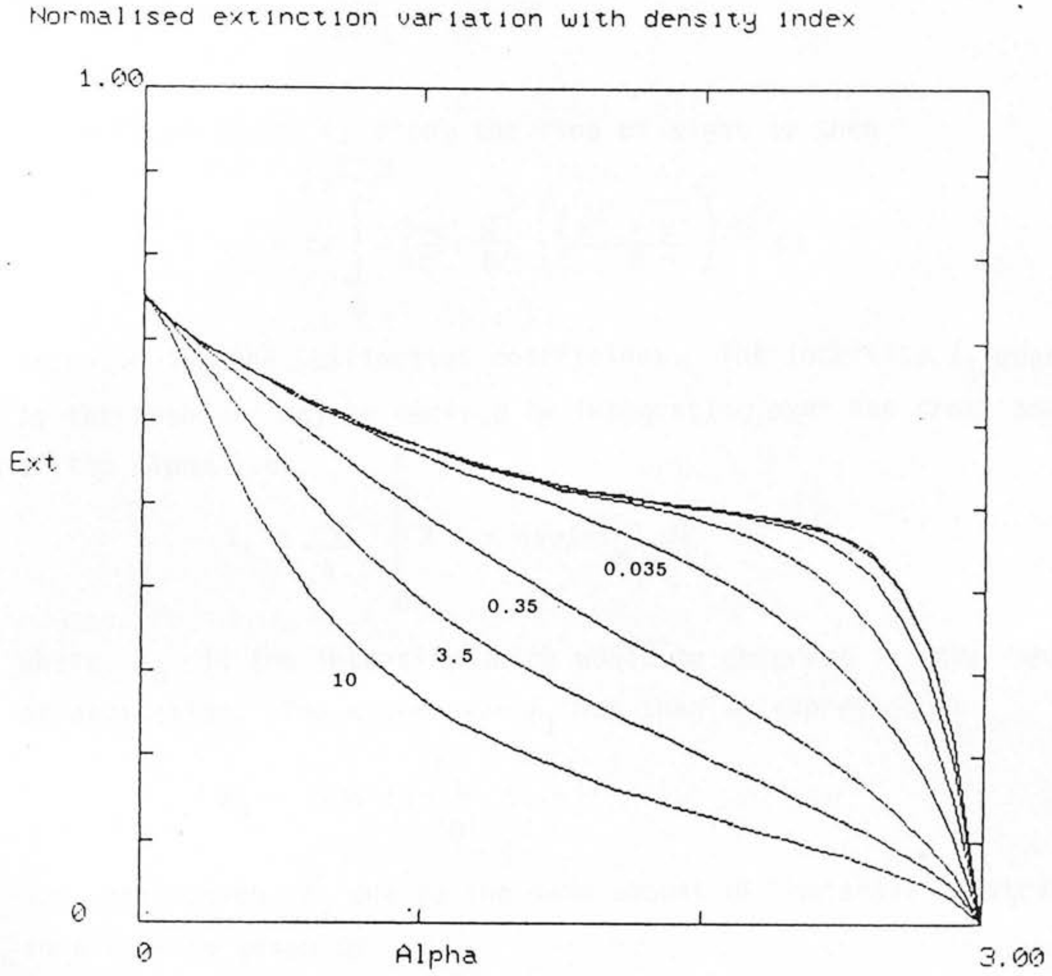


Fig. 6.5 The extinction produced by a spherical cloud whose density varies as $r^{-\alpha}$ out to radius R, beyond which it is zero. The extinction is normalised to that which would be produced if the same material were distributed in a disk of radius R. Each curve shows the effect of increasing central concentration on a fixed amount of material. The curves are marked with the optical depth of the disk.

and the total mass M of the cloud, integrating over radius, is

$$M = \left(\frac{4\ \pi\ R^3}{3-\ \alpha}\right)\rho_R$$

The density $\rho(r)$ can then be written in terms of the mass i.e.

$$\rho(r)\ = \frac{3-\alpha}{4\pi\ R^3}M\ \left(\frac{r}{R}\right)^{-\alpha}$$

The optical depth τ_x along the line of sight is then

$$\tau_x\ = 2\kappa \int_0^{\sqrt{R^2-x^2}} \frac{3-\alpha}{4\pi}\cdot\frac{M}{R^3}\ \left(\frac{\sqrt{x^2\ +\ y^2}}{R}\right)^{-\alpha}\ dy$$

where κ is the extinction coefficient. The intensity I_1 measured by the observer may be derived by integrating over the cross section of the cloud i.e.

$$I_1\ = \frac{I_0}{\pi\ ^2}\int_0^R 2\ \pi\ x\ \exp(-\tau_x)\ dx$$

where I_0 is the intensity which would be observed in the absence of extinction. The extinction A_1 may then be expressed as

$$A_1\ = 1.09\ \ln\frac{I_1}{I_0}$$

The extinction A_2 due to the same amount of material distributed in a disk is given by

$$A_2\ = 1.09\ \ln\left(\frac{\kappa M}{4\ \pi\ R^2}\right)$$

The ratio A_1/A_2 depends on both α and the product κ M. It is plotted against α in Fig. 6.5 for different values of κ M. The curves are marked with the extinction which would be produced if the same material were distibuted in a uniform disk extending just to the limits of the beam. For each curve unity on the abcissa then corresponds to this extinction.

As the material becomes more centrally concentrated (increasing α) the extinction falls but at a rate dependent on M. For small values of κM the cloud is optically thin over most of its area, but as κM increases an increasingly larger region in the centre of the cloud becomes opaque. For an index α of -2 and measured optical depths of between 0.1 and 1.0 the total amount of dust would be underestimated by a factor of about 4 times. This is a worst case since it represents a single centrally concentrated cloud whose centre appears in the centre of the observation beam. In practice it is likely that most of the material will not be centrally peaked in the beam in this way. The density of material across the beam will tend to vary more slowly and so the measured dust densities will be closer to the true values.

This dependence on α also means that the observed colours increase at different rates as the amount of extincting material increases. The reddening vector is therefore not linear and the ratio of total to selctive extinction is not constant. These effects have been analysed by considering the model cloud which represents the probable worst case. For different values of the index α , the amount of extinction has been calculated for J, H and K as κM increases. The effect on the two colour diagram is shown in Fig. 6.6a. As α increases, the locus falls below normal reddening vector. (A similar effect can be obtained by calculating the locus between two points on the reddening vector as the relative mix of the two colours is changed.)

For E(H-K) of 0.5 magnitudes the locus drops below the reddening vector by 20 percent for an index α of 2. As mentioned above this is a worst case but it does indicate that the position of points on a two colour diagram cannot be interpreted unambiguously to derive a reddening vector. For data with small measurement errors the vector can however be derived by fitting an upper envelope to the data. It is interesting to note that as the cloud becomes optically thick at all wavelengths the loci start to approach the reddening vector again.

A similar effect is seen in Fig. 6.6b which shows the relationship between the extinction in a single passband and the colour excess. As α increases loci move upwards on the diagram. For α of 2 the shift in A_H is 25 percent for an E(H-K) of 0.5 and for A_J about 20 percent for an E(J-K) of 1 magnitude. As for the two colour diagram this is a worst case. The effect of clumping then is in the

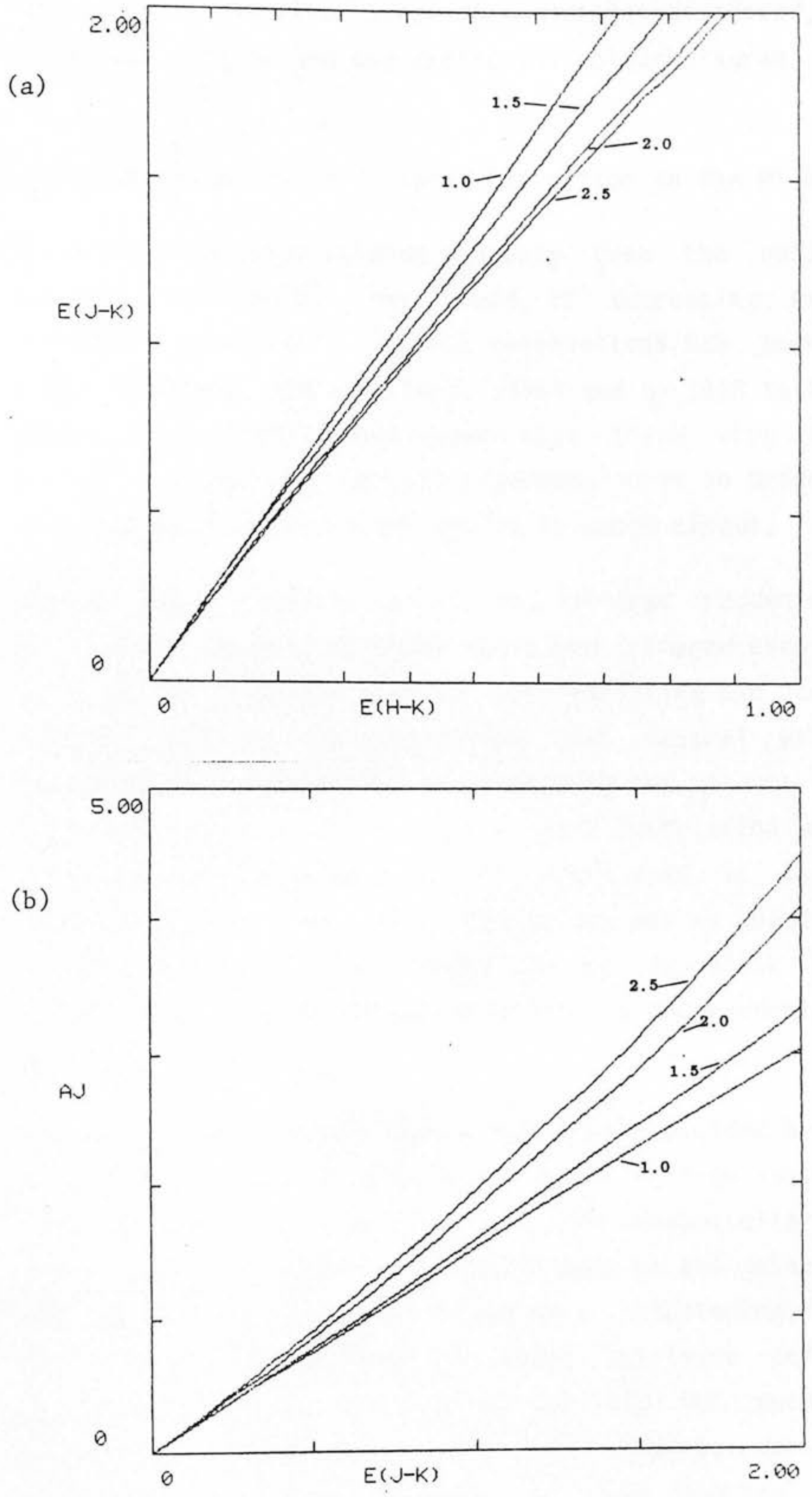


Fig. 6.6 The effect on the two colour diagram (a) and the extinction colour diagram (b) of viewing luminosity through the model spherical cloud. The curves are plotted for different values of index α, which represents the degree of central concentration. The normal reddening vector is close to the $\alpha=1$ locus.

same sense as the other types of geometries discussed, both on the two colour daigram and the extinction colour diagram.

6.3 Observations of Infrared Extinction in the Milky Way

The extinction curve extends smoothly from the optical into the infrared continuing the trend of decreasing extinction with increasing wavelength. By 1945 observations had been extended to 1 μm (Stebbins and Whitford, 1945) and by 1958 to 2 μm (Whitford, 1958). These observations showed that there were apparent local deviations from the general reddening curve in Orion, which it was proposed were due to larger grains in dense clouds.

Johnson (1968) compiled optical and infrared reddening curves for many clouds, several of which exhibited infrared excesses from which large values of R were deduced. Extrapolating the curve from the optical led to the conclusion that neutral extinction, i.e. independent of wavelength, was affecting the colours differently in different regions and that there was a correlation between galactic longitude and the value of R. The stars used in this study were associated with the clouds themselves and as suggested by Johnson could be exhibiting excess radiation due to black body emission, rather than demonstrating variations in the reddening properies of the dust.

The problem of embedded sources was partly avoided by Lee (1970) who made optical and infrared measurements of high luminosity M stars. Although supergiants can also have warm circumstellar dust shells, their high luminosity and confinement to the galactic plane means that they can be observed through more intervening extinction and that small differences in their intrinsic colours are less significant. More importantly Lee did not need to make any assumptions about their intrinsic colours, deriving reddening vectors on two colour diagrams for each spectral class. These vectors were consistent with each other so demonstrating the absence of systematic errors. These data did not confirm the galactic longitude dependence or neutral extinction proposed by Johnson, which were identified as very localised phenomena. The data related the optical to infrared parts of the extinction, a relationship which is still in some doubt. The accuracy of the data is such that E(V-K)/E(B-V) can be well determined but not purely infrared ratios such as E(J-H)/E(H-K).

Many recent studies of galactic infrared extinction have been concerned with studying star formation in dark clouds. This has involved searches for embedded stars visible only in the infrared, which can be dereddened for studies of their intrinsic properties and to derive distances. Studies of variations of grain properties within clouds can also be made using measurements of the infrared extinction curve. Measurements of heavily extincted objects, either embedded in or behind the clouds, allow more accurate determinations of the infrared extinction curve. Lee's data extended to E(J-K) of only 0.8 magnitudes whereas more recent data extend up to a value of 3 and occasionally up to 6 (e.g. Jones and Hyland, 1980 and Elias, 1978b,c). The most suitable wavelengths for such studies are J, H and K since at longer wavelengths, e.g. L, thermal emission from dust can become significant and extinction curves are subject to errors (Glass,1976).

One of the clouds most studied in this way is the ρ Ophiucus complex. Carrasco et al (1973) used infrared measurements to determine R for several embedded sources and found that it appeared to increase with optical depth. From this they inferred that the grains were larger in the highly obscured regions shielded from radiation but Cohen and Kuhi (1979) argued against this on the basis that T Tauri stars would evaporate the mantles off nearby large grains and showed that the effect was more likely caused by errors in the assumed colours of the embedded sources. To make estimates of the extinction curve using stars whose intrinsic properties are well known, Elias (1978c) measured the optical and infrared colours of a large sample of stars and obtained CVF spectra to distinguish late type giants by their CO absorption. His model (Elias 1978a) of the stellar distribution in the disk showed that most background objects would be late type giants and so he was able to differentiate between embedded and background objects. He found that the background objects did not exhibit variable reddening and he derived the excees colour ratio E(J-H)/E(H-K)=1.60 ±0.04. The variable reddening reported by Carrasco et al was evidently a result of the properties of the sources themselves and not the nature of the dust elsewhere in the cloud. Scattering could also cause the embedded sources to exhibit a different extinction law (Jones and Merrill 1976), but only if a significant amount of extinction is close to the star.

Similar anomalous reddening was reported in other clouds. Glass and Penston (1975) estimated E(J-H)/E(H-K) for the R Cr A association and deduced a value of 2.7, much larger than the theoretical value of 1.6 from the van de Hulst curve 15. This value was uncertain due to lack of knowledge about the intrinsic colours of the sources and no firm conclusions about the dust properties could be made. The peculiar reddening in the Cha T association reported by Grasdalen et al (1975) was shown to be present only in particular stars and the overall reddening was consistent with the theoretical curve. The anomalous reddening in clouds reported by Grasdalen (1974) and Rydgren et al (1976) was shown to be due to the embedded sources (Elias, 1978c) with the background stars obeying a normal extinction law.

The power of the (J-H), (H-K) two colour diagram in identifying embedded sources and background sources, from which could be derived the reddening law, was demonstrated by Hyland (1981). He plotted Elias's data (1978b and 1978c) for the Taurus and ρ Oph clouds and showed that the background stars lay along a well defined reddening vector whereas the embedded sources exhibited H-K excesses for a given J-H colour. For the Taurus cloud particularly, the embedded sources lay along a well defined sequence, the origin of which was obscure. The Oph embedded sources displayed more scatter but the background stars lay along the same reddening vector indicating that clouds with different properties can display the same extinction law for background sources.

Another source for which reddening is important and which may be different to the dark clouds mentioned above is the Galactic Centre region. Becklin et al (1978) derived an infrared extinction law by comparing the colours of Galactic Centre sources, likely to be giants and supergiants, with the colour of the general background. A value of 1.5 was deduced for E(J-H)/E(H-K), close to the theoretical value of 1.6, but this was determined from a single source whose intrinsic J-H colour was extrapolated from its H-K colour. Their data showed however that the infrared extinction law in this region is similar, within the errors, to that observed elsewhere.

The first accurate measurements of the excess colour ratio E(J-H)/E(H-K) were made by Elias (1978b and 1978c), who deduced values of 1.56±0.05 and 1.60±0.04 for the Taurus and ρ Oph clouds respectively. This value did not vary between the clouds or with position in the cloud and agreed well with the theoretical value of 1.6. A different value of 2.09±0.10 was derived by Jones and Hyland (1980). They included data from a number of sources, selected to give an upper left envelope on the two colour diagram. Some of the sources were very red and these were taken from observations of the Bok globule in the Southern Coalsack (Jones et al 1980), from sources detected in a search for OH IR stars and from IRS 11 in the Galactic Centre.

The advantage of choosing data that define an upper envelope is that the dependence on intrinsic colour is removed since the upper envelope is likely to be made up of reddened K and M giants. In the presence of scatter it will tend to lead to a higher value of E(J-H)/E(H-K) however. Jones and Hyland find that their data are compatible with an upper left envelope fitted to the data of Elias (1978b and 1978c), taking into account a factor of 0.92 for the different filter systems used. Since Elias was able to determine the intrinsic colours of the reddened stars such a procedure does not seem necessary especially as a mean line of smaller slope seems to fit the background stars well (Hyland 1981). In defence of the application of an upper left envelope to such data, effects which may alter the colours of the sources are likely to move them below the true reddening vector. These effects include thermal emission from dust and gas, scattering of the shorter wavelength radiation into the beam and uncertainties in the spectral types of the stars.

Lower values consistent with the theoretical value are also given by Koorneef (1982), who quotes data obtained by Glass giving a value of 1.68 accurate to a few percent. Observations of early type stars (Whittet and van Breda 1980), although containing a large amount of scatter, suggest a value of about 1.7. In a compromise of the existing data Koorneef (1983b) obtains a value of 1.7. The theoretical value of 1.6, although cast in doubt by the observations of Jones and Hyland, has not been invalidated in general. The possibility exists that there may be systematic variations in the value of E(J-H)/E(H-K). Although care has been taken by different workers in transforming relationships between different filter systems, the possibility also exists that systematic errors exist when comparing data especially when very red colours are involved.

6.4 Previous Determinations of Total Extinction in the Milky Way

Determination of the total extinction in a single wavelength band is much more difficult to deterimine than excess colour ratios, since to be determined directly the intrinsic luminosity and distance of the source must be known. In practice other assumptions are made and R, the ratio of total to selective extinction, is commonly estimated by using infrared measurements to extrapolate the reddening curve to zero reciprocal wavelength. Other methods of determining total extinction include observing variable extinction across clusters and comparing diameters of clusters with their brightness.

The determination of R, or other similar ratios such as A_V/E(V-K) or A_J/E(J-K), is useful since it allows sources to be dereddened from their observed colour and spectral type. The dereddened brightness then allows the distance to the source to be detemined. If the source is associated with a cloud, or if several sources can be observed in front of and behind a cloud, then the distance to the cloud can be determined. R is related to the grain characteristics and so can serve as a useful parameter by which to identify grain differences within and between clouds (e.g. Johnson 1968 and Carrasco et al 1973).

The earlier infrared measurements (e.g. Lee 1970) were useful in extrapolating the extinction curve but were not accurate enough to allow estimates to be made of the total extinction in the infrared. Where it is required to know the total infrared extinction the theoretical van de Hulst curve 15 has usually been fitted to the data (e.g. Becklin 1978) even though there has been little evidence that it is applicable in the infrared. Total extinction is often expressed in V magnitudes even though the sources concerned are frequently unobservable in the visual (e.g. Becklin 1978). The earlier data allow this to be determined with reasonable accuracy, e.g. from Lee's data $A_V/E(J-K)$ equals 6.1±1.5. The source S-1 in ρ Oph has been used by both Vrba et al (1975) and Elias (1978b) to give values for $A_V/E(J-K)$ of 4.9 and 4.6 respectively which are smaller than the theoretical value of 6.3. Considering the differences in E(J-H)/E(H-K) for embedded and background sources such a difference in the value of $A_V/E(J-K)$ for a single embedded source may not be surprising.

Estimates of the total extinction in the infrared have usually been made by using the theoretical curve. This curve can be represented in the infrared by a power law

$$A_\lambda \quad \alpha \quad \lambda^{-\eta}$$

of index -1.9 whereas Jones and Hyland (1980) derive an index of -2.7, correcting for the bandpasses of their filters. The ratio $A_J/E(J-K)$ may be derived from these two estimates, giving values of 1.70 and 2.14 respectively. The value deduced from the observational data depends on the assumed extinction law and so colour excesses cannot be used derive total extinction directly. The ratio can be estimated very approximately from Lee's data which gives a value of 1.67 ±0.68 although if a power law were fitted to the data lower errors could be derived. A single source in the Taurus cloud was used by Elias (1978c) to set limits on $A_K/E(H-K)$, which he deduced lay between 0.9 and 1.8. By assuming an extinction law index of -1 he deduced a value of 1.1 for this ratio. As Whittet and van Breda (1980) point out this index is assumed to have a wide range of values varying between -1 (e.g. Elias) and -4 for the small grain Rayleigh approximation.

Extinction measurements both in the optical and infrared are frequently referred to A_V using relationships involving measurements at B. As Hyland (1981) points out, observations of anomalous reddening in the optical can be interpreted solely as a consequence of variations in the B extinction. Derivation of A from E(B-V) can therefore lead to incorrect conclusions and he suggests that A_V /E(V-K) should be used instead since E(V-K)is less dependent than E(B-V) on variations in the grain population. For normal extinction, optical and infrared excess colours are related (Whittet and van Breda 1980) by

$$R = A_V/E(B-V) = 1.1\ E(V-K)/E(B-V)$$

For heavily extincted sources V measurements are not easily made, nor relevant in many cases and so it is desirable to obtain purely infrared ratios of total to selective extinction.

6.5 Interstellar Extinction Law in Galaxies

That extinction is an important factor in the observable characteristics of galaxies is obvious from their visual appearance, particularly in unusual galaxies such as NGC 253 and M 82 but also in the pronounced dust lanes and spiral features of normal spiral galaxies. The dust plays an important role in the structure as well as the appearance of galaxies, since it is linked to the star formation process by shielding clouds from radiation and acting as a cooling agent.

The extinction curve has been observed in the Magellanic Clouds which have a low dust density, allowing the curve to be determined in the optical and ultraviolet. The ultraviolet law has been observed to be different (e.g. Nandy et al 1981) but the infrared reddening appears similar. Koorneef (1982) related the infrared to optical reddening, deriving a value of 0.65 with large errors for E(J-K)/E(B-V). Morgan and Nandy (1982) also found the infrared extinction to be similar to that of the Milky Way and derived a value of 1.36 ±0.19 for E(J-H)/E(H-K), rather lower than but still consistent with the theoretical value and the value derived observationally by Elias (1978c). They concluded that the different populations of small grains, which cause the differences in the ultraviolet extinction curve, have little effect at longer wavelengths.

Peculiar geometries can be useful in determining the extinction properties of dust. Lynds (1970) observed a background galaxy through a dust lane in M51 to estimate the dust density. In a similar way Keel (1983) determined R for the dust in NGC 3314 (really two galaxies along the same line of sight) by observing the Sb component through the foreground Sc. The foreground dust caused dips in the brightness profile of the background galaxy and also colour variations. Relating these allowed R to be derived giving a value of approximately 3.5 in agreement with the Milky Way value.

The geometry of a foreground dust lane against background bulge light has been used to estimate the ratio of total to selective extinction in a few galaxies. Lindblad (1942) measured A_V/E(B-V) in NGC 7331 by comparing the colour and magnitude variations on the near and far sides of the galaxy. He concluded that the value was similar to that in the Milky Way and that small grains were also responsible for the extinction in NGC 7331. Hodge and Kennicutt (1982) estimated A_B/E(B-V) to be 4.3 0.6 and Hoessel and Melnick (1980) derived a value of 2.7 0.2 for A_G/E(G-R). These values are also in good agreement with the Milky Way values. The same technique will be used in Section 6.8 to derive infrared ratios of total to selective extinction for some of the galaxies observed.

Aaronson's (1977) multiaperture infrared observations of galaxies included some with anomalously red colours which he attributed to extinction but no deductions about the nature of dust in galaxies could be made. In a similar study of galaxies selected to minimise the effects of extinction, Greirsmith et al (1982) also found some which showed signs of reddening by dust. Two of these were observed by Jones et al (1983) at selected points in and out of the dust lane. They found that the J, H and K colours were consistent with the direction of the reddening vectors deduced for the Milky Way. The observational errors were such that both the vectors of Elias (1978c) and Jones and Hyland (1980) fitted the data.

Spiral galaxies have strong dust features but some early type galaxies also exhibit dust lanes (e.g. Hawarden et al 1981). The nature of dust in such galaxies is interesting because its history may be very different to that of spiral galaxies. The dust is not extensive enough to affect the integrated infrared colours (Longmore and Sharples 1982) but surface photometry shows colour variations. Visual surface photometry of IC 4320 (Warren Smith and Berry 1983) indicates a very low value of R (1.9) equivalent to a strong wavelength dependence for the extinction. This they attributed to the absence of large grains.

By analogy with the Magellanic Clouds it is possible that the infrared extinction law for such galaxies is normal but this has not been investigated except for the peculiar galaxy NGC 5128. J, H and K measurements (Harding et al 1981) through the dust lane produced points on the two colour diagram which lie below the Jones and Hyland reddening vector. This was attributed to the star dust mixture and the unknown infrared scattering properties of dust. The dust lane was also observed at wavelengths less that 1 μm (Kunkel and Bradt 1971, Rodgers 1978) and an extinction law different to that of the Milky Way was derived. It diverged most between 0.7 and 0.9 um rather than in the short wavelengths where deviations might have been expected. This discrepancy and differences in the estimate of A_V were attributed by Harding et al to saturation effects. The fact that the discrepancy appeared at the longer wavelengths, where the saturation effects would be least, may have been due to the normalisation process. These problems with visual data illustrate the advantages of infrared observations where optical depths and scattering have less effect.

The properties of some individual clouds in late type spirals have been investigated by Elmegreen (1980) using multicolour high resolution data. She found that the internal properties of the clouds and their location above the plane could be determined using such data together with radiative transfer models. The denser clouds were confined to the plane and the densest occured on the inner edges of spiral branches. There was evidence for high latitude clouds (150 to 300 pc) of relatively low extinction. These had an average visual extinction of about 2 compared with 3 for the spiral arm features and up to 10 for the densest features. The densest clouds observed had densities of about 150 cm^{-3} similar to but not as dense as the molecular clouds in the Milky Way. The association of young stars may have made denser clouds less obvious, by decreasing their contrast relative to the background. These types of clouds are well suited to near infrared observation since the colours are likely to be easily observable without the optical depth in the infrared becoming excessive. The existence of high latitude clouds makes the interpretation of disk colours difficult but their location will not affect the colours of the bulge seen through the disk or the colours of the bulge bisected by a dusty disk.

6.6 Effect of Scattering in the Infrared

When the extinction and luminosity are widely separated the extinction coefficient is simply the sum of the absorption and scattering coefficients. This is because effectively none of the scattered radiation reenters the observation beam. When the dust and luminosity are closely associated then some of the scattered radiation can reenter the beam. An extreme example of this is circumstellar dust shells for which all of the dust is contained within the beam; in this case no radiation is removed from the beam by scattering. Scattering may also cause more radiation to enter the beam than is removed so producing reflection nelulae. If significant scattering will alter the observed colour and extinction ratios.

The effect of scattering is very dependent on the particular geometry of the source and even in a simple spherically symmetric case the radiative transfer problem is very complex requiring, for example, the use of Monte Carlo computational techniques (Witt and Stephens 1974). In galaxies the dust and stars are mixed in a variety of geometries and in certain situations it might be possible for scattering to affect the observed colours. The albedo of grains decreases with increasing wavelength; if the particle size is small compared to the wavelength then scattering cross sections decrease as the fourth power of the wavelength (e.g. Martin 1978). For example, for particles of characteristic size 0.2 um with refractive indices corresponding to dirty ice the albedo at J is less than half that at V; typical values (Martin 1978) are given in Table 6.1. Wickramasinge (1972) computed albedos for other possible types of interstellar grains including graphite and dirty silicates. For these also the albedo in the infrared is much lower than in the optical although for pure silicates it is large (0.79) even at 1.5 µm.

Wavelength	$x = 2\pi a/\lambda$	albedo for n=1.33-0.1i	albedo for n=1.33=0.33i
U	3.5	0.62	0.43
B	2.9	0.59	0.41
V	2.3	0.54	0.37
J	1.01	0.25	0.16
H	0.76	0.16	0.11
K	0.57	0.07	0.05

Table 6.1 The albdeo for grains of size a = 0.2 um for two refractractive indices. The data are derived from Martin (1978, P 63) and is shown for the normal optical and infrared filter bands.

Since the albedo of grains is significantly less in the infrared than in the optical, scattering can be ignored if it can be shown that it is not significant in the optical in similar situations. A consequence of the lower albedo in the infrared is the rarity of infrared reflection nebulae in situations where they could be expected in the optical. A few reflection nebula have been observed e.g. in ρ Oph (Elias 1978b) and in OMC2 (Lee et al 1981) but they are faint compared to the brightness of other infrared sources in such objects. This contrasts with the optical appearance of clouds, which are often dominated by reflection nebulosity, and supports the view that scattering can be neglected if it does not dominate the optical appearance.

The data, from which quantitative deductions about the extinction will be drawn, represent observations of the bulge luminosity seen through dust in the disk. The extinction and dominant luminosity are therefore fairly widely separated and the geometry is roughly similar to that of the double galaxy NGC 3314. Calculations by Keel (1983) show that the proportion of scattered light, in a geometry similar to that of a bulge seen through a disk, is not significant.

Warren Smith and Berry (1983) argue that scattering in the optical is unlikely to be significant for the dust lane of IC 4320. They find that the extinction is very wavelength dependent, whereas scattering would make it less wavelength dependent than normal by increasing the shorter wavelength radiation. Also scattering would cause the apparent extinction to vary across the dust lane and this they did not observe. Calculations they performed on proposed grains reinforced this view. The absence of scattering in the optical for such situations means also that it will be insignificant in the infrared.

If scattering is significant, the effect will be to increase the observed intensity at the shorter wavelengths. This would move points on the two colour diagram below the reddening vector and above the locus on the extinction excess colour diagram. These shifts are in the same sense as produced by geometrical effects and so the interpretation of the data as limits, or the fitting of envelopes to the data will remain valid.

6.7 Determination of Colour Excess Ratio

As demonstrated in Chapter 5 a major contributor to the colour variations observed in the galaxies is extinction due to dust. NGC 4565, NGC 5907 and NGC 7331 show extinction effects large enough to allow an investigation of the properties of dust in oher galaxies. NGC 7814 and NGC 4216 also show the effects of extinction but to a lesser degree and so no H measurements were made. The J-K measurements of the former set of galaxies were augmented by H-K measurements to confirm the interpretation of the colours as extinction and to derive colour excess ratios for comparison with similar measurements obtained for the Milky Way.

The data were obtained in three passbands, J, H and K. It is customary to plot these on (J-H),(H-K) two colour diagrams and to derive the excess colour ratio E(J-H)/E(H-K). Since the independent sets of data obtained are J-K and H-K, however, the colour excess ratio E(J-K)/E(H-K) will be derived instead. This is easily related to the more usual ratio by

$$E(J-H)/E(H-K) = E(J-K)/E(H-K) - 1$$

The analysis of Section 6.2 showed that the measured ratio can be less than the true colour excess ratio due to extinction distributed within the luminosity and to clumping. The measured points therefore tend to lie below the true reddening vector which can be derived in noise-free data by fitting an upper envelope to the points on the two colour diagram. This is similar to the method used by Jones and Hyland (1980) but their reasons for fitting the envelope are somewhat different (Section 6.3).

For the method to be successful in determining the reddening vector, the intrinsic colour of the underlying luminosity must not vary significantly compared to the extinction reddening. The multi-aperture work of Aaronson (1977) and Griersmith et al (1982) shows that the intrinsic colour variations of the bulges of spiral galaxies due to population and metallicity effects are much less than the extinction reddening in the data described. Even this small effect could be removed by subtracting colours on two sides of a scan if one were relatively dust free. This has not been done because the errors (particularly in H-K) are likely to be more significant than the intrinsic variations. The errors in relative positioning between the J-K and H-K data also make this correction less useful. Where there is a significant amount of light from the disk, the two sides of a scan are no longer complementary but show similar reddening effects. The data themselves show that the colours of the bulge region are relatively flat and featureless except for the effects of extinction.

6.7.1 NGC 7331

The two colour diagrams for various sets of data from NGC 7331 are shown in Fig. 6.7. The minor axis data are taken from within 33 arcseconds of the nucleus to exclude noisy data and to ensure that intrinsic scatter due to the disk contribution is minimised. The major axis and N/S data extend to 50 arcseconds from the nucleus.

A linear regression has been performed on the data to relate E(J-K) to E(H-K). Since comparable errors exist in each set of data, two regressions, E(J-K) onto E(H-K) and E(H-K) onto E(J-K), were calculated. The slope given in Table 6.2 is the slope of the bisector of the results of the two regressions (Topping 1972).

	Data set	No. of data	E(J-K)/E(H-K) (error)	Correlation coefficient
NGC 4565	Scan 9" SE	15	1.83 (0.20)	0.927
NGC 5907	Bulge points	13	3.18 (0.47)	0.881
NGC 7331	Minor axis	13	3.02 (0.39)	0.912
	Major axis	11	1.80 (0.16)	0.964
	N/S	21	1.88 (0.22)	0.898

Table 6.2 The colour excess ratios in NGC 4565, NGC 5907 and NGC 7331 as derived from a least squares fit to selected points, which were weighted by the square of their error (assumed to be at least 0.03 magnitudes).

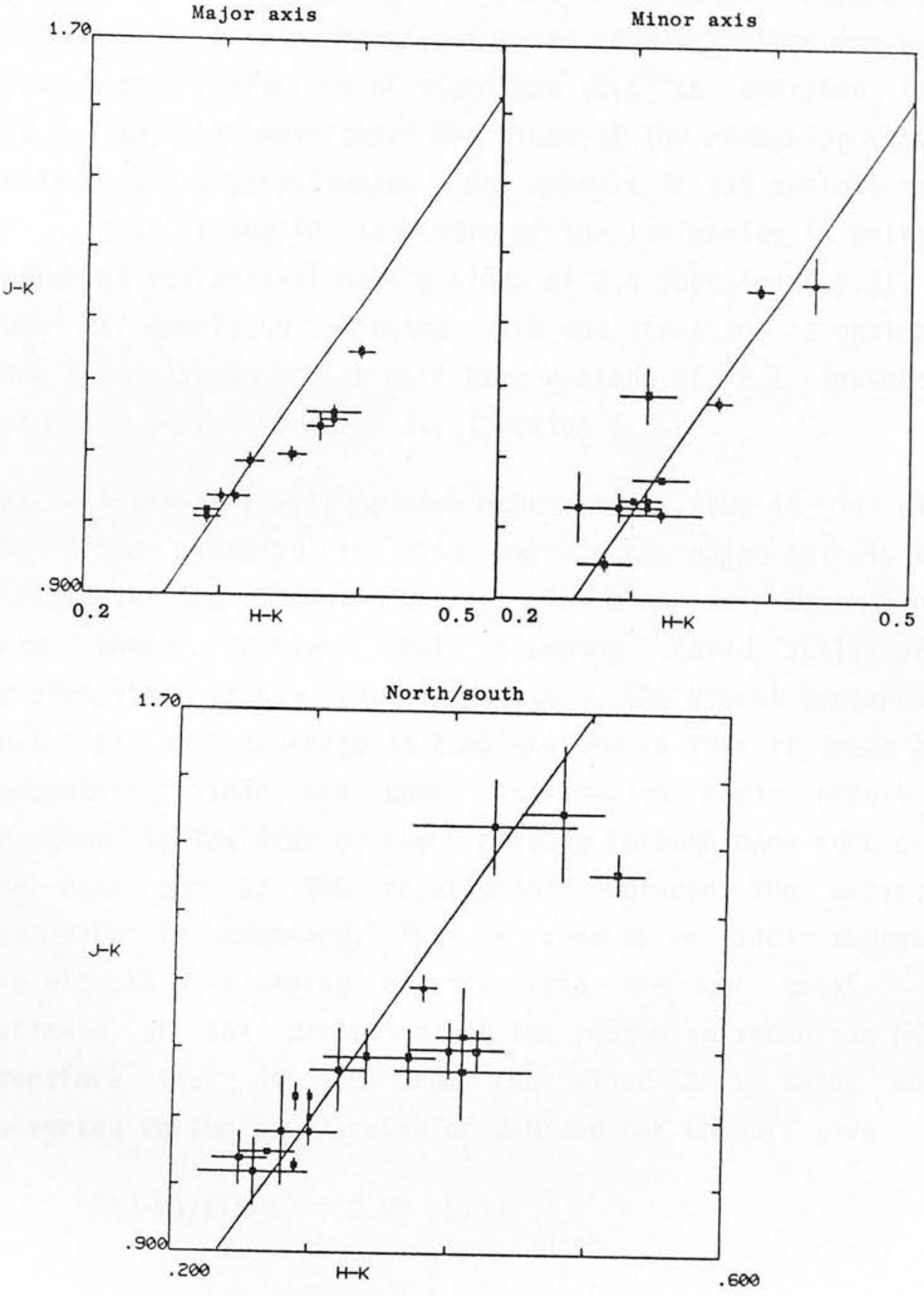


Fig. 6.7 The two colour diagram for the profiles of NGC 7331. The line representsthe best fit to the minor axis data.

The slope of the vector agrees well in the major axis and N/S scans. The minor axis data, on the other hand, provide a better estimate because the inclination of NGC 7331 is such that along the minor axis the dust is seen in projection against the bulge. Along the major axis (and the adjacent N/S scan) the dust is distibuted within the luminosity. Where the bulge is dominant the dust divides the luminosity in half (p=0.5 in Section 6.2.3). Farther out the distribution is more complex and the reddest points may approximate to an even distribution of stars and dust as analysed in Section 6.2.2. In both these cases the slope of the reddening vector can be considerably underestimated. For example if the optical depth at J of a thin cloud in the middle of the luminosity is unity then the reddening vector will have a slope of 2.4 (Section 6.2.3). If the dust is evenly distributed with the stars and is optically thick then the measured vector will have a slope of 2.1 instead of the reddening vector slope of 2.7 (Section 6.2.2).

The real geometry will be much more complex, but it is clear that the slope measured in this way for the major axis is not a true estimate of the reddening vector. The minor axis data do not suffer from these problems but clumping could still affect the determination of the reddening vector. The visual appearance shows that the dust coverage is complete and so even if there are clumpy components within the dust distribution their effect will be lessened by the line of sight passing through many such clumps. In the next section the relationship between the extinction and reddening is examined. This is found to be linear suggesting that the effects of clumping on the data are not great. The best estimate of the direction of the reddening vector in NGC 7331 is therefore that derived from the minor axis data, which when converted to the usual ratio of J-H and H-K colours give

$$E(J-H)/E(H-K) = 2.02 \pm 0.39$$

Section 6.2 shows that geometry effects cause the measured points to lie below the reddening vector. The vector derived from the minor axis data is drawn on the two colour diagrams for the major axis and N/S data (Fig. 6.7).

The two dimensional map data are plotted on two colour scatter diagrams in Fig. 6.8, restricting the data to those which have estimated errors of better than 5 percent in both colours. The derived reddening vector is drawn from the nuclear points and is seen to define an approximate upper envelope given the typical errors of the data.

6.7.2 NGC 4565

Two colour data were obtained for the perpendicular profile and are shown in Fig. 6.9. The luminosity gradients in this region are smaller than along the minor axis and so more accurate data can be obtained in the presence of small pointing errors. The shape of the colour profiles are somewhat different in J-K and H-K, with the red region being narrower in H-K. This could be the result of saturation which would be strongest in the J-K data and can be expected since the galaxy is close to edge-on and the total extinction at J is large (at least 1.1, Section 6.9).

Instrumental effects may also be responsible. Small relative pointing errors in the presence of strong colour gradients could produce different shapes of profiles. Similar effects could be produced if the chopper waveform (the time for the mirror to come effectively to rest) changed between the taking of the J-K and H-K data. If the effective beam profiles were wider when the J-K data were measured, then the region red in J-K would be broader and less intense than the region measured in H-K. A significant correlation coefficient is derived, but this could be the result of measuring different mixtures of saturated and unreddened colours (Section 6.9). The deduced ratio corresponds to a value of E(J-H)/E(H-K)

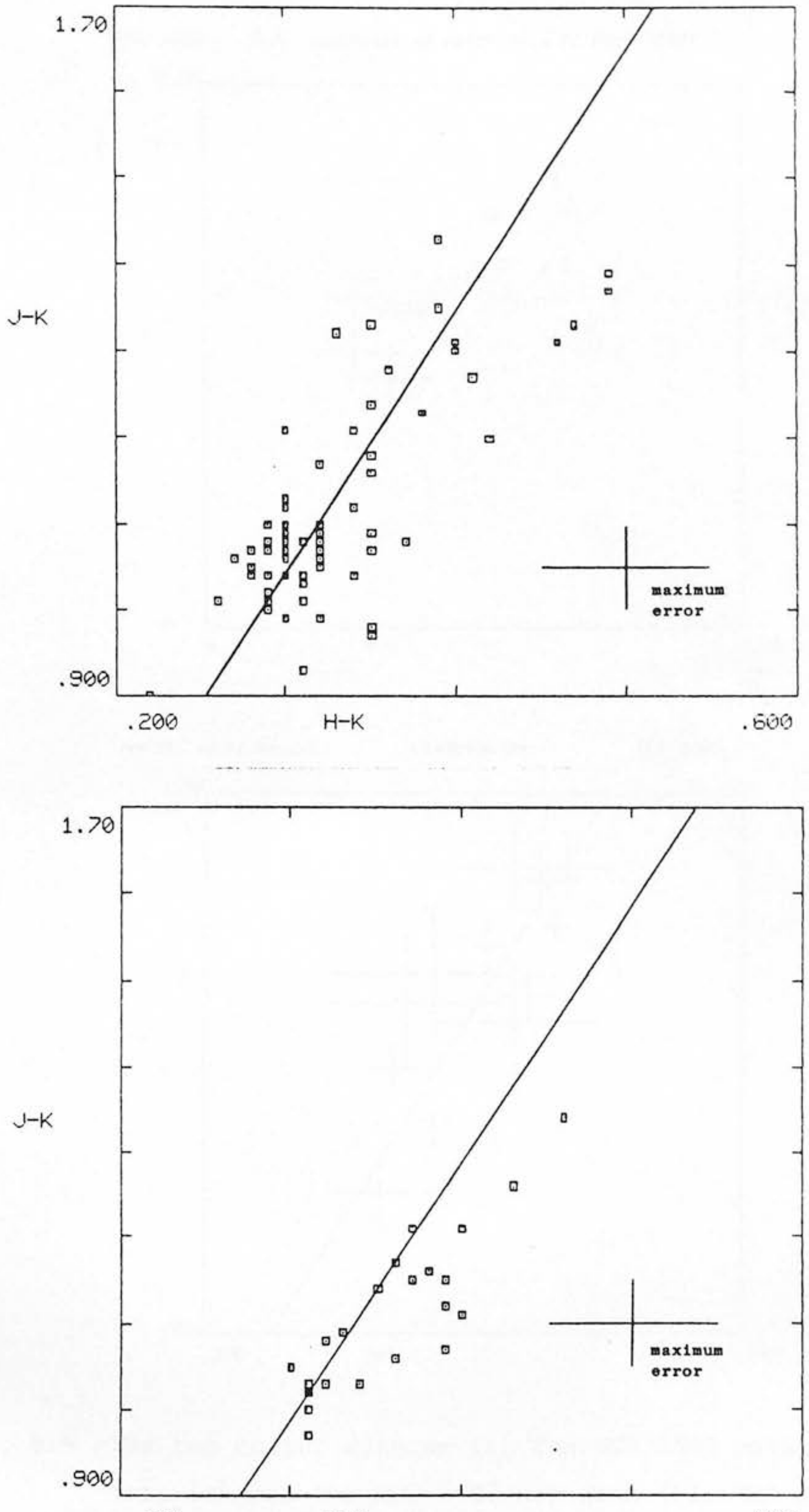


Fig. 6.8 The two colour diagram for the NGC 7331 map data. (a) shows the data from the small map and (b) shows the data from the large map. The data were selected so that the estimated errors were no larger than those shown. The lines are drawn through the points representing the nucleus and have a slope derived from the fit to the minor axis data.

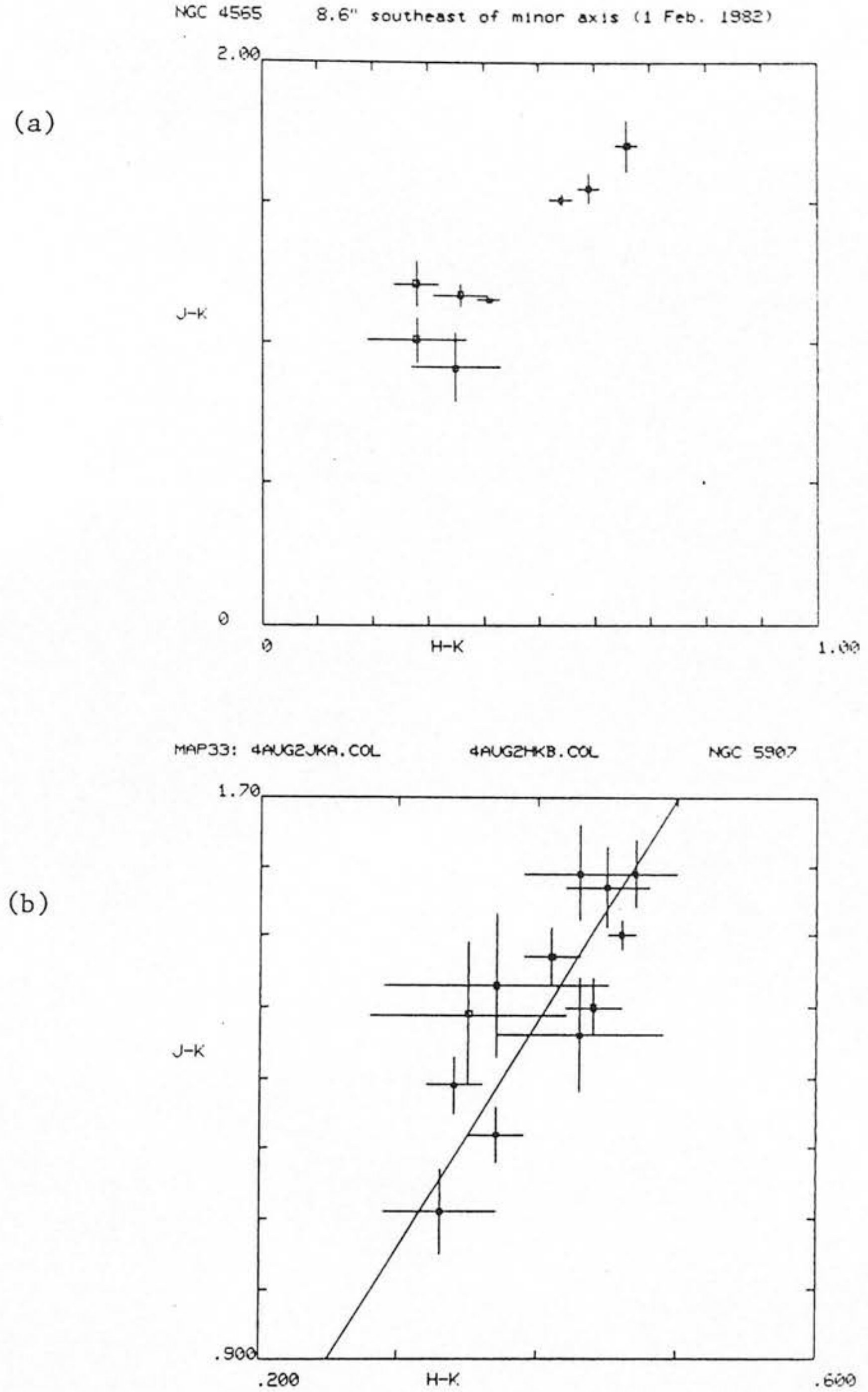


Fig. 6.9 The two colour diagram for the NGC 4565 parallel profile data (a) and the NGC 5907 map data (b). The line is the best fit to the NGC 5907 data.

less than unity; this is not consistent with an extinction decreasing with increasing wavelength and is unlikely to be a property of the dust.

6.7.3 NGC 5907

NGC 5907 was mapped in two dimensions in both J-K and H-K with a 10 arcsecond aperture. In order to define a reddening vector more accurately, the data set has been restricted to those points within 30 arcseconds of the nucleus so that the bulge light and the dust define the brightnesses and colours. The points on the edge of the nucleus are very susceptible to small telescope positioning errors and so have also been excluded. The remaining points, plotted in Fig. 6.9, have been weighted by the square of their errors and used in a regression analysis, the results of which are shown in Table 6.2. This gives the result

$$E(J-H)/E(H-K) = 2.18 \pm 0.47$$

The large error is the result of small pointing errors between the two maps. The errors in the determination of this vector and that for NGC 7331 would probably be much less if the three passbands, rather than only two, could be observed simultaneously. This is analagous to the improvement in colour accuracy obtained by using two simultaneous channels rather than using consecutive measurements.

6.8 Determination of Ratio of Total to Selective Extinction

As discussed in the Section 6.4, the ratio of total to selective extinction in the infrared is poorly determined. In contrast to the ratio of colour excess it is better known for other galaxies. This is because the unreddened brightness may be estimated by reference to unreddened regions of the galaxy under study, without assumptions about distance or intrinsic luminosity. If the galaxy is assumed to have axial symmetry and mirror symmetry about the galactic plane, then the total luminosities along two lines of sight equally above and below the major axis are equal. Visual observations of dust free S0 galaxies along directions perpendicular to the major axis (e.g. Burstein 1979) have shown that this is a valid assumption. Azimuthal features such as spiral arms will not be significant if the dominant luminosity is from the bulge.

Three of the galaxies studied, NGC 4565, NGC 5907 and NGC 7331 show large scale features due to extinction. NGC 7814 and NGC 4216 show similar but less pronounced features. In order to derive the ratio it must be possible to reflect the profiles about the major axis to obtain the unextincted brightness. The observations of NGC 4565 allow this to be done accurately. Both the minor axis profile and the perpendicular profile 9 arcseconds southeast of the nucleus (Fig. 5.11) are seen to be symmetric about their peak (the major axis point) except for dips in the profile where the measured signal is lessened by extinction. The dips are much more apparent in J than K with the H profiles being intermediate and these dips in the brightness profiles correspond to peaks in the colour profiles.

Of these two scans in NGC 4565, the perpendicular profile is more suitable for the determination of the ratio. This is because the amount of reddening and extinction can be expected to be similar in the two scans, whereas the brightness gradient is significantly less than along the minor axis. In fact greater reddening is observed in the PP. This could be due to greater dust column density or less foreground luminosity, but is more likely due to smaller dilution caused by the chopper crossing bright regions. In neither scan does the chop cross the nucleus but for the minor axis scan it does cross brighter regions than the PP. This dilution is exactly the same for both filters, due to the simultaneity of the measurements, and so does not affect the ratio in a random way. The mixing of luminosity reddened by different amounts does however cause a small variation in the ratio of total to selective extinction as shown in Section 6.2. Its effect on the derived value will be discussed below.

The differences , Δ, between the two sides of both the minor axis and the PP are plotted in Fig. 6.10, which shows Δ J against Δ(J-K). The errors are the quadrature sums of the errors of the two points constituting each difference. All the points are included except the point on the minor axis which grazes the nucleus and hence suffers from a large amount of dilution. The trend is linear confirming the impression gained by inspecting the J brightness and J-K colour profiles that red colours correspond to dips in the brightness profile. A linear regression (Table 6.3) has been applied to these two sets of data.

Data Set	Colour	No. of data	Slope	Intercept	r
NGC 4565					
minor axis	J-K	5	1.70 ± 0.15	-0.15 ± 0.04	0.989
perp. profile	J-K	6	1.76 ± 0.11	-0.12 ± 0.04	0.992
" "	H-K	6	1.97 ± 0.36	0.06 ± 0.07	0.938
NGC 7331 N/S	J-K	11	1.54 ± 0.21	-0.00 ± 0.03	0.915

Table 6.3 The results of a linear regression between the excess colours and the brightness deficits for scans in NGC 4565 and NGC 7331. These quantities are the differences in brightness and colour between points equidistant on each side of the scan. r is the correlation coefficient.

The data used in the regression have been weighted inversely with the square of their statistical error, for which a minimum value of 0.03 was assumed. The regression line has not been constrained to pass through the origin since a vertical shift is introduced if the central point of the scan does not coincide with the major axis point. If such an error is present the calculation of the slope is unaffected if the brightess declines exponentially (linear decline in magnitude). This is approximately true for these data and in any case the intercept is small.

The high degree of correlation obtained might suggest that clumping and varying mixtures of foreground and embedded luminosity are not

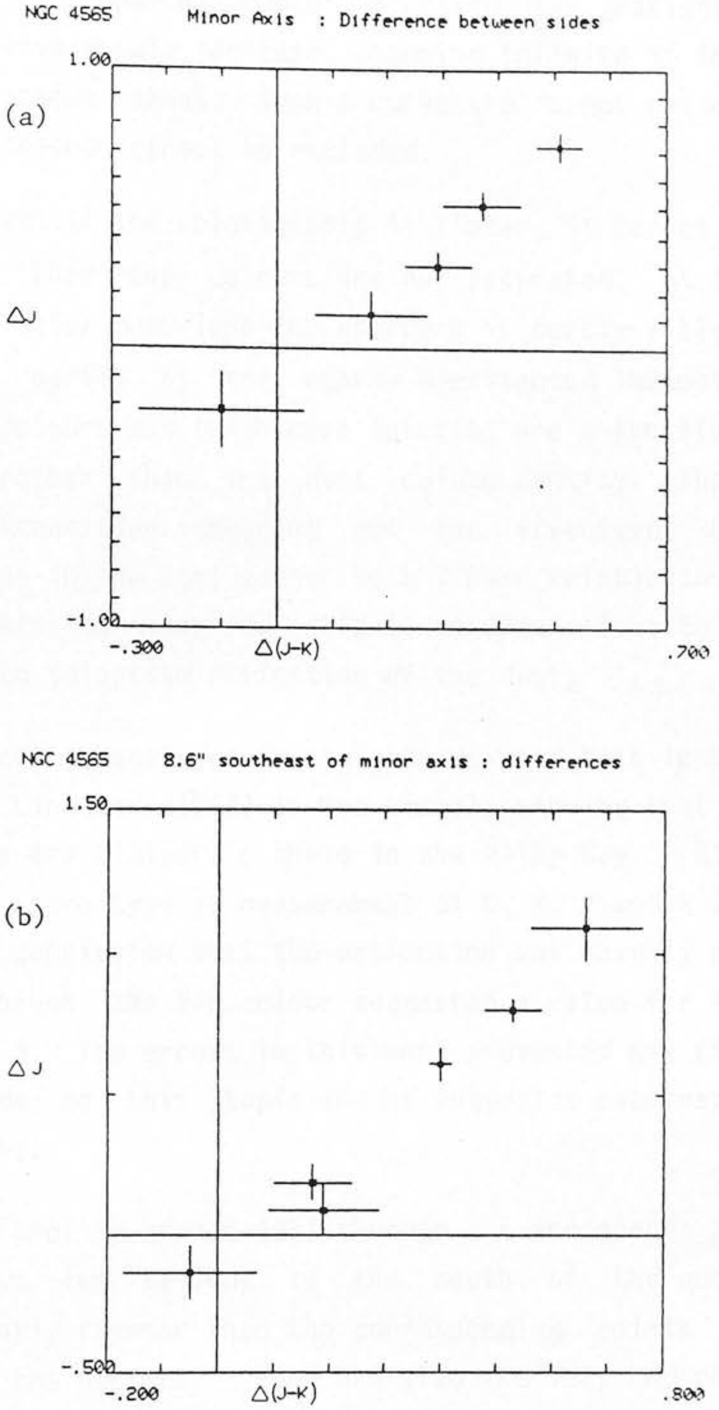


Fig. 6.10 Brightness and colour differences between the two sides of the NGC 4565 minor axis profile (a) and the parallel profile (b).

saturating the colours and affecting the ratio. If this were the case and the different values of the colours and dips in the brightness profiles were due to different dust column densities, then as the optical depth increased the gradient of the Δ J vs Δ (J-K) curve should increase, becoming infinite if the dust were in a thin opaque sheet. Such a curvature is not evident in the data but its presence cannot be excluded.

However even if the relationship is linear, it cannot necesarily be concluded that the colours are not saturated. At the edges of an optically thick dust lane the aperture is partly filled by the dust lane and partly by the nearby unextincted luminosity and so the measured colours and brightness deficits are a function of the beam filling rather than the dust column density. The effect of the chopper transition smearing out the effective beam profiles contributes in the same manner to a linear relationship. Because of these effects the measured ratio is an upper limit to the true ratio of total to selective extinction of the dust.

NGC 7331 can be analysed in the same way and this technique has been used by Lindblad (1942) in the visual, showing that the extinction properties are similar to those in the Milky Way. Blackman (1979) made the same type of measurement at U, B, V and R and came to the tentative conclusion that the extinction was roughly propotional to , although the V-R colour suggested a value for R significantly less than 3. The errors in this work prevented any firm conclusions being made on this topic and he suggested observations at longer wavelengths.

The N/S J profile of NGC 7331 through a 5 arcsecond aperture (Fig. 5.2) shows two regions to the south of the nucleus which are significantly fainter than the corresponding points on the other side of the nucleus. These are also the very red regions measured in J-K. Making the assumpion that the galaxy is axially symmetric, the J deficits can then be calculated and correlated with the J-K colour. Fig. 6.11 shows these two quantities plotted against each other and it is apparent, as in NGC 4565, that they are linearly

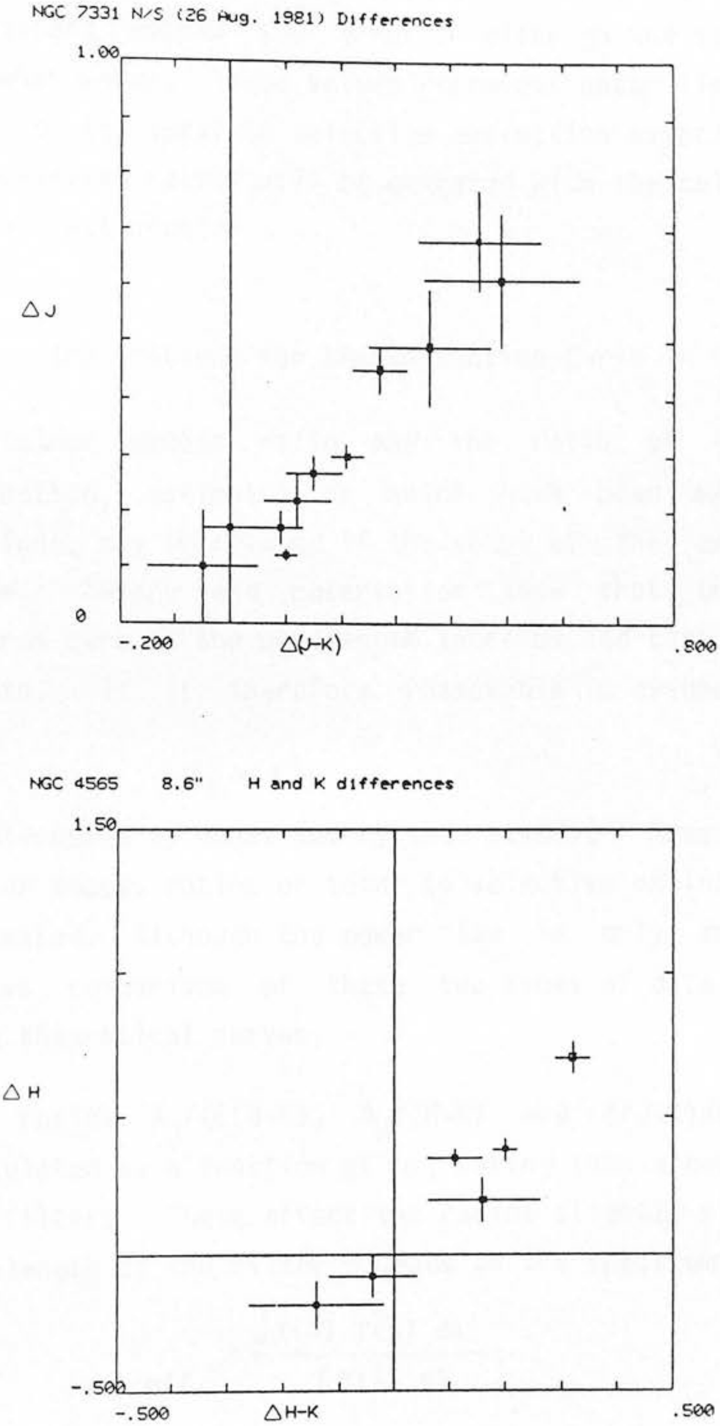


Fig. 6.11 Brightness and colour differences between the two sides of the NGC 7331 N/S profile (a) and the NGC 4565 parallel profile (b).

related.

The measured $\Delta J / \Delta(J\text{-}K)$ ratios for NGC 4565 and NGC 7331 are consistent within the errors, although the value for NGC 7331 is somewhat lower. These values represent upper limits for the true ratio of the total to selective extinction associated with the dust. The measured ratios will be compared with the colour excess ratios in the next section.

6.9 Implications for the Extinction Curve in Galaxies

The colour excess ratio and the ratio of total to selective extinction, estimates of which have been made in the previous sections, may be related if the shape of the extinction curve is known. Theory and observation show that the extinction tends towards zero as the wavelength increses and that the variation is smooth. It is therefore reasonable to assume a power law of the form

$$A_\lambda \propto \lambda^\eta$$

as discussed by Jones and Hyland (1980). Measurements of either colour excess ratios or total to selective extinction allow to be estimated. Although the power law is only an approximation it allows comparison of these two types of data with each other and with theoretical curves.

The ratios $A_J/(E(J\text{-}K)$, $A_H/(H\text{-}K)$ and $E(J\text{-}K)/E(H\text{-}K)$ have been calculated as a function of η, taking into account the passbands of the filters. These affect the ratios slightly since the effective wavelength of the filter depends on the spectrum of the source by

$$\lambda_{\text{eff}} = \frac{\int I(\lambda)\ T(\lambda)\ d\lambda}{\int T(\lambda)\ d\lambda}$$

where $I(\lambda)$ is the source intensity and $T(\lambda)$ is the transmission of the filter (and atmosphere etc.). The ratios were calculated assuming that the filter transmissions were 100 percent inside the passband and zero outside. The source was assumed to be a 3000K blackbody appropriate to cool stars. The use of other temperatures does not affect the ratios significantly. Jones and Hyland used a more sophisticated stellar atmosphere model and concluded that errors in the assumed source spectrum were of little significance.

The finite filter bandpasses can also cause non-linearities in the relationships due to the effective wavelength increasing as the radiation becomes more reddened. Jones and Hyland show that the effect is negligible up to at least E(J-H)=3, which is much greater than the reddening detected in these observations.

A_J/E(J-K) and E(J-K)/E(H-K) are plotted as functions of η in Fig. 6.12. η =0 corresponds to neutral extinction and so A_J/E(J-K) becomes very large as η approaches zero. The largest value of η considered is 4 corresponding to Rayleigh scattering by small particles. In this wavelength region the van de Hulst curve 15 (Johnson 1968) can be parameterised by η=1.7. Jones and Hyland deduce a value of 2.5 from their data on reddening through dense clouds. The colour ratios measured by Elias (1978b and 1978c) correspond to η=1.8 ±0.1, when the colour ratio is transformed onto the UKIRT system.

These values may therefore be compared with the values derived from the observations of NGC 4565, NGC 5907 and NGC 7331. The values deduced from the various sets of data are summarised in Table 6.4. The values of η are all lower limits if the colours are saturated or if clumping and beam filling are important.

Data		from A_J/E(J-K)	from A_H/E(H-K)	from E(J-K)/E(H-K)
NGC 4565	perp. profile	1.6 ± 0.2	2.5 (+0.9,-0.4)	< 0
	minor axis	1.6 ± 0.2		
NGC 5907	"bulge"			2.7(+0.8,-1.1)
NGC 7331	minor axis			2.4 + 0.7
	major axis			< 0
	N/S	1.9 ± 0.5		< 0
Jones and Hyland				2.5 ± 0.2
Elias				1.8 ± 0.1
van de Hulst curve 15				1.7

Table 6.4 The values of η (where $A \propto \lambda^{\eta}$) deduced from these data, the observations of Jones and Hyland (1980), the observations of Elias (1978c) and the theoretical van de Hulst curve 15.

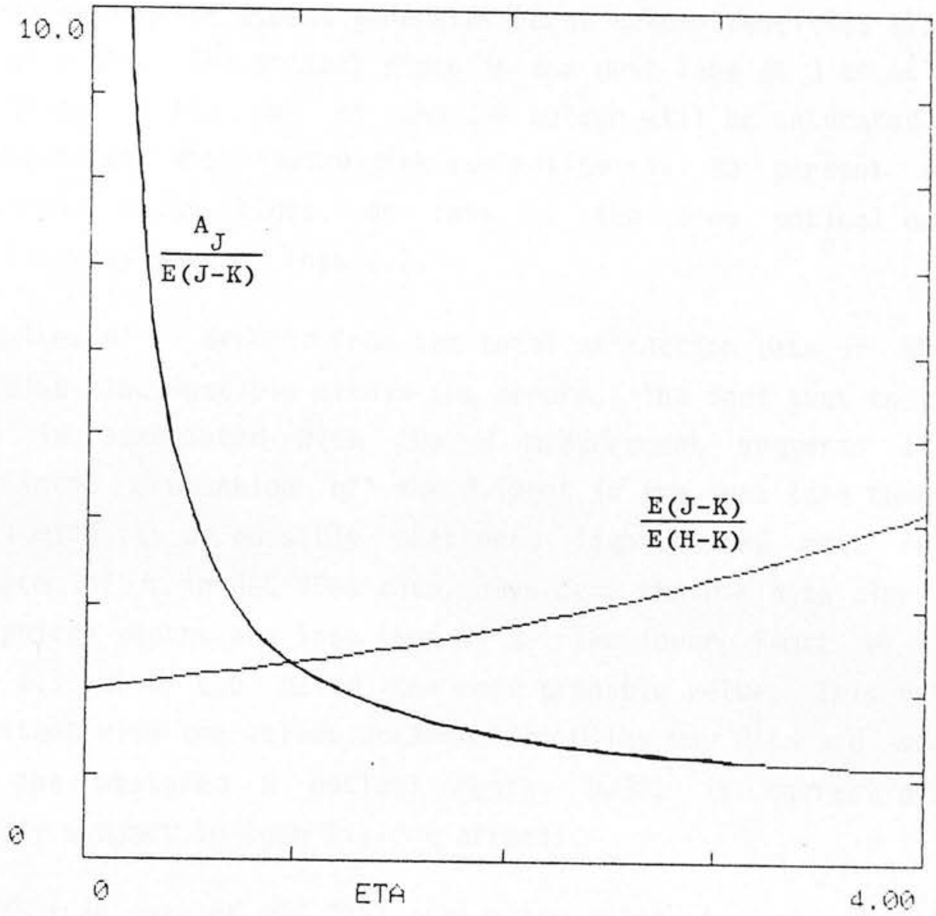


Fig. 6.12 The ratio of total to selctive extinction and the excess colour ratio plotted as functions of η , where the extinction is taken to vary as $\lambda^{-\eta}$.

It is immediately clear that the data are not consistent within the errors, even when the data refer to the same region. This is especially true for the NGC 4565 excess colour ratio data which is incompatible with an extinction power law decreasing with wavelength. The colours in NGC 4565 may well be saturated because the nearly edge-on aspect generates large column densities along the line of sight. The optical depth in the dust lane at J is at least 1.1 (Fig. 6.10) and so the J-K colour will be saturated if the foreground and distributed disk luminosity is 30 percent of the background bulge light, or less if the true optical depth is significantly greater than 1.1.

The values of η deduced from the total extinction data of NGC 4565 are also incompatible within the errors. The fact that the larger value is associated with the H measurement suggests that a significant proportion of the J light in the dust lane comes from disk luminosity or possibly scattered light. The most reliable estimate of η in NGC 4565 data comes from the H-K data alone since the optical depths are less than at J. The lower limit of η is then 2.1 with 2.5 being the most probable value. This value is consistent with the values obained from Milky Way data and suggests that the measured H optical depth, 0.32, is correct although possibly subject to beam filling effects.

The N/S scan data of NGC 7331 give a low value of   but this is not surprising since the ΔJ, $\Delta (J-K)$ relationship for distributed luminosity and dust is only linear if the optical depths are small (Fig. 6.3b). The two colour data from the minor axis, in which the dust is in front of the luminosity, give a larger value of η, although the errors are such that the two values are not inconsistent.

The three highest values of η (which are strictly upper limits) obtained from the data all agree well with each other and with the value deduced by Jones and Hyland, but are not incompatible with the results of Elias within the errors. These three values are likely to the most reliable since they are derived from more favourable geometries or wavelengths than the other values.

To summarise, the near infrared extinction curve has been shown to be similar in the spiral galaxies observed and the Milky Way. Combined with the findings of Morgan and Nandy (1982) the curve is similar, within quite large errors, in a wide range of Hubble types including Sab to Sc spirals and Magellanic types. The errors in the data are such that the observed properties of extragalactic dust are compatible with the conflicting findings of Elias (1978c) and Jones and Hyland (1980). A simultaneous three wavelength system would allow significantly more accurate data to be obtained for plotting on two colour diagrams.

7. Distribution of Interstellar Material

7.1 Dust Distribution in NGC 7331

The distribution of dust within NGC 7331 may be deduced from the distribution of the colour variations since it has been shown that these are caused predominantly by extincton. The infrared brightness and colour are represented by brightness and colour on the Starlink images (Fig. 5.7), which show the two mapped areas (48 by 30 and 120 by 120 arcseconds) in J-K and H-K. (The same data are shown more quantitatively in the contour plots of Fig's 5.5 and 5.6.) In these figures the red areas correspond to red infrared colours and it is apparent that the inner regions of the galaxy are relatively dust free. The red colours occur beyond about 30 arcseconds along the major axis and 15 arcseconds along the minor axis. Farther out the colours show a tendency to approach those of the inner region although there is much scatter.

The J-K and H-K colours show the same overall distribution. The differences in detail are due largely to small pointing differences. The colours vary significantly over small distances as evidenced by the fully sampled major axis scan (Fig. 5.4). The infrared data show a strong resemblence to the visual photographs. Because the infrared optical depths are less the obscured regions appear red rather than dark and the degree of reddening can be used as a quantitative indicator of the dust column density, though subject to the geometry effects described in Section 6.2.

The one dimensional profiles allow the radial dust distribution to be investigated more quantitatively than the two dimensional maps, since they do not depend on background levels and some are aligned with the axes of the galaxy. The scans, both at 5 and 10 arcsecond resolution, show the same general trends and these are seen in both the H-K and the more extensive J-K data. In the inner regions the colours show little variation, but at about 30 arcseconds along the major axis (Fig's 5.3 and 5.4) the J-K colour rapidly reddens by about 0.15 magnitudes in the 10 arcsecond aperture. The N/S scan made in a 5 arcsecond aperture shows excess colours of larger maximum amplitude (about 0.4 magnitudes) and also more variation; at about 40 arcseconds south of the nucleus the colours become as blue as the nucleus before becoming red again. This is due to the uneven distribution of the dust being better resolved in the smaller aperture.

The N/S and E/W scans both suggest that the dust is located in a ring located near the edge of the bulge. The major axis scan shows very clearly the flat colours inside this ring but does not show so clearly an outer limit to the red region. Beyond about 90 arcseconds the colours are about as blue as those of the nucleus but with considerable scatter. The decomposition of the B profiles by Boroson (1981) shows that beyond 60 arcseconds along the major axis the luminosity is mostly from the disk. As the light becomes more disk dominated, when the luminsosity is divided less equally in two, it is likely that the colours become more variable due to population and geometry effects.

The outer extent of the red region is better defined in the minor axis and E/W (15 degrees from the minor axis) profiles, in which the colours become blue again at about 25 and 30 arcseconds respectively from the nucleus. Assuming an inclination of 72 degrees (Boroson 1981) this corresponds to a distance of about 80 arcseconds along the major axis. Here the disk is about 0.5 magnitudes brighter than the bulge and so the colours will depend on the location of the dust within the disk rather than the projection effect of the bulge seen through the disk. The colours beyond about 90 arcseconds from the nucleus along the major axis, or the equivalent distance in other directions, are similar to those of the disk and so it may be concluded that the dust column densities in the outer regions are less than those in the ring, but by an amount difficult to quantify.

The fact that there is an inner edge to the reddened region does not necessarily imply that the dust distribution has such an inner edge. For example progressing along the minor axis, the dust becomes more deeply embedded in the luminosity and so the reddening will become less. At the nucleus, and along the major axis, the dust divides the luminosity equally in two and so if the dust distribution continued inside the bulge with the same column density the reddening at the nucleus would be about half that where the dust was in front of the luminosity. There would therefore be a gradual transition along the minor axis from red to blue with the nucleus itself being of intermediate colour. The minor axis data (Fig. 5.3) show that this is not the case.

Inspecting the major axis data the same conslusion is reached. Along the major axis the dust divides the luminosity equally in two (p=0.5, Section 6.2.3) along the whole scan and so the colour should remain unchanged inside the bulge dominated region if the dust column density were constant. Boroson (1981) shows that the bulge dominates the B light within about 60 arcseconds of the nucleus. The infrared data show that the J-K and H-K colours become red 30 arcseconds from the nucleus and so this transition is not due to dilution of the reddened light by unreddened bulge light. The bulge and disk dominated regions are not so clearly defined in the infrared data (Section 5.2.3), but there is no evidence that the luminsoity is disk dominated to within 30 arcseconds of the nucleus.

Both major axis scans, at 5 and 10 arcsecond point spacing, show that the region within 30 arcseconds of the nucleus is constant in J-K to within 0.02 magnitudes, about the observational error. The minor axis data show that the colour is constant from the far side of the bulge (25 arcseconds northeast) to the inner edge of the of the reddened region (10 arcseconds southwest). From these data it may be concluded that, within 30 arcseconds (2.9 kpc at a distance of 20 Mpc) of the nucleus, the dust column densities are much less (no more than 20 percent) than the densities beyond this region.

The earlier data using a 5 arcsecond aperture along the N/S and E/W directions show similar features to the major and minor axes but the colours of the inner region are about 0.1 magnitudes redder on the north side. This could suggest a larger amount of dust in the inner regions than implied by the major and minor axis data but in this case the colours should redden gradually. The brightness gradient here is steep (about 0.2 magnitudes per arcsecond) and so a small difference in the effective positions of the apertures of the two beams would cause a constant difference between the two sides of the nucleus. This error would represent a shift dependent on the size and direction of the brightess gradient but would not cause random fluctuations or affect the conclusions drawn from these data (Telesco, Gatley and Stewart : Appendix B).

The galaxy may be divided into three axially symmetric regions characterised by their distance from the nucleus along the major axis. The inner 24 arcseconds (2.4 kpc) show colour variations of no more than 0.02 magnitudes in J-K. Between 29 and 90 arcseconds (2.9 kpc and 8.7 kpc) the colours are reddened by 0.14 magnitudes in J-K (standard deviation is 0.04) using the data from the fully sampled major axis scan. Beyond this region the dust surface density seems to decline. The observed excess J-K colours of these regions relative to the nucleus are shown in Table 7.1. The reddened region, 30 to 90 arcseconds, is divided into two equal halves which have equal excess colours within the errors. Although the individual data show marginal evidence for intrinsic scatter, when averaged together within these regions the standard deviations are consistent with the statistical errors of the data.

The inner edge of the dust distribution is well defined, whereas the outer limit may be a gradual decline as observed in the Milky Way molecular gas distribution. The dust distribution may therefore be considered to exhibit a central hole rather than a well defined ring.

Arcseconds	kpc	N	E(J-K)	rms error	st. dev.
0 - 24	0 - 2.4	6	0.00	0.01	0.01
29 - 57	2.8 - 5.5	7	0.13	0.03	0.03
62 - 90	6.0 - 8.7	7	0.16	0.04	0.05
95 - 128	9.2 - 12.4	5	-0.03	0.06	0.08

Table 7.1 The excess J-K colours of regions of the major axis of NGC 7331 relative to the nucleus. The distance is indicated in arcseconds southeast of the nucleus and in kpc assuming a distance of 20 Mpc. N is the number of points in each region and the standard deviation of the excess colours is indicated.

7.2 Optical Depth of Dust in NGC 7331

The dust surface density as a function of radius may be most conveniently determined from the major axis data, where the dust bisects the bulge luminosity, and the near side of the minor axis, where the bulge luminosity is predominantly behind the dust. The three dimensional distribution of the luminosity does not then need to be considered in detail. If the dust is not optically thick the apparent reddening will be half that which would be obtained if the dust were in front of the luminosity (Section 6.2.3). If the dust is optically thick or clumpy then the amount of dust will be an underestimate.

The major axis was measured in a 10 arcsecond aperture on both sides of the nucleus at 10 arcsecond spacing and on one side only at 5 arcsecond spacing. The distribution of the reddening is similar in general shape on both sides of the nucleus. The data spaced by 5 arcseconds will be used in preference to the 10 arcsecond data to derive the densities, because of its fully sampled nature.

The disk dominated region (beyond 60 arcseconds along the major axis) is intermediate between the situation, in which the luminosity is bisected by the dust, and that in which the luminosity and dust are homogeneous , since the scale height of the dust is less than that of the disk stars (e.g. van der Kruit and Searle 1982a). If the dust is not optically thick both approximations indicate an apparent reddening one half that which would be measured if the dust were in front of the luminosity (Section 6.2.3). If the dust is optically thick or clumpy then the amount of dust will be underestimated. This is the same situation as for the bulge dominated region and so the measured colours will be doubled to derive a lower limit to the dust surface density.

If dust clouds are present above the galactic plane, doubling the apparent reddening will cause the dust column densities to be overestimated. Such clouds have been observed in the spheroid of M31 (Gallagher and Hunter 1981) where their central optical depth at I is < 0.3, and in NGC 185 where the optical depth is > 2.3. These clouds are relatively small (about 30 pc diameter) but the technique was not sensitive to large structures. High latitude clouds with sizes larger than 100 pc have been observed in galaxies by Elmegreen and Elmegreen (1980). Overall as many such clouds will lie below the galactic plane as above and so the distribution of clouds perpendicular to the plane will not affect the derived surface densities.

Such clouds can be distinguished by their location on the two colour diagram. If they are relatively thick they can display colours in excess of the maximum possible for centrally distributed dust (p=0.5 in Fig. 6.3) and also be close to the reddening vector rather than below it. The red features at 30 and 50 arcseconds south of the nucleus (Fig. 5.2) are the reddest points in the two colour diagram (Fig. 6.7c). They show colours slightly in excess of the maximum possible with centrally distributed dust and also lie close to the reddening vector while the other points lie mostly below it. It is likely therefore that these points represent clouds lying above the central plane of the luminosity. More detailed data would allow their optical depths and locations within the luminosity to be determined using the radiative transfer techniques of Elmegreen (1980).

For both disk and bulge dominated regions therefore, the dust surface density may be determined by doubling the measured reddening, so long as the dust is distributed on a large scale evenly about the galactic plane. Clumping effects will cause this to be a lower limit as demonstrated in Section 6.2.4. If the line of sight passes through an individual cloud the dust density could be underestimated by a factor of about 3. If the galaxy is highly inclined, the line of sight will pass through many such clumps and the clumping effect will be lessened. The inclinataion of NGC 7331 is large (72 degrees) and so no factor will be applied to account for clumping.

The equivalent excess colour, $E'(J-K)$, produced by the same amount of dust in front of the stars and observed face-on is therefore

$$E'(J-K) = 2\,E(J-K)\cos i$$

where $E(J-K)$ is the measured colour excess and i is the inclination of the galaxy. For NGC 7331, $E(J-K)$ is 0.14 magnitudes between 30 and 90 arcseconds and i is 72 degrees giving

$$E'(J-K) = 0.09$$

Section 6.8 shows that there is no reason to believe that the infrared extinction law in other galaxies is different to that of the Milky Way and so the equivalent A_V can be deduced in principle. Jones and Hyland (1980) show that such a transformation is not well determined but the van de Hulst curve 15 can be considered sufficiently accurate. (It also gives values close to those obtained by Koorneef (1983b) in a compromise of the existing data). This gives (Johnson 1968)

$$A_V = 6.4\,E(J-K)$$

and so the equivalent face-on visual extinction is

$$A'_V = 0.55$$

A comparable value may be deduced from the N/S scan. Here E(J-K) in the reddest region is about 0.45 magnitudes and so A_V is 0.9 magnitudes since, as discussed above, these dust clouds are probably in front of most of the luminosity and the factor of 2 is not then required. The N/S data may not be as easily averaged as the major and minor axis data since the relative position of the dust within the luminosity varies along the scan.

The minor axis scan shows an excess J-K colour of 0.3 magnitudes which is consistent with the major axis value of 0.14, since in this case the dust is in front of the luminosity. The value of A'_V deduced from the minor axis data is 0.59, in very good agreement with the major axis value of 0.55.

7.3 Dust and Gas Densities in NGC 7331

The corrected face-on optical depths derived in the previous section are related directly to grain number density and more indirectly to dust mass and gas densities. The extinction curve has been shown to be similar to that of the Milky Way and so it is reasonable to suppose that the grain sizes and shapes are also similar. Relationships between extinction and dust and gas densities derived from Milky Way observations are therefore likely to be valid for the galaxies oberved. There is reason to believe that different gas to dust ratios may occur in the Magellanic Clouds (e.g. Koorneef 1982) but it is unreasonable to suppose that dust in other spiral galaxies should be similar to dust in the Magellanic Clouds rather than in the Milky Way. The relationships between extinction and densities are not well determined but allow comparisons to be made between the Milky Way and the observed galaxies.

Although dust mass is more directly related to the extinction, the relationship between gas densities and extinction has been better determined, since gas densities, unlike dust density, can be measured directly. Extinction may be measured by excess colours, usually E(B-V), or for dense clouds by star counts of the background field. Gas densities can be measured by 21 cm emission or Lyman absorption for HI, by ultraviolet transitions for low density H_2 and by millimetre emission of other molecules, usually CO.

HI column densities along the line of sight to reddened stars have been measured by 21 cm emission and Lα absorption (e.g. Kerr 1969). The advantage of the Lα technique is that gas behind the star does not contribute but the ultraviolet observations are limited to relatively low densities ($A_V < 1.5$). Jenkins and Savage (1974) have compared stellar colour excesses with L absorption strengths to derive the relationship

$$N(\text{H total})/E(B-V) = 7.5 \times 10^{21} \text{ atom cm}^{-2}$$

The Lα absorption allows only the column density of HI to be measured and they make corrections for ionisation near the stars and molecular hydrogen along the line of sight, which was measured for some of the stars by Spitzer et al (1973). The correlation between column gas densities and excess colours was better when the H_2 densities were included, implying that the dust density is proportional to the total gas density and not the density of either atomic or molecular hydrogen.

The accuracy of this relationship has been improved by further measurements (Bohlin et al 1978) giving the relationships

$$N(\text{H total}) \ / \ E(B-V) = 5.8 \times 10^{21} \text{ atom cm}^{-2} \text{ mag}^{-1} \qquad (7.1)$$

The dust densities themselves have been estimated using simple dust grain models (Jenkins and Savage 1974) and the ratio of gas to dust mass density has been estimated as about 170.

The relationships above can be applied directly only to low levels of infrared extinction, i.e. $E(J-K) < 0.2$, but the conclusion that the dust is related to the total gas density indicates that extrapolation to larger densities may be justified. In dense clouds the gas is predominantly molecular but gas densities cannot be measured by the ultraviolet absorption because of the very large optical depths. The H_2 molecule does not have any dipole moment and so rotational transitions, which could give rise to observable emission lines, are not allowed. The infrared quadrupole transition is observed (Shull and Beckwith, 1982) but only in special circumstances (shocked gas) and so is of no use in deriving column densities in dark clouds. After H_2 the next most abundant molecule in the interstellar gas is CO and this does have observable dipole transitions in the millimetre region.

Derivations of molecular gas densities from CO line strengths are subject to many uncertainties including relative abundances, excitation temperatures of the gas, saturation effects and possible deviations from local thermodynamic equilibrium (LTE). The uncertainties are such that one method of estimating molecular gas densities is to assume that the gas density to extinction ratio is as measured in less dense clouds, and then to relate the implied gas density to the observed CO line strength (Dickman 1978). Many estimates of molecular gas density are based on these results, although other methods have been applied such as LTE analysis (Frerking et al 1982) and mass estimates from the virial theorem (Young and Scoville 1982a). Many uncertainties remain and it has been suggested (e.g. Leung and Liszt 1976 and Blitz and Shu 1980) that molecular gas densities have been considerably overestimated.

The constancy of the gas to dust ratio has therefore only been demonstrated directly for relatively low density clouds ($A_V < 1.5$). There is no evidence that this ratio is not also valid for denser clouds which would produce measurable infrared reddening; these clouds would be unlikely to be more than 2 times denser than those for which the relationship has been verified observationally. However the ratio has been shown to be different in some objects e.g. in ρ Oph the ratio is about 3 times higher (Savage and Mathis 1979), believed to be the result of different dust properties at visual wavelengths. The Magellanic Clouds appear to have a higer gas to dust ratio than the Milky Way. However other spiral galaxies may be expected to have gas to dust ratios similar to those found in the Milky Way and also similar dust properties (Chapter 6).

To derive gas densities from the infrared reddening, equation 7.1 may be rewritten

$$N(\mathrm{H\ total}) / E(J\text{-}K) = 1.2 \times 10^{22}\ \mathrm{cm}^2\ \mathrm{mag}^{-1}$$

using the value E(J-K)/E(B-V) = 0.48 from the van de Hulst curve 15 (Johnson 1968). Although this relationship is poorly determined observationally, the errors introduced in estimates of the gas density will be small compared to the other uncertainties.

The molecular gas density in the 30 to 90 arcsecond region of NGC 7331 can be derived from the equivalent face-on excess J-K colour of 0.09 magnitudes. This gives the value

$$\begin{aligned} N(\mathrm{H\ total}) &= 1.1 \times 10^{21}\ \mathrm{atoms\ cm}^{-2} \\ &= 9\ M_{\odot}\mathrm{pc}^{-2} \end{aligned}$$

This value may be compared with the gas surface density determined by the CO line strength (Young and Scoville 1982c). This has a value of 27 $M_{\odot}\mathrm{pc}^{-2}$ in a 50 arcsecond beam centred 45 arcseconds away from the nucleus. These values can be considered to be in good agreement since the value derived from the extinction is a lower limit and, as discussed above, large uncertainties exist in the relationships used to derive the gas densities from CO line strengths.

Assuming a distance of 20 Mpc for NGC 7331, the mass of interstellar material in the reddened 2.9 to 8.7 kpc region may be calculated. The lower limit of 9 $M_{\odot}\mathrm{pc}^{-2}$ gives a mass of $2 \times 10^9 M_{\odot}$. The total mass within 8.7 kpc of the nucleus may be calculated from the rotational velocity at 8.7 kpc. The mass M within radius R at which the rotational velocity is V is given by

$$M = V^2\ R/G$$

where G is the gravational constant. This assumes Keplerian motion, i.e. the mass is assumed to be centrally concentrated but it is still a good estimate of the mass when the mass is distributed in a disk (e.g. Fall 1981). This gives a total mass inside 8.7 kpc of 1.1 x $10^{11} M_\odot$, for a rotational velocity of 230 km s^{-1} (Rubin et al 1965 and Bosma 1981). The proportion of material in the ISM within 8.7 kpc is therefore at least 2.6 percent and probably about double this as deduced by Young and Scoville (1982c) from the CO data.

The mass within the central 2.9 kpc may be similarly estimated from the rotational velocity of 180 km s^{-1}, giving a value of 2 x $10^{10} M_\odot$. Because at this radius the mass is distributed in a more complex manner than at larger radii, this mass estimate is more subject to error but it does indicate that the inclusion of this region in the total mass does not greatly affect the proportion of mass in the ISM. A more complete treatment may be made by fitting a mass model to the rotation curve as done for NGC 7331 by Bosma (1981). Integrating the surface density distribution yields a total mass compatible with the figure derived above. The mass model itself may be subject to large errors due to the high inclination of NGC 7331 and so cannot significantly improve the lower limit provided by the simple model.

7.4 The Interstellar Medium in NGC 7331 and the Milky Way

The large scale structure of the interstellar medium (ISM) in the Milky Way was first determined by observations in the 21 cm line of HI. The disk has been mapped in several surveys (Burton 1976) by measuring the intensity and spectrum of the emission as a function of direction. The spectrum allows the velocity of various components of the gas to be determined and from this can be deduced the distance from the Galactic Centre of the emitting gas.

The column densities as measured can be converted to surface densities by knowledge of the scale height of the gas. Baker and Burton (1975) estimated the width of the HI distribution to be 280 pc, between half density points, when the gas was represented by a one component model. The peak surface density of the HI gas was estimated to be about 3 $M_\odot pc^{-2}$ using this technique.

Until CO surveys were performed it was thought that the HI distribution was representative of the distribution of the ISM as a whole. CO surveys show that molecular gas is more common than atomic gas in the galactic plane and follows a very different distribution. These surveys (e.g. Scoville and Solomon 1975, Burton et al 1975, Gordon and Burton 1976) were made undersampling the sky due to the beamsizes of about 50 arcseconds obtained with the 36 ft NRAO telescope. They showed that the CO and hence the H_2 was largely confined to a region between 4 and 8 kpc from the Galactic Centre, the so-called molecular ring, and was closely confined to the galactic plane.

The peak volume density was estimated to be about 2 molecules cm^{-3} at 5.5 kpc, about 10 times the atomic hydrogen density. As discussed above the determinations of molecular gas densities are subject to many uncertainties and the dominance of H_2 with respect to HI is still in some dispute (Blitz and Shu 1980). It is clear however that the total gas density and the molecular gas in particular are depleted within 4 kpc. The surface density, as would be observed from outside the galaxy, can be determined using estimates of the scale height of the gas. Burton and Gordon (1976) estimated the thickness of the molecular gas distribution to be 117 pc, i.e. about half that of the HI, from which Gordon and Burton (1976) deduced that the peak molecular gas density was about 9 $M_\odot$ pc^{-2}.

Cohen and Thaddeus (1977) used a small 1.2 m telescope to obtain fully sampled data in an 8 arcminute beam. They measured the z distribution finding that the half thickness at half maximum was 59 pc producing surface densities in good agreement with Gordon and Burton. However they found that the thickness increases outside 7 kpc so making the 5.5 kpc peak less well defined. They concluded that the surface density is approximately constant between 5 and 8 kpc. This is very similar to the shape of the dust distribution in NGC 7331, which maintains an approximately constant surface density between 2.9 and 8.7 kpc.

The volume densities derived by Cohen and Thaddeus are between 1.4 and 4 times less than those measured in the previous surveys. They postulate that this is due to a weak substratum of CO emission best detected in the Gordon and Burton survey; this is a further source of error in the estimates of molecular gas density. This error arises because the densities are derived by counting the number of clouds in the optically thick ^{12}CO line and assuming a representative mass for the clouds as determined by the optically thin ^{13}CO line. Thinly distributed material is not therefore accurately detected by the cloud counting method. The opposite is true for densities derived from the reddening, since optically thin material contribues more to the reddening of distributed background luminosity than optically opaque material.

The survey of Cohen and Thaddeus also differs inside 4 kpc where they find that the CO emission is close to zero whereas Gordon and Burton find that the emission within the inner regions is almost half that at 5.5 kpc. Cohen and Thaddeus suggest that this discrepancy may be due to undersampling or to errors in the assumptions of circular rotation and symmetry. The existence of non-circular motions makes the determination of radial distribution increasingly difficult as the radius decreases. The dust distributon in NGC 7331 shows a rapid decline inside 2.9 kpc and is very similar in shape to the molecular gas distribution of the Milky Way as determined by Cohen and Thaddeus. With regard both to this respect and the slow rate of decline at larger radii, the dust distribution in NGC 7331 and the molecular gas distribution of the Milky Way as determined by Cohen and Thaddeus (1977) are very similar.

7.5 Atomic or Molecular Gas

In the discussion above the distribution of the dust in NGC 7331 and of the molecular gas in the Milky Way have been compared and found to be very similar. The density of the gas (9 $M_{\odot}pc^{-2}$) in NGC 7331 inferred from the dust density is also very similar to the molecular gas density in the Milky Way (30 $M_{\odot}pc^{-2}$ Solomon and Sanders 1977, 12 $M_{\odot}pc^{-2}$ Gordon and Burton 1976), considering that the former is a lower limit. These facts do not however necessarily imply that the gas associated with the dust in NGC 7331 is molecular in form.

If a thickness of 100 pc is assumed for the gas layer in NGC 7331, the volume density is about 2.3 H nuclei cm^{-3}. This is comparable to the average density of gas in the Milky Way between 4 and 8 kpc, which is 8 cm^{-3} according to Solomon and Sanders (1977), and greater than the likely lower limit of 2.5 cm^{-3} deduced by Liszt et al (1981). The density of the molecular clouds themselves, as defined by the region within which there is detectable CO emission, is much larger, typically 300 cm^{-3} (Solomon and Sanders 1977). Over a diameter of 50 pc the optical depth at V would be over 20 and so the dense parts of molecular clouds would be opaque in the near infrared and would not contribute to the reddening. The cores are even more dense but in the outer regions the column densities may be as low as 2 cm^{-3} or less (Burton 1979) and so could be responsible for reddening in the near infrared.

The filling factor of giant molecular clouds over the area of the Milky Way seen face-on may be estimated using the properties described by Solomon and Sanders (1977). They find that the average dimension (square root of the area) is 40 pc and estimate that there are 4000 such clouds between 4 and 8 kpc. Assuming they are spherical they would occupy 4 percent of the area in this region seen face-on. At an inclination of 72 degrees this fraction would increase to 14 percent if the clouds are assumed spherical. (If they are flattened along the plane then the observed filling factor will also be larger.) The characteristic size of 40 pc is based on a convenient antenna temperature of 3K in the ^{13}CO line. This represents emission in a region with a temperature of 10K and a density of 300 cm^{-3} and colder or less dense gas would not contribute to the measured cloud sizes. Their maps show larger areas with less intense emission between the dense clouds.

Molecular cloud complexes can have significantly larger dimensions than the characteristic dimension of 40 pc. The Rosette molecular complex (Blitz and Thaddeus 1980) has a linear size of about 100 pc and is extended in the plane of the Galaxy. It appears to consist of clumps of molecular gas embedded within an HI cloud occupying an area on the sky 5 to 10 times greater than the area of the molecular complex (Raimond 1966). The Taurus complex has also been shown to have large amounts of lower density material associated with the molecular emission (Heiles and Habing 1974).

This picture is similar to the distribution of dust clouds described by Hollenbach (1971), who made a statistical analysis of the dark clouds catalogued by Lynds (1962). The denser clouds were found to be surrounded by less dense material. It is therefore very likely that the dust associated with molecular gas as observed in the Milky Way would fill large proportions of the 500 and 1000 pc apertures used to observe NGC 7331.

The inferred gas density in NGC 7331 may be compared with typical cloud and intercloud HI densities in the Milky Way. Radhakrishnan et al (1972) found that average HI clouds have a column density of 3 x 10 cm^{-3} and that the column density in the solar neighbourhood is about 1.4 x 10 cm^{-3}, equivalent to an A_V of 0.07. The distribution of material between cold clouds and warm intercloud regions is disputed by Greisen (1973) but similar column densities are derived. The peak volume density of HI in the galactic plane is 0.4 atoms cm^{-3} (Burton 1976) which corresponds to a column density of 1.2 x 10 cm^{-3} and an A_V of 0.06 for a layer thickness of 100 pc. These values are less than those observed in NGC 7331 by a factor of about 10.

The peak HI column density in the cloud associated with the Rosette molecular complex is 2.9 x 10 cm^{-3}, equivalent to an A_V of 1.4. This is 2.5 times greater than the average value observed in the 30 to 90 arcsecond region of NGC 7331 and so represents the approximate density required to produce the observed infrared reddening, when clumping and beam filling effects are taken into account. HI clouds associated with molecular complexes are therefore good candidates for the origin of the observed infrared reddening, whereas the diffuse HI interstellar medium is not able to produce the observed colours.

If the gas in NGC 7331 were atomic rather than molecular, then some mechanism would have to exist in NGC 7331 to prevent the formation of molecules in similar density regimes to those of the Milky Way, since the deduced total gas surface density in NGC 7331 is comparable to the molecular gas surface density of the Milky Way. The phase of the gas depends on the density, the temperature and the properties of the associated dust, which can shield the molecules from dissociating radiation and provides sites for the formation of molecules. The infrared extinction properties of the dust in NGC 7331 have been shown to be similar to those of dust in the Milky Way and so it is reasonable to assume that its affect on the gas is also similar. There is no reason to assume that the radiation field is different in the dusty regions of NGC 7331 and the Milky Way and so again the conclusion is reached that the gas is predominantly molecular.

This interpretation is confirmed by the observations of CO emission (Young and Scoville 1982c) and HI emission (Bosma 1981) in NGC 7331. The total column density of the gas inferred from the infrared extinction is 9 $M_{\odot}pc^{-2}$, the molecular gas density from the CO data is 27 $M_{\odot}pc^{-2}$ and the atomic gas density from the 21 cm data is 6 $M_{\odot}pc^{-2}$ at a distance of about 50 arcseconds from the nucleus. When clumping effects are considered (Section 6.2.4) the gas density inferred from the infrared reddening is consistent with the molecular gas density but not with the atomic gas density.

The column density of the molecular gas is derived from a more uncertain set of assumptions than the total gas column density or the atomic gas column density. The determination of total gas column density from reddening is based on relationships derived from 21 cm data (e.g. Bohlin et al 1978) and some additional assumptions about the exact relationship between infrared and visual reddening.

The inference may therefore be made that the atomic gas can comprise no more than 50 percent of the total gas in the 3 to 9 kpc region of NGC 7331. The total gas column density in the inner 3 kpc of NGC 7331 has been shown to be no more than 20 percent of that in the outer region. Taking a proportion of the figure derived by Young and Scoville for the molecular gas density at 45 arcseconds, the total gas density in the inner region can be no more than 6 $M_{\odot}pc^{-2}$, assuming that the dust clumping is similar in these regions.

This upper limit is compatible with the figure of 4 $M_{\odot}pc^{-2}$ found by Bosma for HI and indicates that the proportion of molecular gas in the inner region is small. The size of the synthesised aperture used by Bosma was 25 by 45 arcseconds (RA and declination) and so the nuclear point lies within the region, in which there is no strong reddening. The beam size used for the CO observations, 50 arcseconds, is such that the nuclear point will include a contribution from the region in which there is molecular emission. This suggests that the figure of 27 $M_{\odot}pc^{-2}$ derived by Young and Scoville cannot significantly overestimate the gas density in the region around 45 arcseconds. Their estimate of the gas surface density is based on the relationship between the CO line intensity with H_2 column density, which as mentioned in Section 7.3 is subject to many uncertainties.

Similar assumptions are used in the estimates of the molecular gas content of the Milky Way and imply that the ISM is dominated by molecular gas outside 5 kpc. This dominance has been questioned by Blitz and Shu (1980) who use theoretical and observational data to argue that the mass of molecular gas has been considerably overestimated. They argue that $N(H_2)/N(^{13}CO)$ is 4 x 10^5 , about 5 times less than the values used by e.g. Solomon and Sanders (1977) who conclude that H_2 dominates over HI in the 5 to 8 kpc region. This lower value implies that in the 5.5 kpc molecular ring of the Milky Way, the ratio of H_2 to HI is about unity and that the central hole in the total gas distribution is much less pronounced. The predominance of molecular gas outside 3 kpc in NGC 7331 implies that the gas in the Milky Way is also predominantly molecular outside 4 kpc because of the similarities in radial distributions of atomic and molecular gas and in the rotations curves.

7.6 Other Galaxies

The inclination of NGC 7331 makes it particularly suitable for studying the radial distribution of the dust by measuring the infrared reddening. The other galaxies studied are all more edge-on making saturation effects more severe. In addition the beam when projected onto the plane of the galaxies covers a large range of radial distances. Estimates of the surface densities can be made although the correction to equivalent face-on values is subject to some uncertainty due to uncertainties in the values of cos i.

NGC 4216 was scanned in the N/S and E/W directions, 20 degrees away from the major and minor axes respectively (Hamabe and Okamura 1981). The J-K colour along the N/S scan (Fig. 5.2), apart from a gradient consistent with the changing relative location of the dust within the luminosity as the near side is approached from the far side, shows little variation larger than the statistical errors. Similarities between the two sides of the scan suggest that some of these colour variations may be real and related to extinction or population effects but higher sensitivity data are needed to analyse them properly.

The E/W scan, 20 degrees away from the minor axis, shows strong evidence for reddening on the east (near) side. Comparison with the optical photograph (Fig. 5.1) shows that this corresponds to an obvious dust lane as is also indicated by the asymmetry in the J profile. The three points between 14 and 21 arcseconds are 0.16 ±0.03 magnitudes redder than the nucleus. The inclination of NGC 4216 is estimated as 83 degrees (de Vaucouleurs 1958) and so the equivalent face-on visual extinction A'_V is 0.12 using the technique applied to NGC 7331. The distance to NGC 4216 has been estimated as 21.4 Mpc (Simkin 1975) and so the radial extent of the reddening is 1.5 to 2.2 kpc. The high inclination, however, makes both the radial distance and A'_V very uncertain.

In both NGC 4565 and NGC 5907 the colours show evidence for saturation. Lower limits to the visual optical depth of the dust in the densest parts of the dust lanes are about 4 magnitudes in both cases. The inclinations are so high, close to 90 degrees, that correction to equivalent face-on values is not possible.

7.7 Molecular Gas Distribution in Galaxies

Molecular gas was not observed in other galaxies until 1975 but since then CO and other molecules have been observed in over 40 galaxies. The observations are limited in resolution since the largest available telescopes operating in the millimetre region, the 11m National Radio Astronomy Observatory and the 16m Five College Radio Astronomy Observatory dishes, have a diffraction limited resolution of about one arcminute when observing the 2.6 mm $J = 0 \rightarrow 1$ transition of CO. Sensitivity has limited the amount of data obtainable and this has been aggravated by the low filling efficiency of concentrated sources in the large apertures.

The earlier searches for CO emission concentrated on the central regions of nearby galaxies. Many were detected but in an unbiased survey (Rowan Robinson et al 1980) of spiral galaxies no strong dependence between molecular gas content and morphological type or other characterisitics was found. Elliptical galaxies (Morris and Rickard 1982) and Magellanic type irregulars (Israel et al 1982a) are known to be much less abundant in molecular gas than spiral galaxies.

The first maps of the molecular gas distribution in other galaxies showed that in NGC 253, M51 and M82 (Rickard et al 1977a) the gas was centrally condensed and more abundant than HI. The data suggested a central source as in the Milky Way but were not extensive enough to distinguish between this model and a centrally condensed disk distribution. Such a disk model does fit the observed emission in Maffei 2 (Rickard et al 1977b) in which CO emission was detected out to a distance of 6 kpc. Subsequent observations of M51, together with two other late type spirals IC 342 abd NGC 6946, (Rickard and Palmer 1981) showed that the CO distribution could be represented by a central source superimposed on a fairly flat disk, which corresponds in extent to the optical emission rather than the HI distribution. In these galaxies the 65 arcsecond beam corresponds to dimensions of 3.0, 3.2 and 1.4 kpc respectively and so the resolution was adequate to distinguish a molecular ring feature in the presence of a central source. The resolution was also adequate to detect spiral structure but this was not observed although azimuthal asymmetries were present.

IC 342 and NGC 6946 were also observed at slightly higher resolution by Young and Scoville (1982a). Although their data are consistent with those of Rickard and Palmer, they interpret the emission as being representative of a disk which exhibits the same exponential fall off as the stellar disk observed in B. They agree with Rickard and Palmer in that there is no evidence for spiral features and show that there is no central hole in the molecular gas distribution (or molecular ring) as observed in the Milky Way. The data, although consistent with the exponential representing the B luminosity, can be fitted equally well by other functions but there is at least a rough correspondence between the B luminosity and the CO flux.

Since the B light in late type galaxies originates predominantly in young stars, the conclusion is drawn by Young and Scoville that the star formation rate per nucleon of interstellar gas is approximately constant. Observations of 8 other late type galaxies (Young and Scoville 1982b) reinforce this interpretion and show that the ratio is approximately constant both within and between galaxies over two orders of magnitude, even including the Seyfert galaxy NGC 1068. This relationship is not true for the HI densities. They conclude that the star formation process within molecular clouds is relatively unaffected by external influences such as shocks or cloud collisions.

No evidence for central holes or molecular rings was found in these galaxies, but was found in the CO observations of the Sb galaxies NGC 2841 and NGC 7331 (Young and Scoville 1982c); this is discussed in Section 7.9. Young and Scoville argue that, excluding the central hole, the CO surface density in NGC 7331 bears the same relationship to the B luminosity as was found for the Sc galaxies. NGC 2841, at an assumed distance of 7.6 Mpc rather than 20 Mpc, allows greater spatial resolution to be obtained from the 50 arcsecond beam. Its lower inclination (65 degrees rather than 72 degrees) also means that the beam, projected onto the plane of the galaxy, does not cover such a large range of radii. The CO flux was found to be approximately constant at the 4 points between 1.7 and 6.8 kpc. Upper limits on the central point and the single point at larger radius are about 40 percent of the flux in the molecular ring. In NGC 2841 therefore there is evidence from the CO data that the proportionality between molecular gas density and B flux does not apply. This apparent difference in the shapes of the molecular gas distributions in NGC 2841 and NGC 7331 is discussed in Section 7.8.

The relationship between molecular clouds and spiral structure in the Milky Way is still a controversial subject. Spiral structure is seen in the HI distribution of the Milky Way (e.g. Shane 1972) and in other galaxies, e.g. M81 (Rots 1975, Rots and Shane 1975). It might therefore be expected that CO, being associated with star formation regions, should exhibit strong spiral structure. The earlier CO surveys showed no evidence for a close association (Scoville et al 1979) but fully sampled data have been interpreted (Cohen et al 1980) as showing that molecular clouds are largely confined to spiral arms. The relationship has important implications for the lifetime and origin of molecular clouds.

The available resolution of CO data does not allow CO spiral structure, if present, to be resolved in any but very close galaxies. Combes et al (1977) and Boulanger et al (1981) have, however, found that CO emsission in M31 is correlated with spiral features, although the gas densities are considerably less than in the Milky Way and giant molecular cloud complexes to not appear to be present. Knowledge of the dust distribution with respect to spiral arms in other galaxies could help answer this question. However in NGC 7331 the spiral features are weak; they are most prominent in the parts of the galaxy corresponding to the outer regions of the large area map where the data are most noisy and so no conclusions can be drawn.

An important feature of the molecular gas distribution in the Milky Way is the strong central concentration. The HI distribution in the Inner Galaxy is flat but the CO data indicate that most of the molecular emission within 2 kpc is confined to the inner 300 pc. Liszt and Burton (1978) describe a "165 km s^{-1}" expanding arm observed in CO which they consider to be part of the rotating nuclear disk, also seen in the HI data (Burton and Liszt 1978). Bania (1980) states that many if not all of the high velocity features within 8 degrees of the Galactic Centre can be accounted for by the rotating disk model of Burton and Liszt. The amount of material within this nuclear disk has been estimated to be $3.5 \times 10^8 M_0$ (Bania, 1977) and at least $10^9 M_\odot$ (Liszt and Burton 1978). This density of material is surprisingly high but has been confirmed using a different method (absorption by HCO^- and HCN) by Linke et al (1981), who obtained an even higher value of about $8 \times 10^9 M_\odot$. The implied surface densities are several thousand $M_\odot pc^{-2}$; they would therefore be optically thick and undectable by infrared reddening, unless associated with a significant area of optically thin dust.

However if such a feature were present in NGC 7331, the central deficieny in the molecular gas distribution would not have been detectable from the CO data because of the low spatial resolution. This would not have been the case with the current infrared data (Fig. 7.1), demonstrating again the advantage of higher spatial resolution, as achieved in the current infrared observations.

7.8 A Model for the NGC 7331 Radial H_2 Distribution

The dust distribution obtained from the current infrared observations has been derived from major axis data at a resolution

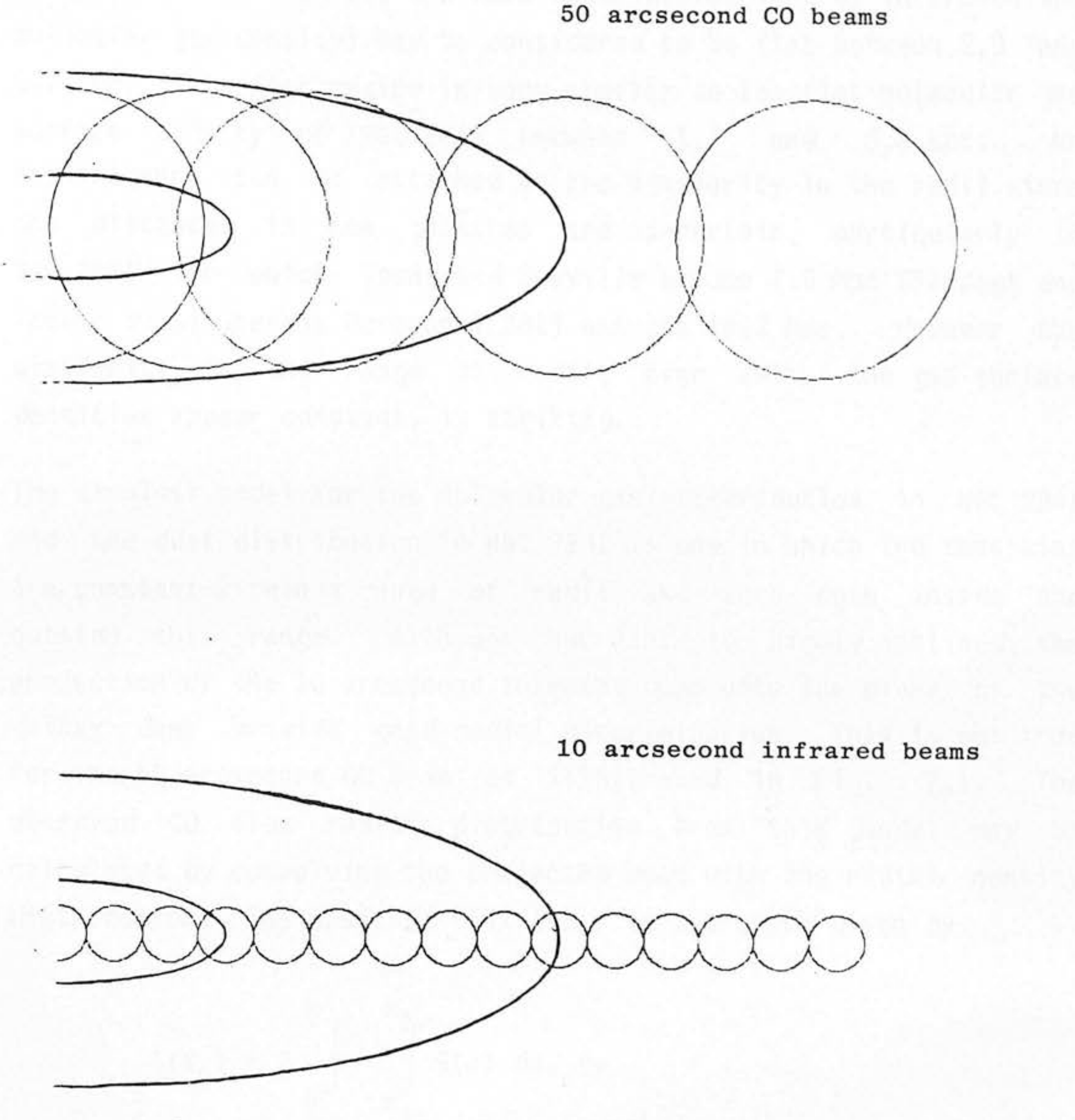


Fig. 7.1 The 50 arcsecond CO beams of Young and Scoville (1982b) and the 10 arcsecond infrared beams shown against the inclined plane of NGC 7331. The ellipses represent the inner and outer limits (30" and 90") of the molecular gas distribution of model **B** discussed in the text.

of 1.0 kpc compared to the CO data at a resolution of 6.8 kpc. The CO data for NGC 2841 is at a resolution of 1.7 kpc and so may be compared tentatively with the infrared data of NGC 7331. As discussed in Section 7.1 the dust distribution (and by inference the molecular gas density) may be considered to be flat between 2.9 and 8.7 kpc. This flat region is very similar to the flat molecular gas surface density of NGC 2841 between 1.7 and 6.8 kpc. No significance can be attached to the similarity in the radii since the distances to the galaxies are uncertain, particularly to NGC 2841 for which Young and Scoville assume 7.6 Mpc (Sandage and Tamann 1975) whereas Boroson (1981) assumes 15.7 Mpc. However the similarity in the range of radii, over which the gas surface densities appear constant, is striking.

The simplest model for the molecular gas distribution in NGC 2841 and the dust distribution in NGC 7331 is one in which the densities are constant within a range of radii and zero both inside and outside this range. Although NGC 7331 is highly inclined, the projection of the 10 arcsecond infrared beam onto the plane of the galaxy does provide good radial discrimination. This is not true for the 50 arcsecond CO beam as illustrated in Fig. 7.1. The observed CO flux radial distribution from this model may be calculated by convolving the projected beam with the radial density distribution. The observed flux $I(X_0)$ is therefore given by

$$I(X_0) = 2 \int_0^a \int_{x_1}^{x_2} S(r)\, dx.\, dy$$

where $r = (x + (y/\cos i)^2)^{1/2}$

$X_1 = X_0 - (a^2 + y^2\ )^{1/2}$

$X_2 = X_0 + (a^2 + y^2\ )^{1/2}$

S(r) is the surface luminosity at radius r, a is the radius of the beam, i is the inclination of the galaxy to the plane of the sky and X_0 is the position of the beam centre along the major axis. In the calculation the beam is assumed to be flat-topped and so represents the infrared beam more accurately than the diffraction limited CO beam.

Using this model the radial profiles have been calculated for both the 10 arcsecond (infrared) and 50 arcsecond (CO) beams in the cases where

(A) S(r) is proportional to the B luminosity (Boroson 1981) beyond 45 arcseconds and is zero within this radius; this is the interpretation given by Young and Scoville.

(B) S(r) is constant between 30 and 90 arcseconds and zero elsewhere.

(C) S(r) is proportional to the B luminosity of the disk outside 30 arcseconds; inside this region S(r) is zero.

The profiles calculated from these models are compared with the infrared and CO data in Fig. 7.2. The vertical scaling is chosen to match the models to the 45 arcsecond CO point. The infrared points represent the average excess colours of the data grouped as in Table 7.1. The error bars represent the estimated standard error of the mean.

Although model (A) fits the CO data well when projection effects are not taken into account, the fit is less accurate when the projected luminosity is integrated over the large beam. Models (A) and (B) fit the inner three CO points equally well. Both overestimate the contribution from the central point; this may be due to the diffraction limited CO beam being peaked rather than flat-topped as assumed in the model. The 90 arcsecond CO point is fitted better by model (B) than model (A). The point at 135 arcseconds is intermediate between the two models suggesting that the outer limit of the gas density is not as rapid as assumed in model (B).

The infrared data are much better described by model (B) than model (A) since there is no evidence for an inner peak to the measured density. The surface density of the dust and gas might be expected to undergo a gradual transition from zero to proportionality with the B luminosity and so the sharp inner peak would not be present in

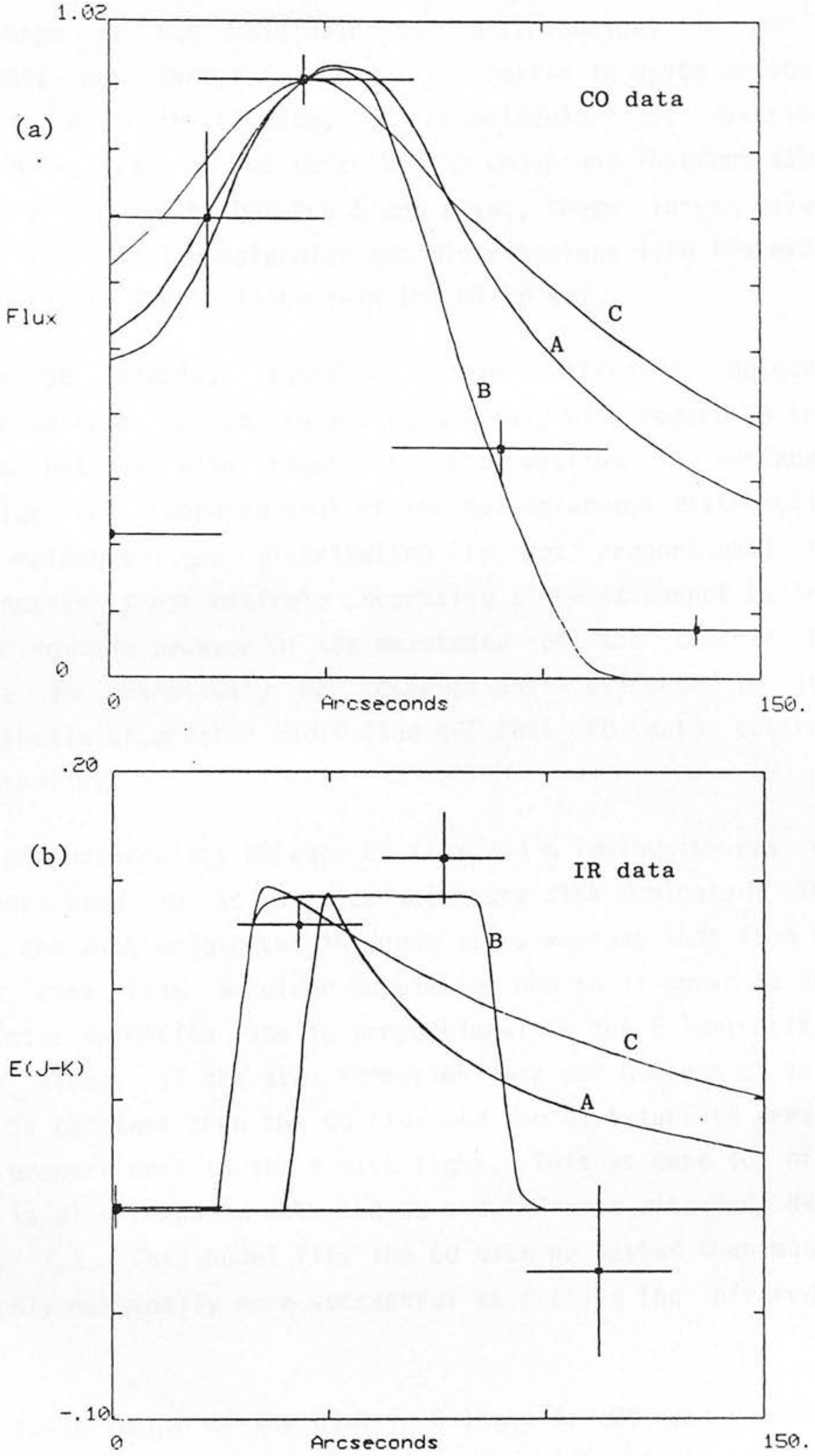


Fig. 7.2 The profiles calculated from the models of the molecular gas distribution described in the text. Convolved with a 50 arcsecond beam they are compared with the CO data (a), and convolved with a 10 arcsecond beam with the infrared data (b).

the data, but a decline with increasing radius should still be seen. The shape of the molecular gas distributions in NGC 7331 and NGC 2841 may therefore be very similar in spite of the apparent differences in the CO data. If the molecular gas distribution of the Milky Way is as described by Cohen and Thaddeus (1977), i.e. relatively constant between 5 and 8 kpc, these three galaxies may have very similar molecular gas distributions with the exception of the Galactic Centre feature in the Milky Way.

These Sb galaxies therefore have different molecular gas distributions to Sc galaxies, not only with regard to the central holes, but also with regard to the decline in surface density outside the supposed peak of the molecular gas distribution. That the molecular gas distribution is not proportional to the B luminosity is not entirely surprising since it cannot be true in the inner regions because of the existence of the central holes and there is essentially no observational evidence to support the hypothesis at greater radii (the NGC 2841 CO data contradict the hypothesis).

The proportionality between CO flux and B luminosity has only been demonstrated in Sc galaxies which are disk dominated. The B light from the disk originates in young stars whereas that from the bulge must come from an older population and so it could be argued that the star formation rate is proportional to the B luminosity from the disk alone. If the star formation rate per nucleon of interstellar gas is constant then the CO flux and the dust surface density should be proportional to the B disk light. This is case (C) of the model and is also compared with the CO and infrared data for NGC 7331 in Fig. 7.2. This model fits the CO data no better than model (A) and is only marginally more successful at fitting the infrared data than (A).

The scale length of the disk in B light is 104 arcseconds (Boroson 1981) which is about 1.5 times the width of the reddened region. This region is divided into two for comparison purposes in Table 7.1; these two halves have J-K colours of 0.13±0.01 and

0.16±0.02, i.e. equal within the errors, whereas they should differ by a factor of 1.3 if the dust density were proportional to the B light. The colours are inconsistent at the 95 percent confidence level with a difference corresponding to one scale length. Such a model also cannot fit the CO data for NGC 2841, since the scale length of the disk B light is 71 arcseconds (Boroson 1981) whereas the CO flux is constant over 135 arceconds.

Therefore unlike Sc galaxies, the star formation rate per nucleon of interstellar gas is not constant in these two Sb galaxies which possess central holes in their gas distribution. This ratio may however be constant beyond 180 arcseconds in NGC 2841 and 60 arcseconds in NGC 7331 but higher quality data would be necessary to test this. Either some external triggering mechanism for star formation must be present in the inner regions of these galaxies, or the intrinsic nature of the molecular clouds must vary significantly with radius. Such external factors could include passage of galactic shocks or cloud collisions which could be a funcion of radius.

7.9 Molecular Rings in Galaxies

The molecular ring at 5.5 kpc is one of the most striking features of the molecular gas distribution in the Milky Way. Similar features have been detected in NGC 7331 by these infrared data and by the CO data of Young and Scoville (1982c), who also detected a ring in NGC 2841. Although the feature is commonly referred to as a ring it should be more properly be considered a central hole since it has not been observed to decline more rapidly for larger galactocentric radii than the stellar distribution. Young and Scoville have presented arguments on the origin of these features which can be examined in the light of the higher resolution infrared data.

Ring-like structures are also seen in the HI distribution of some galaxies. M81 shows HI features well correlated with the visible spiral structure. These features reach a peak density at about

4 kpc radius and inside this region there is relatively little emission (Rots 1975, Rots and Shane 1975). M31 shows an even more pronounced ring (Unwin 1980) which appears to be related to the weak spiral features. Central gas deficiencies therefore appear to be related to the overall spiral structure of galaxies.

The spiral density wave theory, originally proposed by Lindblad (e.g. Lindblad 1959) and further developed by Lin (e.g. Lin, Yuan and Shu 1969) provides a possible explanation for the molecular ring in the Milky way. The theory attempts to explain the existence and survival of global spiral patterns in differentially rotating disks which would otherwise disrupt the pattern within a galactic revolution. The theory states that the spiral pattern can rotate with a constant angular frequency Ω_p even though the material making up the disk rotates differentially with angular frequency Ω (R), which is a function of the galactocentric distance R.

The stars may be considered to exhibit small deviations from circular motion in the form of epicyclic motions of frequency $\kappa(R)$, superimposed on the circular rotation of frequency $\Omega(R)$ (Wielen 1974). The spiral pattern can be considered as a density wave moving through the rotating disk material and can exist within the limits

$$\Omega - \kappa/m < \Omega_p < \Omega + \kappa/m$$

where m is 2 for the common two armed spial pattern. In this case the pattern can only propagate between the Lindblad inner and outer resonances defined by

inner : $\Omega_p = \Omega - \kappa/2$
outer : $\Omega_p = \Omega + \kappa/2$

Corotation, where $\Omega = \Omega_p$, also defines a resonance which effectively defines the outer limit for the spiral features and allows Ω_p to be estimated by the rotational velocity of the outermost HII regions.

The inner Lindblad resonance (ILR) has been estmated to occur at 3.2 kpc in the Milky Way (Lin, see Wielen 1974). The epicyclic frequency κ is a function of radius

$$\kappa = 2\,\Omega(1 + 1/2\; R/\Omega\; d\Omega/dR)$$

For rigid body rotation, Ω is independent of (R) and so κ is also independent of R. The condition for the ILR cannot therefore be met and so the ILR can only occur outside the rigid body rotation region although the exact location is difficult to determine (Roberts et al 1975). Sanders (1979) has performed calculations using the derived parameters for the Milky Way and has reproduced the observed hole in the molecular gas distribution between 1 and 4 kpc. Near the ILR, which can be quite broad, clouds of molecular gas execute large radial motions and lose energy in collisions though maintaining angular momentum. They therefore tend to move to larger or smaller radii leaving a depletion near the ILR.

The other mechanism, considered by Young and Scoville as being responsible for the central deficiency of molecular gas, is more efficient star formation in the inner regions of a galaxy during its initial collapse, resulting in a large central bulge. Einsato et al (1980) argue that the rotation curves of the Milky Way and other galaxies imply that the stellar disk also has a central hole as has been observed in some spiral galaxies by Kormendy (1977). This bulge can be detected both by brightess profiles and the shape of the rotation curve which exhibits a bump if a massive central bulge is present. The rotation curve of the Milky Way exhibits such a bump (e.g. Burton and Gordon 1978) and so the molecular gas distribution could be explained by either this mechanism or the ILR. Sc galaxies have been found not to have central deficiencies of molecular gas (Young and Scoville 1982a and 1982b) but as these galaxies have neither massive bulges nor ILRs, neither mechanism could be eliminated.

The Sb galaxies NGC 2841 and NGC 7331 do have central deficiencies in molecular emission (these observations and Young and Scoville

1982c) and provide evidence that the ILR is not responsible. The H α rotation curve of the inner region of NGC 7331 (Rubin et al 1965) shows a bump at 50 arcseconds and approximately rigid body rotation within this radius. Young and Scoville found the maximum CO emission at 45 arcseconds and argued that this indicated the central deficiency was related to the massive bulge and not associated with the ILR which could only occur outside 45 arcseconds. The spatial resolution, 50 arcseconds, still left some room for doubt in this interpretation.

The higher resolution infrared data show that the inner limit to the high dust densities, and by implication molecular gas densities, occurs at 30 arcseconds rather than 45 arcseconds. This radius is comfortably within the rigid body rotation region and so strengthens the conclusion of Young and Scoville that the ILR is not responsible for the central deficiency. The other galaxy shown by Young and Scoville to have a central deficiency, NGC 2841, shows molecular emission extending inwards to 45 arcseconds, which is also comfortably within the rigid body rotation region. The rotation curve of NGC 2841 (Bosma 1981) shows a bump at 120 arcseconds and rigid body rotation within this radius. When the location of the bump of the rotation curve is compared to the inner radius of the molecular gas distribution, these two galaxies appear more similar than the CO data alone suggest.

These findings suggest but do not prove that the same mechanism is responsible for the Milky Way molecular gas distribution. The calculations of Sanders (1979) indicate the spiral density wave theory can account for both the deficiency between 1 and 4 kpc and the central concentration of molecular gas. The simple model of more efficient star formation in the inner region of the Milky Way does not account for the central concentration, but the Galactic Centre is peculiar in other respects which cannot be accounted for by a simple application of spiral density wave theory. There is therefore no reason to believe that the Milky Way and these two Sb galaxies have different mechanisms responsible for the molecular gas distribution.

7.10 Conclusions

(a) NGC 7331 has a dust and molecular gas distribution which exhibits a central hole within about 3 kpc of the nucleus. No evidence is found for a feature near the nucleus similar to that at the Galactic Centre.

(b) The dust and gas surface densities remain relatively constant between 3 and 9 kpc and have a lower limit of 9 $M_{\odot}pc^{-2}$. The true value is probably a few times greater than this due to clumping and geometry effects and so is compatible with the peak value of 26 M pc found by Young and Scoville. Within 3 kpc of the nucleus the total gas surface density can be no more than 20 percent of the density in the 3 to 9 kpc region.

(c) Comparison of the dust distribution with the HI and H_2 distributions confirms that within the 3 to 9 kpc region the gas is mostly in molecular form.

(d) The higher resolution of the infrared data compared to the CO data shows that the dust and gas surface density distribution is relatively constant between 3 and 9 kpc. It does not follow either the total B luminosity or the disk B luminosity as found by Young and Scoville for Sc galaxies, implying that the star formation rate is not a function solely of gas density.

(e) The infrared and CO data are consistent when the differences in resolution are taken into account and show that the radial distribution of molecular gas in NGC 7331 is similar to that of NGC 2841.

(f) The inner extent of the dust distribution is well within the rigid body rotation region of NGC 7331 and strengthens the conclusion of Young and Scoville that the inner Lindblad resonance is not responsible for the central deficiency.

(g) The mass in the ISM in NGC 7331 inside 9 kpc is at least $2 \times 10^{9} M_{\odot}pc$ (2 percent of the total mass) and probably about twice this value.

(h) The distribution of the ISM in other galaxies can, in certain situations, be determined more accurately by near infrared observations of the extinction, rather than by lower resolution measurements of the gaseous emission.

ACKNOWLEDGEMENTS

Much of the work described in this thesis was performed while I was a member of the staff at UKIRT, Hawaii. I would like to thank the many people there, who provided encouragement and enthusiastic assistance in constructing the instrument and making the observations. Particular thanks should be given to David Beattie and Ian Gatley for providing the inspiration for this project, Charlie Telesco and Andy Longmore for useful discussions, and John Clark, Alf Nield and Jay Tsutsumi for their efforts in the workshop.

In Edinburgh I would like to thank my supervisor, David Emerson, for his advice and my colleagues at the Royal Observatory for their forbearance during the writing up of the thesis.

BIBLIOGRAPHY

Aaronson,M. : 1977, Ph. D. Thesis, Harvard University.

Aaronson,M. : 1979, in "Infrared Astronomy", IAU Symp. 96., ed. C.G. Wynn-Williams and D.P. Cruikshank, p297.

Aaronson,M., Cohen,J.G., Mould,J. and Malkan,M. :1978, Ap. J., 223,824.

Aaronson,M. and Mould,J. : 1980, Ap. J., 240, 804.

Aaronson,M., Mould,J., and Huchra,J. : 1980, Ap. J., 249, 76.

Adams,D.J., Adamson,A.J. and Giles,A.B. : 1983, M.N.R.A.S., 202, 241.

Allen,C.W. : 1973, "Astrophysical Quantities", Athlone Press, London.

Allen,D.A., and Cragg,T. : 1983, M.N.R.A.S., 203, 777.

Arp,H.C. : 1965, Ap. J., 141,43.

Baker,P.L. and Burton,W.B. : 1975, Ap. J., 198, 281.

Baldwin,J.R., Frogel,J.A. and Persson,S.E. : 1973, Ap. J., 184, 427.

Bania,T.M. : 1977, Ap. J., 216, 381.

Bania,T.M. : 1980, Ap. J., 242, 95.

Becklin,E.E, Matthews,K., Neugebauer,G., and Willner,S.P. : 1978 Ap. J., 220, 831.

Blackman,C.P. : 1979 , M.N.R.A.S., 186, 717.

Blanco,B.M., Blanco,V.M., and McCarthy,M.F. : 1978, Nature, 271, 638.

Blanco,V.M., McCarthy,M.F., and Blanco,B.M. : 1980, Ap. J., 242, 938.

Blitz,L. : 1979, Ap. J., 231, L115.

Blitz,L and Shu,F.H. : 1980, Ap. J., 238, 148.

Blitz,L. and Thaddeus,P. : 1980, Ap. J., 241, 676.

Blitz,L, Israel,F.P., Neugebauer,G., Gatley,I., Lee,T.J., and Beattie,D.H. : 1981, Ap. J., 249, 76.

Bohlin,R.C., Savage,B.D. and Drake,J.F. : 1978, Ap. J., 224, 132.

Boisse,P., Gispert,R., Coron,N., Wijnbergen,J.J., Serra,G., Ryter,C. and Puget,J.L. : 1981, Astr. Ap., 94, 265.

Boroson,T. : 1981, Ap. J. Suppl., 46, 177.

Bosma,A. : 1981a, Astr. J., 86, 1791 and 1825.

Boulanger,F., Stark,A.A. and Combes,F : 1981, Astr. Ap., 93, L1.

Bouw,G.D. and Parsons,S.B. : 1971, Coll. on Supergiant Stars, Trieste, ed. M. Hack, p22.

Burstein,D. : 1979, Ap. J., 234, 829.

Burton,W.B. : 1976, Ann. Rev. Astr. Ap., 14, 275.

Burton,W.B. : 1979, in "The Large Scale Characteristics of the Galaxy", ed. W.B.Burton, IAU Symposium 84, p257.

Burton,W.B., Gordon,M.A, Bania,T.M., Lockman,F.J. : 1975, Ap. J., 202, 30.

Burton,W.B. and Gordon,M.A. : 1976, Ap. J., 207, L189.

Burton,W.B. and Gordon,M.A. : 1978, Astr. Ap., 63, 7.

Burton,W.B. and Liszt H.S. : 1978, Ap. J., 225, 815.

Carrasco,L., Strom,S.E., and Strom,K.M. : 1973, Ap. J., 182, 95.

Carruthers, G. : 1970, Ap. J., 161, L181.

Cohen,J.G., Frogel,J.A. and Persson,S.E. : 1978, Ap. J., 222, 165.

Cohen,J.G., Frogel,J.A., Persson,S.E., and Elias,J. : 1981, Ap. J., 249, 481.

Cohen,M., and Kuhi,L.V. : 1979, Ap. J. Suppl., 41, 743.

Cohen,R.S. and Thaddeus,P. : 1977, Ap. J., 217, L155.

Cohen,R.S., Cong,H., Dame,T.M. and Thaddeus,P. : 1980, Ap. J., 239, L53.

Combes,F., Encrenaz,P.J., Lucas,R. and Weliachew,L : 1977, Astr. Ap., 61, L7.

Danver,C-G. : 1942, Ann. Obs. Lund, No. 10.

Davis,M., Feigelson,E and Latham,D.W. : 1980, Astr. J., 85, 131.

de Vaucouleurs,G. : 1948, Ann. d'Ap., 11, 247.

de Vaucouleurs,G. : 1958, Ap. J., 127, 487.

de Vaucouleurs,G. : 1978, Ap. J., 233, 730.

de Vaucouleurs,G, de Vaucouleurs,A. and Corwin,H.G. : 1976, Second Reference Catalogue of Bright Galaxies, Univ. of Texas, Austin.

Dickman,R.L. : 1978, Ap. J. Suppl., 37, 407.

Dyck, H.M. and Jones,T.J. : 1978, Astr. J., 83, 594.

Einsato,J, Tenjes,P., Barabanov,A.V. and Zasov,A.V. : 1980, Ap. Space Sci., 67, 31.

Elias,J.H : 1978a, Ap. J., 223, 859.

Elias,J.H : 1978b, Ap. J., 224, 453.

Elias,J.H : 1978c, Ap. J., 224, 857.

Elias,J.H., Frogel,J.A., Matthews,K. and Neugebauer,G. : 1982, Astr. J., 87, 1029.

Elias,J.H., Frogel,J.A., Hyland,A.R. and Jones,T.J. : 1983, Astr. J., 88, 1027.

Elmegreen,B.G. and Elmegreen,D.M. : 1981, in "Interstellar Molecules", IAU Symp. 87, ed. B.H.Andrew, p191.

Elmegreen,D.M. : 1980, Ap. J. Suppl., 43, 37.

Fall,S.M. : 1981, in "Structure and Evolution of Normal Galaxies", ed. S.M.Fall and D. Lynden-Bell, Cambridge Univ. Press.

Frankston,M. and Schild,R. : 1976, Astr. J., 81,500.

Frerking,M.A., Langer,W.D. and Wilson,R.W. : 1978, Ap. J., 262, 590.

Frogel,J.A. and Persson,S.E. : 1973, Ap. J., 192, 351.

Frogel,J.A., Persson,S.E., Aaronson,M. and Matthews,K. : 1978, Ap. J., 220, 75.

Frogel,J.A., Persson,S.E. and Cohen,J.G. : 1980 Ap. J. 240, 785.

Frogel,J.A., Blanco,V.M., McCarthy,M.F., and Cohen,J,G. : 1982 Ap. J., 252, 133.

Gallagher,J.S. and Hunter,D.A. : 1981, Astr. J., 86, 1312.

Gillett,F.C., Deneriak,E.L., and Joyce,R.R. : 1977, Opt. Eng., 16, 544.

Glass,I.S. : 1974, M.N.R.A.S., 171, 19p.

Glass,I.S. : 1976, M.N.R.A.S., 175, 191.

Glass,I,S. : 1979a, M.N.R.A.S., 186, 317

Glass,I.S. : 1979b, M.N.R.A.S., 187, 305.

Glass,I.S. and Feast : 1973, M.N.R.A.S, 163, 245.

Glass,I.S., and Penston,M.V. : 1975, M.N.R.A.S., 172, 227.

Gordon,M.A. and Burton,W.B. : 1976, Ap. J., 208, 346.

Grasdalen,G. : 1974, Ap. J., 193, 373.

Grasdalen,G., Joyce,R., Knacke,R.F.,Strom,S.E., and Strom,K.E. : 1975, Astr. J., 80, 117.

Greisen,E.W. : 1973, Ap. J., 184, 379.

Griersmith,D., Hyland,A.R. and Jones,T.J : 1982, Astr. J., 87, 1106.

Hackwell,J.A. and Schwiezer,F : 1983, Ap. J., 265, 643.

Hall,D.N.B, Aikens,R.S, Joyce,R. and McCurnin,T : 1975, Appl. Opt., 14, 450.

Hamabe,M., Kodaira,K., Okamura,S. and Takase,B. : 1979, Publ. Astr. Soc. Japan, 31, 431.

Hamabe,M., Kodaira,K., Okamura,S. and Takase,B. : 1980, Publ. Astr. Soc. Japan, 32, 197.

Hamabe,M. and Okamura,S. : 1981, Ann. Tokyo Astr. Soc. Japan, 18, 191.

Hamajima,K., Ichikawa,T., Ishida,K., Hidayat,B. and Raharto,M. : 1981, Publ. Astr. Soc. Japan, 33, 591.

Harding,P., Jones,T.J., and Rodgers,A.W. : 1981, Ap. J., 251, 530.

Hartley,K.F. : 1983, Starlink User Note 23.7.

Haslett,J.W. : 1972, IEEE Trans. on Electron Devices, 19, 943.

Hawarden,T.G., Elson,R.A.W, Longmore,A.J., Tritton,S.B. and Corwin,H.G. : 1981, M.N.R.A.S., 196, 747.

Hayakawa,S., Ito,K., Matsumoto,T., Murakami,H. and Uyama,K. : 1978, Publ. Astr. Soc. Japan, 30, 369.

Hegyi,D.G. and Gerber,G. : 1979, in "Photometry, Kinematics, Dynamics of Galaxies", ed. D.S.Evans, Univ. of Texas, Austin, Texas.

Heiles,C. and Habing,H.J. : 1974, Astr. Ap. Suppl., 14, 1.

Herzberg,G. : 1950, "Spectra of Diatomic Molecules", van Nostrand Reinhold, New York.

Hodge P.W., and Kennicutt R.C. : 1982, Astr. J., 87, 264.

Hoessel,J., and Melnick,J. : 1980, Astr. Ap., 84, 317.

Hollenbach,D.J., Werner,M.W. and Salpeter,E.E. : 1971, Ap. J., 163, 165.

Holmberg,E. : 1958, Medd. Lunds Astr. Obs., Ser. II, No. 136.

Huchtmeier,W.K., Seiradakis,J.H and Tammann,G.A. : 1980, Astr. Ap., 89, 95.

Humphreys,R.M. : 1970, Astrophys. Lett., 6, 1.

Hyland,A.R. : 1981, in "Infrared Astronomy", IAU Symp. 96, ed. G.C. Wynn-Williams and D.P.Cruikshank, p125.

Ichikawa,T., Hamajima,K., Ishida,K., Hidayat,B. and Raharto,M : 1982, Publ. Astr. Soc. Japan, 34, 231.

Innanen,K.A. : 1973, Ap. Space Sci., 22, 393.

Israel,F.P., Gatley,I., Matthews,K., and Neugebauer,G. : 1982a, Astr. Ap., 105, 229.

Israel,F.P., de Graauw,T., Lidholm,S., van de Stadt,H. and de Vries,C. : 1982b, Ap. J., 262, 100.

Jenkins,E.B., and Savage,D.B. : 1974, Ap. J., 187, 243.

Jensen,E.B. and Thuan,T.X. : 1982, Ap. J. Suppl., 50, 421.

Johnson,H.L. : 1963, Ap. J., 141, 923.

Johnson,H.L. : 1966a, Ap. J., 143, 187.

Johnson,H.L. : 1966b, Ann. Rev. Astr. Ap., 4, 193.

Johnson,H.L. : 1968, in "Nebulae and Interstellar Matter"', ed. B.M. Middlehurst and L. Haller, Univ. of Chicago Press, p167.

Jones,T.J, Hyland,A.R., Robinson,G., Smith,R and Thomas,J. : 1980, Ap. J., 242, 132.

Jones,T.J, and Hyland,A.R. : 1980, M.N.R.A.S., 192, 359.

Jones,T.J, and Hyland,A.R. : 1982, M.N.R.A.S., 200, 509.

Jones,T.J, Hyland,A.R., and Gatley,I. : 1983, M.N.R.A.S., 204, 1263

Jones,T.W. and Merrill,K. : 1976, Ap. J., 209, 509.

Kawara,K., Kozasa,T., Sato,S., Kobayashi,Y. and Jugaku,J : 1982, Publ. Astr. Soc. Japan, 34,389.

Keel,W.C. : 1983, Astr. J., 88, 1579.

Kerr,F.J. : 1968, "Stars and Stellar Systems", 7, 575, Univ. Chicago Press.

Kerr,F.J. : 1969, Ann. Rev. Astr. Ap., 7, 39.

Knapp,G.R. : 1983, in "Kinematics, Dynamics and Structure of Milky Way", ed. Shuter,W.L.H., Reidel Publ. Co.

Koorneef,J. : 1982, Astr. Ap., 107, 247.

Koorneef,J. 1983a, Astr. Ap. Suppl., 51, 489.

Koorneef,J. 1983b, Astr, Ap., 128, 84.

Kormendy,J. : 1977, Ap. J., 217, 406.

Kormendy,J. and Bruzual,A.G. : 1978, Ap. J., 223, L63.

Kormendy,J and Norman,C.A. : 1979, Ap. J., 233, 539.

Kormendy,J. and Illingworth,G. : 1982, Ap. J., 256, 460.

Krumm,N. and Salpeter,E.E. : 1979, Astr. J., 84, 1138.

Kruse,P.W. : 1970, in "Semiconductors and Semimetals", Vol. 5., ed. Willardson,R.K. and Beer,A.C., Academic Press.

Kunkel,W.E. and Bradt,H.V. : 1971, Ap. J., 170, L7.

La Rocca,A.J. : 1978, The Infrared Handbook, ed. Wolfe and Zessis, publ. Envionmental Research Institute of Michigan.

Lebofsky,M. and Rieke,G. : 1979, Ap. J., 229, 111.

Lee,T.A. : 1970, Ap. J., 162,217.

Lee,T.J., Gatley,I., Stewart,J.M., Lonsdale,C.J., and Wynn-Williams,G.C. : 1981, presentation at IAU Symp. 96.

Leung,C.M. and Liszt,H.S. : 1976, Ap. J., 208, 732.

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

Lindblad,B : 1959, Handbuch der Physik, 53, 21.

Lindblad, B. : 1942, Stockholm Obs. Ann., 14, No. 3.

Linke,R.A., Stark,A.A., Frerking,M.A. : 1981, Ap. J., 243, 147.

Liszt,H.S. and Burton,W.B. : 1978, Ap. J., 226, 790.

Liszt,H.S, Xiang,D., Burton,W.B. : 1981, Ap. J., 249, 532.

Longmore,A.J., and Sharples,R.M. : 1982, M.N.R.A.S., 201, 111.

Lonsdale,C.J., Becklin,E.E., Lee,T.J. and Stewart,J.M. : 1982, Astr. J., 87, 1819.

Lynds,B.T. : 1962, Ap. J. Suppl., 7, 1.

Lynds,B.T. : 1970, in "The Spiral Structure of our Galaxy", IAU Symp. 38, ed. W.Becker and G.Contopoulos, p26.

Manduca,A. and Bell,R.A. : 1979, Publ. Astr. Soc. Pacific, 91, 848.

Martin,P.G. : 1978, "Cosmic Dust : its Impact on Astronomy", Clarendon Press, Oxford.

Merrill,K.M. and Stein,W.A. : 1976, Publ. Astr. Soc. Pacific, 88, 285.

Mikami,T., Ishida,K., Hamajima,K. and Kawara,K. : 1982, Publ. Astr. Soc. Japan, 34, 223.

Morgan,D.H. and Nandy,K. : 1982, M.N.R.A.S., 199, 979.

Morris,M and Rickard,L.J. : 1982, Ann. Rev. Astr. Ap., 20, 517.

Mould,J. : 1976, Astr. Ap., 48, 443.

Mould,J. and Hyland,A.R. : 1976, Ap. J., 208, 399.

Murphy,H.P., Schild,R.E. and Weeks,T.C.: 1983, M.N.R.A.S., 202, 1127.

Nandy,K., Thompson,G.I., Jamar,C., Monfils,A. and Wilson,R. : 1975, Astr. Ap., 44, 195.

Nandy,K., Morgan,D.H., Willis,A.J., Wilson,R., and Gondhalekar,P.M. : 1981, M.N.R.A.S., 196, 955.

Nassau and van Albada, 1949, Ap. J., 109, 391.

Nilson,P. : 1973, Uppsala General catalogue of Galaxies, Nova Acta Ser. V, Vol. 1.

Nussbaum,A and Phillips,R.A. : 1976, "Contemporary Optics for Scientists and Engineers", Prentice Hall.

Ocli 1970 : The Infrared Handbook, Optical Coating Laboratory Inc., Santa Rosa, California.

Oda,N., Maihara,T., Sugiyami,T., and Okuda,H. : 1979, Astr. Ap., 72, 309.

Persson,S.E., Aaronson,M. and Frogel,J.A. : 1977, Astr. J., 82, 729.

Persson,S.E., Frogel,J.A. and Aaronson,M. : 1979, Ap. J. Suppl., 39, 61.

Persson,S.E., Aaronson,M., Cohen,J,G., Frogel,J.A., and Matthews,K. : 1983, Ap. J., 266, 105.

Price,S.D. : 1981, Astr. J., 86, 193.

Radeka,V. : 1967, Proc. Gatlinburg Conf. on Semiconductor Radiation Detectors and Circuits, P 393.

Radhakrishnan,V., Brooks,J.W., Goss,W.M., Murray,J.D., Schwarz,U.J.: 1972, Ap. J. Suppl., 24, 1.

Raimond,E. : 1966, Bull. Astr. Inst. Neth., 18, 191.

Rickard,L.J., Palmer,P., Morris,M., Turner,B.E., Zuckerman,B. : 1977a, Ap. J., 213, 673.

Rickard,L.J., Turner,B.E., and Palmer,P. : 1977b, Ap. J., 218, L51.

Rickard,L.J. and Palmer,P. : 1981, Astr. Ap., 102, L13.

Rieke,G.H. and Lebofsky,M.J. : 1978, Ap. J., 220, 137.

Rieke,G.,H. and Lebofsky,M.J. : 1979, Ann. Rev. Astr. Ap., 17, 477.

Rieke,G.H., Montgomery,E.F., Lebofky,M.J., and Eisenhardt,P.R. : 1981, Appl. Optics, 20, 814.

Roberts,W.W. : 1969, Ap. J., 158, 123.

Roberts,W.W., Roberts,M.S. and Shu,F.H. : 1975, Ap. J., 196, 381.

Rodgers, A.W. : 1978, Ap. J. Letters, 219, L17.

Rots,A.H. : 1975, Astr. Ap., 45, 25.

Rots,A.H. and Shane,W.W. : 1978, Astr. Ap., 45, 25.

Rowan-Robinson,M., Phillips,T.G. and White,G. : 1980, Astr. Ap., 82, 381.

Rubin,V.C., Burbridge,E.M., Burbridge,G.R., Crampin,D.J. and Prendergast,K.H : 1965, Ap. J., 141, 759.

Rydgren,A.E. : 1976, Publ. Astr. Soc. Pacific, 88, 111.

Rydgren,A.E., Strom, S.E., and Strom,K.M. : 1976, Ap. J. Suppl., 30, 307.

Sancisi,R. : 1976, Astr. Ap., 53, 159.

Sandage,A. : 1961, The Hubble Atlas of Galaxies, Carnegie Inst., Washington D.C.

Sandage,A. : 1973, Ap. J., 183, 711.

Sandage,A. and Tamann,G.A. : 1975a, Ap. J., 194, 559.

Sandage,A. and Tamann,G.A. : 1975b, Ap. J., 196, 313.

Sandage,A. and Tamann,G.A. : 1981, A Revised Shapley Ames Catalogue of Bright Galaxies, Carnegie Inst., Washington D.C.

Sanders,D.B. : 1979, in "The Large Scale Characteristics of the Galaxy", ed. W.B.Burton, IAU Symp. 84, 383.

Sanduleak,N. : 1976, Astr. J., 81, 350.

Savage,B.D. and Mathis,J.S. : 1979, Ann. Rev. Astr. Ap., 17, 73.

Schmidt,M. : 1965, "Stars and Stellar Systems", 5, 513, Univ. Chicago Press.

Schweizer,F. : 1976, Ap. J. Suppl., 31, 313.

Scoville,N.Z., and Solomon,P.M. : 1975, Ap. J., 199, L105.

Scoville,N.Z., Solomon,P.M. and Sanders,D.B. : 1979, in "The Large Scale Characteristics of the Galaxy", ed. W.B.Burton, IAU Symposium 84, p277.

Scoville,N.Z., Young,J.S. and Lucy,L.B.: 1983, Ap. J., 270, 443.

Selby,M. : 1981, "Transmission Profiles of Short Wave Set of OCLI UKIRT Filters at 77K", available from UKIRT office, Royal Obs. Edinburgh.

Sesnic,S.S. : 1972, IEEE Trans. on Electronic Devices, 19, 933.

Shane,W.W. : 1972, Astr. Ap., 16, 118.

Shull,J.M. and Beckwith,S. : 1982, Ann. Rev. Astr. Ap., 20, 163.

Simkin,S.M. : 1967, Astr. J., 72, 1032.

Simkin,S.M. : 1975, Astr. J., 80, 415.

Simonson,S.C. and Mader, G.L. : 1973, Astr. Ap., 27, 337.

Smith,W.J. 1978 : "The Infrared Handbook", ed. Wolfe and Zessis, publ. Envionmental Research Institute of Michigan.

Solomon,P.M., and Sanders,D.B.: 1977, "Giant Molecular Clouds in the Galaxy", Gregynog Astr. Workshop, ed. P.M.Solomon and M.G.Edmunds (Pergamon Press, Oxford, 1980) p41.

Spitzer,L., Drake,J.F., Jenkins,E.B., Morton,D.C., Rogerson,J.B. and York,D.G. : Ap. J., 181, L116.

Stebbins,J. and Whitford,A.E. : 1945, Ap. J., 102, 318.

Stewart,J.M. : 1983a,"UKIRT Instrumentation Computer", UKIRT Report No. 12, Royal Observatory Edinburgh.

Stewart,J.M. : 1983b,"UKIRT Spectro-photometry System", UKIRT Report No. 17, Royal Observatory Edinburgh.

Storey,J.W.V. : 1981, Rev. Sci. Inst., 52, 769.

Strom,K.M., Strom,S.E., Wells,D.C. and Romansihin,W. : 1978, Ap. J., 220, 62.

Struck-Marcell,C., and Tinsley,B.M. : 1978, Ap. J., 221, 562.

Suits,G.H. 1978 : The Infrared Handbook, ed. Wolfe and Zessis, publ. Envionmental Research Institute of Michigan.

Telesco,C.M, Gatley,I. and Stewart,J.M. : 1982, Ap. J., 263, L13.

Thuan,T.X. : 1983, AP. J., 268, 667.

Topping,J. : 1972, "Errors of Observations and Their Treatment", Chapman and Hall.

Trumpler,R.J. : 1930, Lick Obs. Bull., 14, 154

Unwin,S.C. : 1980, M.N.R.A.S., 192, 243.

van de Hulst,H.C. : 1949, Rech. Astr. Utrecht, 11, parts 1 and 2.

van der Kruit,P.C. : 1979, Astr. Ap. Suppl., 38, 15.

van der Kruit,P.C, and Searle,L. : 1981a, Astr. Ap., 95, 106.

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 Houten,C.J. : 1961, Bull. Astr. Soc. Neths.,16, 1.

Veeder,G.J. : 1974, Astr. J., 79, 1056.

Vladimirescu,A., Kaihe Zhang, Newton,A.R., Pederson,D.O., and Sangiovanni-Vincentelli : 1981, SPICE Version 2G User's Guide, Dept. of Electrical Engineering and Computer Sciences, Univ. of California, Berkeley.

Vrba,F.J., Strom,K.M., Stom,S.E. and Grasdalen,G.L. : 1975, Ap. J., 197, 77.

Warren Smith,R.F. and Berry,D.S. : 1983, M.N.R.A.S., 205, 889.

Watenabe,M., Kodaira,K., Okamura,S. : 1982, Ap. J. Suppl., 50, 1.

Watson,J. : 1970, "An Introduction to Field Effect Transistors", Siliconix Ltd.

Whitford A.E. : 1958, Astr. J., 63, 201.

Whittet,D.C.B. and van Breda,I.G. : 1980, M.N.R.A.S., 192, 467.

Wickramasinghe,N.C. : 1972, Interstellar Matter, Swiss Soc. of Astr. Ap., Second Adv. Course.

Wielen,R. : 1974, Publ. Astr. Soc. Pacific, 86, 341.

Williams,P.M. : 1983, "List of InSb Standards for UKIRT", available from UKIRT office, Royal Obs. Edinburgh.

Willner, Becklin,E.E. and Visvanathan,N. : 1972, Ap. J., 175, 699.

Wimmers,J.T. and Smith,D.S. : 1982, Proc. S.P.I.E., 364-14.

Witt,A.N., and Stephens,T.C. : 1974, Ap. J., 79, 948.

Wolfe, W.L. 1978 : The Infrared Handbook, ed. Wolfe and Zessis, publ. Envionmental Research Institute of Michigan.

Wyatt,C.L., Baker,D.J., and Frodsham,D.G. : 1974, Infrared Physics, 14, 176.

Wynn-Williams,C.G. and Becklin,E.E. : 1974, P.A.S.P., 86, 5.

Wynn-Williams,C.G., Becklin,E.E., Matthews,K., and Neugebauer,G. : 1979, M.N.R.A.S., 189, 163.

Young,J.S. and Scoville,N. : 1982a, Ap. J., 258, 467.

Young,J.S. and Scoville,N. : 1982b, Ap. J., 260, L11.

Young,J.S. and Scoville,N. : 1982c, Ap. J., 260, L41.

Zwicky,F. : 1955, Publ. Astr. Soc. Pacific, 67, 232.

APPENDIX A TABULATED DATA

The measured K magnitudes and colours (J-K and H-K) are tabulated below. Those magnitudes measured with the Mark II cryostat (10 arcsecond aperture) may be converted to magnitudes/arcsecond by the addition of 4.7+0.2. This conversion cannot be applied to those data measured with the Mark I cryostat (4 and 5 arcsecond apertures).

The colour errors are calculated as described in Section 4.5.2. When only one set of K data exists (i.e. H-K was not measured) the K errors are the statistical standard errors of the mean. When two sets of K data exist, the errors are either the statistical error or the difference in the two determinations divided by root 2, whichever is the larger.

The colours measured with the Mark I cryostat may suffer from systematic errors due to uncertainties in the zero points. Those measured with the Mark II cryostat should not suffer from systematic errors larger than 0.03 magnitudes.

" N	K	+/-	H-K	+/-	J-K	+/-
0.0	9.27	0.01	0.30	0.00	1.10	0.02
-4.8	10.29	0.01	0.32	0.01	1.17	0.00
-9.5	11.00	0.04	0.30	0.00	1.13	0.01
-14.3	11.52	0.02	0.29	0.00	1.13	0.02
-19.0	11.85	0.02	0.32	0.02	1.17	0.06
-23.8	12.10	0.02	0.38	0.01	1.29	0.02
-28.6	12.36	0.06	0.48	0.03	1.55	0.10
-33.3	12.67	0.03	0.52	0.02	1.46	0.03
-38.1	12.89	0.01	0.42	0.02	1.20	0.03
-42.9	13.15	0.03	0.41	0.05	1.17	0.07
-47.6	13.61	0.04	0.43	0.06	1.53	0.07
-52.4	13.93	0.05	0.45	0.07	1.37	0.12
-57.1	13.98	0.05	0.47	0.07	1.22	0.08
-61.9	14.20	0.06	0.21	0.10	1.02	0.09
-66.7	14.09	0.05	0.20	0.06	0.98	0.07
4.8	10.16	0.04	0.26	0.04	1.02	0.05
9.5	11.00	0.08	0.25	0.02	1.04	0.04
14.3	11.50	0.03	0.29	0.00	1.03	0.01
19.0	11.78	0.04	0.27	0.02	1.05	0.00
23.8	12.02	0.04	0.30	0.02	1.08	0.02
28.6	12.22	0.04	0.37	0.02	1.19	0.04
33.3	12.52	0.06	0.34	0.03	1.19	0.04
38.1	12.75	0.05	0.40	0.03	1.20	0.04
42.9	13.06	0.06	0.41	0.02	1.22	0.07
47.6	13.40	0.05	0.28	0.02	1.08	0.08
52.4	13.85	0.05	0.36	0.06	0.88	0.07
57.1	14.14	0.06	0.25	0.09	0.90	0.06

Table A.1 The N/S profile of NGC 7331 measured in a 5 arcsecond aperture on 26 Aug. 1981. The standard star used was BS 8541. The K magnitudes are the mean of those measured with J and H.

" E	K	+/-	J-K	+/-
0.0	9.21	0.00	1.10	0.01
-4.8	10.46	0.02	1.16	0.03
-9.5	11.44	0.01	1.26	0.01
-14.3	12.12	0.03	1.47	0.06
-19.0	12.65	0.04	1.45	0.05
-23.8	13.16	0.02	1.43	0.05
-28.6	13.58	0.08	1.25	0.09
-33.3	14.29	0.06	0.99	0.07
-38.1	14.82	0.07	1.08	0.24
4.8	10.82	0.02	1.07	0.03
9.5	11.53	0.01	1.11	0.02
14.3	11.95	0.02	1.18	0.03
19.0	12.70	0.01	1.10	0.04
23.8	13.31	0.02	1.08	0.08
28.6	13.78	0.05	1.15	0.09
33.3	13.95	0.03	1.25	0.06
38.1	14.41	0.07	1.40	0.08

Table A.2 The E/W profile of NGC 7331 observed in a 5 arcsecond aperture on 26 Aug. 1981. The standard star used was BS 8541.

" NW	K	+/-	J-K	+/-	H-K	+/-
0	8.69	0.00	1.04	0.00	0.30	0.01
10	9.82	0.10	1.02	0.00	0.29	0.01
20	10.55	0.07	1.01	0.01	0.29	0.01
30	10.95	0.03	1.09	0.01	0.32	0.01
40	11.24	0.03	1.25	0.01	0.40	0.01
50	11.67	0.09	1.16	0.01	0.38	0.02
60	12.12	0.05	1.14	0.03	0.34	0.02
70	12.42	0.03	1.08	0.02	0.35	0.03
80	12.83	0.11	1.10	0.03	0.38	0.03
90	12.81	0.10	0.87	0.03	0.26	0.04
100	13.21	0.04	1.05	0.05	0.24	0.04
110	13.37	0.06	0.98	0.06	0.31	0.04
120	13.77	0.05	0.95	0.07	0.26	0.04
130	14.14	0.06	1.09	0.09	0.26	0.05
-10	9.73	0.04	1.04	0.01	0.31	0.01
-20	10.48	0.02	1.04	0.01	0.30	0.01
-30	10.88	0.02	1.10	0.01	0.35	0.01
-40	11.19	0.01	1.14	0.02	0.37	0.01
-50	11.67	0.02	1.15	0.01	0.38	0.01
-60	11.90	0.02	1.25	0.01	0.39	0.01
-70	12.31	0.02	1.21	0.01	0.35	0.02
-80	12.64	0.05	1.26	0.04	0.32	0.03
-90	13.01	0.05	1.11	0.02	0.29	0.03
-100	13.58	0.11	1.04	0.06	0.28	0.04
-110	13.77	0.03	0.96	0.05	0.39	0.04
-120	13.92	0.03	1.00	0.06	0.33	0.06
-130	14.04	0.05	1.04	0.09	0.34	0.06

Table A.3 The major axis profile of NGC 7331 observed in a 10 arcsecond aperture on 5 August 1982. The nucleus magntudes were taken from the measurements of 4 August.

" NW	K	+/-	J-K	+/-
0.0	8.64	0.00	1.03	0.00
4.8	8.98	0.00	1.04	0.00
9.5	9.74	0.00	1.05	0.01
14.3	10.11	0.01	1.04	0.01
19.0	10.43	0.01	1.05	0.01
23.8	10.63	0.01	1.05	0.01
28.6	10.79	0.01	1.15	0.02
33.3	10.99	0.01	1.14	0.02
38.1	11.09	0.02	1.21	0.02
42.9	11.31	0.02	1.16	0.03
47.6	11.54	0.02	1.20	0.03
52.4	11.72	0.03	1.13	0.04
57.1	11.86	0.03	1.21	0.04
62.1	11.90	0.02	1.23	0.04
66.2	11.99	0.00	1.15	0.02
71.3	12.18	0.02	1.18	0.00
76.7	12.42	0.02	1.18	0.04
80.8	12.55	0.02	1.29	0.05
85.9	12.73	0.00	1.14	0.05
90.0	12.74	0.03	1.20	0.05
95.3	13.10	0.03	1.05	0.06
100.4	13.56	0.04	0.93	0.07
109.6	13.74	0.03	1.07	0.04
119.1	13.89	0.05	1.09	0.06
128.3	14.03	0.04	0.92	0.08

Table A.4 The fully sampled major axis profile of NGC 7331 observed at J and K in a 10 arcsecond aperture on 12 July 1982. The standard star used was HD 203856.

" NE	K	+/-	J-K	+/-	H-K	+/-
0	8.65	0.02	1.04	0.00	0.30	0.01
5	9.15	0.07	1.04	0.00	0.31	0.01
10	10.13	0.05	1.03	0.01	0.31	0.01
15	10.89	0.02	1.04	0.01	0.31	0.01
20	11.47	0.04	1.02	0.01	0.32	0.03
25	12.08	0.02	1.03	0.05	0.26	0.01
30	12.49	0.03	0.95	0.01	0.28	0.02
35	12.80	0.03	0.88	0.06	0.25	0.02
40	13.07	0.04	1.00	0.05	0.28	0.04
45	13.33	0.05	0.91	0.06	0.23	0.04
50	13.56	0.05	0.97	0.08	0.24	0.06
55	13.82	0.07	0.97	0.09	0.31	0.06
60	14.09	0.07	0.76	0.11	0.28	0.05
-5	9.15	0.18	1.07	0.00	0.32	0.02
-10	10.42	0.04	1.18	0.01	0.36	0.01
-15	11.15	0.07	1.36	0.04	0.43	0.01
-20	11.73	0.07	1.35	0.01	0.39	0.01
-25	12.15	0.07	1.19	0.04	0.31	0.02
-30	12.38	0.05	1.03	0.02	0.29	0.02
-35	12.72	0.03	1.05	0.05	0.33	0.03
-40	13.13	0.04	1.10	0.07	0.31	0.02
-45	13.33	0.05	1.12	0.07	0.29	0.03
-50	13.51	0.09	1.12	0.11	0.32	0.06
-55	13.66	0.06	1.00	0.09	0.32	0.05
-60	13.61	0.08	0.94	0.08	0.22	0.05

Table A.5 The minor axis profile of NGC 7331 observed in a 10 arcsecond aperture on 4 August 1982. The standard star used was GL 748.

"	K	+/-	J-K	+/-
0.0	9.76	0.01	1.15	0.01
-3.0	10.39	0.09	1.17	0.01
-6.1	11.19	0.03	1.14	0.01
-6.1	11.13	0.04	1.17	0.00
-9.1	11.61	0.05	1.20	0.01
-12.1	11.88	0.05	1.09	0.05
-15.2	12.05	0.03	1.11	0.02
-18.2	12.24	0.01	1.08	0.02
-21.2	12.29	0.03	1.08	0.03
-24.2	12.39	0.02	1.10	0.07
-27.3	12.53	0.04	1.02	0.05
-30.3	12.64	0.01	0.96	0.04
-33.3	12.75	0.04	1.01	0.05
-36.4	12.77	0.05	1.08	0.06
-39.4	12.77	0.06	1.27	0.08
-42.4	12.82	0.04	1.13	0.07
-45.5	12.90	0.05	1.14	0.09
-48.5	12.93	0.04	1.33	0.08
-51.5	12.91	0.07	1.27	0.09
-54.5	12.87	0.09	1.36	0.11
3.0	10.33	0.01	1.14	0.01
6.1	11.22	0.02	1.09	0.02
9.1	11.63	0.04	1.07	0.06
12.1	12.00	0.01	1.13	0.02
15.2	12.12	0.01	1.10	0.03
18.2	12.28	0.01	1.14	0.02
21.2	12.40	0.01	1.14	0.02
24.2	12.46	0.03	1.16	0.03
27.3	12.57	0.01	1.19	0.04
30.3	12.68	0.01	1.09	0.05
33.3	12.77	0.02	1.19	0.04
36.4	12.87	0.03	1.10	0.03
39.4	12.90	0.01	1.09	0.05
42.4	12.88	0.03	1.16	0.05

Table A.6 The major axis profile of NGC 4565 observed in a 4 arcsecond aperture on 30 and 31 January 1982. The standard star used was BS 4550.

"	K	+/-	J-K	+/-
0.0	9.87	0.02	1.15	0.00
3.0	10.93	0.03	1.40	0.04
6.1	11.96	0.02	1.48	0.04
9.1	12.44	0.02	1.34	0.06
12.1	12.85	0.02	1.14	0.04
15.2	13.24	0.03	1.11	0.06
18.2	13.67	0.03	1.07	0.09
21.2	14.18	0.05	1.32	0.11
-3.0	11.07	0.01	1.08	0.02
-6.1	11.76	0.01	0.97	0.02
-9.1	12.30	0.04	0.97	0.04
-12.1	12.83	0.01	0.85	0.05
-15.2	13.29	0.04	0.94	0.08
-18.2	13.78	0.04	1.17	0.12

Table A.7 The minor axis profile of NGC 4565 observed on 31 January 1982 in a 4 arcsecond . The standard star used was BS 4550.

" NW	K	+/-	J-K	+/-	H-K	+/-
0.0	11.31	0.03	1.15	0.01	0.41	0.02
4.3	11.69	0.04	1.51	0.02	0.54	0.02
8.6	12.32	0.02	1.70	0.09	0.66	0.02
12.8	12.62	0.03	1.55	0.05	0.59	0.02
17.2	12.80	0.03	1.17	0.04	0.36	0.05
21.4	13.16	0.05	1.21	0.08	0.28	0.04
25.7	13.65	0.06	1.01	0.08	0.28	0.09
30.1	14.23	0.06	0.91	0.12	0.35	0.08
-4.3	11.48	0.01	1.11	0.01	0.34	0.01
-8.6	11.89	0.01	1.04	0.04	0.34	0.02
-12.9	12.35	0.02	1.02	0.02	0.28	0.04
-17.2	12.77	0.04	1.00	0.06	0.20	0.08
-21.4	13.25	0.04	1.02	0.06	0.32	0.06
-25.7	13.74	0.05	1.06	0.09	0.42	0.07

Table A.8 The perpendicular profile of NGC 4565, parallel to the minor axis and intersecting the major axis 9 arcseconds SE of the nucleus. A 4 arcsecond aperture was used and the standard star was BS 4550.

" N	K	+/-	J-K	+/-
0.0	10.42	0.03	0.98	0.01
3.6	10.98	0.03	0.98	0.00
7.1	11.72	0.02	0.95	0.01
10.7	12.49	0.03	0.90	0.01
14.3	13.03	0.03	0.91	0.05
17.9	13.34	0.03	0.88	0.05
19.0	13.39	0.02	0.97	0.05
21.4	13.52	0.03	0.96	0.04
25.0	13.74	0.02	0.92	0.04
28.6	13.92	0.04	0.85	0.05
32.1	13.98	0.05	1.01	0.07
35.7	14.20	0.03	0.82	0.05
39.3	14.35	0.03	0.86	0.06
-3.6	11.15	0.02	0.98	0.01
-7.1	11.70	0.01	0.95	0.02
-10.7	12.34	0.01	0.98	0.03
-14.3	13.01	0.01	0.99	0.01
-17.9	13.36	0.01	0.97	0.03
-21.4	13.56	0.01	1.03	0.06
-25.0	13.73	0.02	1.02	0.05
-28.6	13.87	0.02	0.95	0.06
-32.1	13.92	0.03	1.10	0.05
-35.7	14.17	0.03	1.00	0.08
-39.3	14.29	0.02	1.02	0.08

Table A.9 The N/S profile of NGC 4216 observed in a 4 arcsecond aperture on 29 January 1982. The standard star used was BS 4550.

" E	K	+/-	J-K	+/-
0.0	10.42	0.03	0.98	0.01
-3.6	11.43	0.04	0.89	0.01
-7.1	12.55	0.04	0.88	0.03
-10.7	13.22	0.06	0.85	0.04
-14.3	13.68	0.04	0.87	0.05
-17.9	14.18	0.05	0.85	0.04
-21.4	14.57	0.04	0.86	0.06
-25.0	14.90	0.04	0.98	0.13
0.0	10.42	0.03	0.98	0.01
3.6	10.51	0.00	1.01	0.00
7.1	11.80	0.01	1.00	0.01
10.7	12.81	0.04	1.00	0.02
14.3	13.44	0.02	1.14	0.03
17.9	13.95	0.02	1.16	0.06
21.4	14.29	0.03	1.14	0.06
25.0	14.69	0.04	0.73	0.07

Table A.10 The E/W profile of NGC 4216 observed in a 4 arcsecond aperture on 29 January 1982. The standard star used was BS 4550.

" NE	K	+/-	J-K	+/-
0.0	9.61	0.02	1.51	0.00
5.1	10.70	0.04	1.28	0.01
10.1	12.07	0.01	1.19	0.02
15.2	12.98	0.03	1.21	0.01
20.2	13.71	0.04	1.08	0.06
26.9	14.31	0.06	1.08	0.08
0.0	9.61	0.02	1.51	0.00
-5.1	10.84	0.01	1.24	0.00
-10.1	12.16	0.02	1.10	0.04
-15.2	12.98	0.02	1.17	0.04
-20.2	13.64	0.03	1.16	0.05

Table A.11 The NE/SW profile of NGC 7814 observed in a 5 arcsecond aperture on 31 August 1981. The standard star used was BS 4550.

" NW	K	+/-	J-K	+/-
0.0	9.55	0.01	1.51	0.00
5.1	10.35	0.03	1.45	0.00
10.1	11.39	0.02	1.39	0.02
15.2	11.99	0.01	1.40	0.01
20.2	12.43	0.03	1.38	0.04
26.9	12.80	0.02	1.38	0.04
30.3	12.93	0.03	1.40	0.05
-5.1	10.32	0.03	1.46	0.01
-10.1	11.32	0.01	1.36	0.01
-15.2	11.94	0.00	1.35	0.02
-20.2	12.51	0.02	1.32	0.03
-25.3	12.74	0.02	1.37	0.05
-30.3	13.00	0.02	1.25	0.04
-35.4	13.16	0.02	1.32	0.04
-40.4	13.31	0.01	1.39	0.05

Table A.12 The NW/SE profile of NGC 7814 observed in a 5 arcsecond aperture on 31 August 1981. The standard star used was BS 4550.

	15" E	10" E	5" E	0	5" W	10" W	15" W
25" N	11.96 0.97 0.07	11.58 0.90 0.05	10.98 1.10 0.03	10.71 1.07 0.02	10.63 1.07 0.02	10.85 1.12 0.03	11.04 1.31 0.04
20" N	11.75 0.94 0.06	11.22 1.08 0.04	10.82 1.04 0.03	10.43 1.01 0.02	10.41 1.06 0.02	10.65 1.12 0.02	10.99 1.30 0.04
15" N	11.50 0.93 0.04	10.84 1.10 0.03	10.44 1.07 0.02	10.16 1.06 0.01	10.15 1.05 0.01	10.60 1.16 0.02	11.02 1.32 0.04
10" N	11.27 0.97 0.04	10.67 1.06 0.02	10.09 1.04 0.01	9.65 1.08 0.01	9.78 1.08 0.01	10.51 1.17 0.02	11.05 1.31 0.04
5" N	11.21 0.86 0.03	10.40 1.02 0.02	9.60 1.04 0.01	8.86 1.09 0.00	9.47 1.13 0.01	10.59 1.21 0.02	10.99 1.23 0.04
0	10.87 1.09 0.03	10.33 1.00 0.02	9.31 1.01 0.01	8.68 1.04 0.00	9.15 1.06 0.01	10.38 1.18 0.02	11.11 1.33 0.04
5" S	10.77 1.04 0.02	10.20 1.01 0.01	9.50 1.01 0.01	9.24 1.03 0.01	9.77 1.07 0.01	10.77 1.28 0.03	11.27 1.37 0.05
10" S	10.94 0.99 0.03	10.44 0.99 0.02	9.92 1.04 0.01	9.89 1.03 0.01	10.40 1.07 0.02	11.06 1.35 0.04	11.47 1.37 0.06
15" S	10.92 0.98 0.03	10.47 1.01 0.02	10.26 1.04 0.02	10.34 1.05 0.02	10.74 1.24 0.03	11.22 1.43 0.05	11.69 1.28 0.07
20" S	10.68 1.21 0.03	10.60 1.04 0.02	10.53 1.08 0.02	10.68 1.09 0.02	11.04 1.27 0.04	11.50 1.29 0.06	11.89 1.22 0.08
25" S	10.68 1.33 0.03	10.72 1.08 0.02	10.71 1.10 0.02	10.88 1.20 0.03	11.29 1.39 0.05	11.60 1.29 0.06	12.05 1.20 0.09

Table A.13 NGC 7331 small area map in J-K observed on 11 July 1982. At each position is indicated the K magnitude, the J-K colour and the error of the colour. The standard star used was HD 203856.

	15" E	10" E	5" E	0	5" W	10" W	15" W
25" N	12.13	11.66	11.24	10.84	10.76	10.91	11.07
	0.30	0.22	0.30	0.28	0.30	0.30	0.46
	0.06	0.04	0.03	0.02	0.02	0.02	0.02
20" N	11.94	11.43	10.92	10.64	10.54	10.79	11.13
	0.24	0.37	0.31	0.26	0.27	0.34	0.40
	0.05	0.03	0.02	0.02	0.01	0.02	0.03
15" N	11.67	11.13	10.66	10.30	10.25	10.56	11.13
	0.31	0.29	0.29	0.30	0.28	0.35	0.33
	0.04	0.03	0.02	0.01	0.01	0.01	0.03
10" N	11.24	10.94	10.32	9.87	9.98	10.40	11.12
	0.35	0.27	0.30	0.30	0.29	0.32	0.40
	0.03	0.02	0.01	0.01	0.01	0.01	0.03
5" N	11.06	10.70	10.00	9.21	9.60	10.31	11.10
	0.32	0.29	0.29	0.30	0.30	0.30	0.38
	0.02	0.02	0.01	0.00	0.01	0.01	0.02
0	10.93	10.42	9.29	8.69	9.34	10.36	11.30
	0.32	0.29	0.31	0.31	0.32	0.35	0.47
	0.02	0.01	0.00	0.00	0.00	0.01	0.03
5" S	10.76	10.34	9.65	9.10	9.80	10.34	11.29
	0.34	0.29	0.31	0.31	0.32	0.36	0.49
	0.02	0.01	0.01	0.00	0.01	0.01	0.03
10" S	10.79	10.34	9.96	9.87	10.35	10.84	11.45
	0.30	0.32	0.31	0.31	0.35	0.39	0.44
	0.02	0.01	0.01	0.01	0.01	0.02	0.03
15" S	10.82	10.51	10.20	10.21	10.76	11.07	11.65
	0.35	0.31	0.30	0.32	0.35	0.39	0.41
	0.02	0.01	0.01	0.01	0.02	0.02	0.04
20" S	10.86	10.62	10.51	10.64	10.94	11.35	11.85
	0.34	0.28	0.32	0.35	0.41	0.40	0.35
	0.02	0.02	0.01	0.02	0.02	0.03	0.05
25" S	10.99	10.74	10.71	10.87	11.21	11.55	11.86
	0.35	0.31	0.32	0.42	0.49	0.46	0.37
	0.02	0.02	0.02	0.02	0.03	0.04	0.05

Table A.14 NGC 7331 small area map in H-K observed on 3 August 1982. At each position is indicated the K magnitude, the H-K colour and the error of the colour. The standard star used was HD 201941.

	40" E	30" E	20" E	10" E	0	10" W	20" W	30" W	40" W
60" N	14.69 0.41 0.88	13.95 0.70 0.48	13.50 1.16 0.38	12.85 1.22 0.22	11.94 0.93 0.08	11.85 1.09 0.08	11.90 1.31 0.09	12.63 1.17 0.17	13.30 1.50 0.39
50" N	13.37 1.75 0.48	13.54 1.29 0.42	12.97 1.33 0.26	12.55 1.14 0.16	11.61 1.01 0.06	11.45 1.16 0.06	11.99 1.09 0.09	12.40 1.09 0.13	12.89 1.29 0.23
40' N	13.36 1.30 0.36	13.73 0.83 0.41	12.77 1.24 0.20	12.22 1.06 0.11	11.21 1.11 0.05	11.10 1.21 0.04	11.66 1.27 0.07	12.56 1.22 0.16	13.12 1.19 0.27
30" N	13.45 1.24 0.38	12.96 1.05 0.22	12.50 1.05 0.14	11.68 1.03 0.07	10.88 1.09 0.03	10.91 1.15 0.03	11.54 1.41 0.07	12.38 1.31 0.15	13.07 1.56 0.32
20" N	13.24 1.07 0.29	13.33 0.88 0.29	12.21 1.05 0.11	11.14 1.06 0.04	10.47 1.02 0.02	10.72 1.15 0.03	11.53 1.32 0.07	12.32 1.40 0.15	13.23 0.93 0.27
10" N	12.92 1.02 0.21	12.65 1.00 0.16	11.76 1.06 0.07	10.65 1.08 0.03	9.72 1.05 0.01	10.61 1.16 0.03	11.73 1.19 0.08	12.51 1.20 0.16	13.37 1.14 0.33
0	12.69 1.49 0.22	12.46 1.11 0.14	11.48 1.02 0.06	10.17 0.97 0.02	8.71 1.02 0.00	10.63 1.21 0.03	11.91 1.30 0.10	12.67 1.00 0.16	13.25 1.51 0.37
10" S	13.29 0.84 0.27	12.46 1.05 0.14	11.34 1.03 0.05	10.31 1.00 0.02	9.92 1.03 0.01	11.08 1.26 0.04	12.19 1.20 0.12	12.84 0.87 0.18	13.21 1.70 0.40
20" S	13.27 0.73 0.26	12.42 1.15 0.14	11.23 1.07 0.05	10.61 1.03 0.02	10.66 1.12 0.03	11.59 1.34 0.07	12.33 1.12 0.13	12.96 1.42 0.27	13.60 1.74 0.60
30" S	13.15 1.22 0.28	12.33 1.06 0.12	11.23 1.17 0.05	10.92 1.14 0.04	11.21 1.34 0.05	12.05 1.00 0.09	12.79 0.88 0.17	13.18 1.08 0.27	13.18 1.58 0.36
40" S	13.83 0.44 0.40	12.34 1.09 0.13	11.39 1.20 0.06	11.24 1.17 0.05	11.74 1.23 0.08	12.46 1.39 0.17	13.09 0.94 0.24	13.63 1.10 0.42	13.89 1.87 0.86
50" S	13.09 1.22 0.27	12.35 1.08 0.13	11.73 1.15 0.07	11.79 1.23 0.08	12.49 1.15 0.15	12.84 1.21 0.21	13.23 1.21 0.30	13.97 1.44 0.69	13.79 1.34 0.55
60" S	13.83 1.35 0.57	12.38 1.00 0.13	11.98 1.13 0.09	12.22 1.25 0.12	12.61 1.09 0.16	13.03 1.21 0.25	13.49 0.86 0.33	13.38 0.67 0.28	13.94 4.00 5.74

Table A.15 NGC 7331 large area map in J-K observed on 11 July 1982. At each position is indicated the K magnitude, the J-K colour and the error of the colour. The standard star used was HD 203856.

	40" E	30" E	20" E	10" E	0	10" W	20" W	30" W	40" W
60" N	14.70	13.95	13.39	13.17	12.33	12.14	12.15	12.76	13.33
	-0.49	0.32	0.27	0.24	0.46	0.40	0.44	0.33	0.22
	0.47	0.27	0.16	0.13	0.06	0.05	0.05	0.09	0.15
50" N	14.43	13.88	13.02	12.58	11.86	11.65	12.09	12.37	13.20
	0.06	0.14	0.32	0.34	0.35	0.40	0.34	0.34	0.37
	0.39	0.24	0.11	0.08	0.04	0.03	0.05	0.06	0.14
40" N	13.44	13.50	12.83	12.05	11.23	11.27	11.85	12.55	13.17
	0.73	-0.01	0.30	0.29	0.40	0.40	0.40	0.40	0.70
	0.19	0.16	0.09	0.05	0.02	0.02	0.04	0.08	0.15
30" N	13.59	13.27	12.51	11.53	10.95	11.07	11.80	12.57	13.36
	0.57	0.18	0.32	0.36	0.33	0.39	0.42	0.33	0.19
	0.21	0.14	0.07	0.03	0.02	0.02	0.04	0.08	0.15
20" N	13.76	13.16	12.01	10.96	10.53	10.93	11.74	12.47	13.19
	0.18	0.41	0.35	0.36	0.31	0.37	0.46	0.28	0.44
	0.22	0.13	0.05	0.02	0.01	0.02	0.04	0.07	0.14
10" N	13.37	12.61	11.61	10.57	9.82	10.80	11.85	12.54	13.22
	0.21	0.23	0.33	0.32	0.30	0.38	0.39	0.43	0.50
	0.15	0.08	0.03	0.01	0.01	0.01	0.04	0.08	0.14
0	13.00	12.65	11.20	9.92	8.87	10.82	12.04	12.65	13.36
	0.36	0.14	0.35	0.31	0.31	0.37	0.38	0.38	0.22
	0.11	0.08	0.02	0.01	0.00	0.02	0.05	0.08	0.15
10" S	12.84	12.28	11.23	10.11	9.91	11.19	12.17	12.63	13.35
	0.33	0.32	0.34	0.31	0.32	0.43	0.41	0.34	0.47
	0.10	0.06	0.02	0.01	0.01	0.02	0.05	0.08	0.16
20" S	12.92	12.12	11.15	10.51	10.72	11.58	12.53	12.95	13.40
	0.63	0.31	0.39	0.31	0.39	0.43	0.25	0.34	0.45
	0.11	0.05	0.02	0.01	0.01	0.03	0.07	0.11	0.17
30" S	13.32	11.95	11.16	10.85	11.28	12.08	12.70	13.23	13.93
	0.33	0.37	0.36	0.35	0.46	0.37	0.42	0.45	0.34
	0.15	0.04	0.02	0.02	0.02	0.05	0.09	0.14	0.27
40" S	13.15	11.98	11.34	11.19	11.82	12.57	13.02	13.47	13.01
	0.35	0.46	0.38	0.36	0.34	0.21	0.44	0.44	0.33
	0.13	0.05	0.02	0.02	0.04	0.07	0.12	0.18	0.11
50" S	13.49	12.10	11.64	11.71	12.32	12.83	13.30	13.90	14.33
	0.11	0.35	0.38	0.37	0.36	0.28	0.22	0.28	0.36
	0.17	0.05	0.03	0.03	0.06	0.09	0.14	0.25	0.38
60" S	12.57	12.37	11.87	12.14	12.77	13.14	13.28	13.07	14.82
	0.29	0.36	0.41	0.47	0.21	0.16	0.48	0.33	-0.03
	0.07	0.06	0.04	0.05	0.09	0.12	0.15	0.12	0.55

Table A.16 NGC 7331 large area map in H-K observed on 4 August 1982. At each position is indicated the K magnitude, the H-K colour and the error of the colour. The standard star used was GL 748.

	12" NE	8" NE	4" NE	0	4" SW	8" SW	12" SW
25" SE	12.27 1.24 0.15	11.96 1.35 0.12	11.71 1.38 0.10	11.70 1.11 0.08	11.67 1.08 0.08	11.81 0.93 0.08	12.22 0.98 0.12
21" SE	12.21 1.37 0.16	11.93 1.33 0.12	11.68 1.34 0.09	11.51 1.03 0.06	11.60 1.08 0.07	11.83 0.91 0.08	12.07 0.93 0.10
17" SE	12.13 1.44 0.15	11.88 1.66 0.15	11.49 1.63 0.10	11.38 1.08 0.06	11.38 1.04 0.06	11.63 0.93 0.07	11.94 0.76 0.08
12" SE	12.05 1.35 0.13	11.82 1.67 0.14	11.39 1.34 0.07	11.12 1.21 0.05	11.21 1.00 0.05	11.50 0.86 0.06	11.79 0.96 0.08
8" SE	12.03 1.49 0.15	11.73 1.64 0.13	11.10 1.46 0.06	10.71 1.09 0.03	10.90 1.05 0.04	11.31 0.97 0.05	11.77 0.95 0.08
4" SE	11.88 1.41 0.12	11.54 1.75 0.12	10.58 1.43 0.04	9.49 1.15 0.01	10.43 0.97 0.02	11.10 0.95 0.04	11.66 0.85 0.06
0				9.31 1.05 0.01			
4" NW	11.79 1.37 0.11	11.45 1.52 0.09	10.66 1.34 0.04	10.03 1.07 0.02	10.51 1.01 0.02	11.15 0.92 0.04	11.62 0.88 0.06
8" NW	11.92 1.39 0.12	11.64 1.53 0.11	11.23 1.46 0.07	10.82 1.02 0.03	10.96 0.95 0.04	11.25 0.96 0.05	11.68 0.86 0.07
12" NW	12.11 1.26 0.13	11.85 1.56 0.13	11.58 1.43 0.09	11.24 0.97 0.05	11.20 1.02 0.05	11.46 0.98 0.06	11.80 0.87 0.07
17" NW	12.17 1.36 0.15	11.90 1.64 0.15	11.70 1.28 0.09	11.38 0.99 0.05	11.35 0.90 0.05	11.58 0.97 0.06	11.91 0.87 0.08
21" NW	12.28 1.11 0.14	12.05 1.47 0.15	11.82 1.62 0.14	11.56 1.01 0.07	11.45 0.95 0.06	11.72 0.98 0.07	12.05 0.93 0.10
25" NW	12.59 1.12 0.18	12.41 1.57 0.22	12.07 1.43 0.15	11.70 0.99 0.07	11.56 1.02 0.07	11.87 0.92 0.08	12.23 0.83 0.11

Table A.17 NGC 4565 map in J-K observed on 1 February 1982. At each position is indicated the K magnitude, the J-K colour and the error of the colour. The standard star used was BS 4550.

	20" E	10" E	0	10" W	20" W
60" N			15.64 0.46 1.67	13.29 2.26 0.54	12.10 1.21 0.09
50" N			14.33 0.83 0.56	12.87 0.96 0.15	11.63 1.30 0.06
40" N		15.92 1.86 4.42	13.58 1.29 0.35	11.96 1.18 0.07	11.38 1.56 0.05
30" N	15.90 2.12 5.30	15.32 0.47 1.25	12.69 1.10 0.14	11.30 1.29 0.04	11.65 1.59 0.07
20" N	15.67 -0.20 1.55	13.46 1.11 0.29	11.68 1.11 0.05	11.10 1.47 0.04	12.18 1.39 0.10
10" N	14.17	12.48 1.08 0.11	10.75 1.23 0.02	11.23 1.59 0.05	12.61 1.18 0.13
0	13.75 0.72 0.32	11.64 0.94 0.05	10.13 1.50 0.02	11.74 1.41 0.07	13.14 1.12 0.21
-10" S	12.88 1.12 0.17	11.24 1.22 0.04	11.01 1.66 0.04	12.60 1.27 0.14	14.54 0.27 0.58
-20" S	12.18 1.43 0.10	11.16 1.40 0.04	11.97 1.36 0.08	13.42 0.87 0.25	14.50 2.33 1.74
-30" S	11.73 1.11 0.06	11.45 1.57 0.06	12.75 1.43 0.18	14.50 1.84 1.18	14.89 0.61 0.87
-40" S	11.37 1.33 0.05	12.18 1.59 0.12	13.68 1.03 0.33	15.09 0.09 0.94	15.37
-50" S	11.60 1.40 0.06	12.76 1.25 0.16	15.05 -0.31 0.87	15.50 0.31 1.42	
-60" S	12.12 1.33 0.09	13.66 0.97 0.32	15.41 0.89 1.55		

Table A.18 NGC 5907 map in J-K through a 10 arcsecond aperture on 4 August 1982. At each position is indicated the K magnitude, the J-K colour and the error of the colour. The standard star used was GL 748.

THE ASTROPHYSICAL JOURNAL, 263:L13–L17, 1982 December 1

THE DISTRIBUTION OF INFRARED OBSCURATION IN NGC 7331: EVIDENCE FOR A MASSIVE MOLECULAR RING

C. M. TELESCO
Institute for Astronomy, University of Hawaii; and Space Science Division, NASA Ames Research Center
AND
IAN GATLEY AND J. M. STEWART
United Kingdom Infrared Telescope Unit
Received 1982 April 20; accepted 1982 August 18

ABSTRACT

Spatial variations in the near-infrared colors of the Sb galaxy NGC 7331 produced by interstellar dust indicate the presence of a massive ring of molecular gas ($6 \times 10^8\ M_\odot$) surrounding the nuclear bulge. The density, mass, radius, and thickness of this ring are comparable to similar structures observed within the Galaxy and M81.

Subject headings: galaxies: photometry — infrared: general — interstellar: matter

I. INTRODUCTION

A well-known advantage of observations in the infrared is that the effects of interstellar extinction are much less than in the visible (e.g., Aaronson 1979). In this *Letter* we present the results of an experiment designed to take advantage of this fact; we have investigated the possibility that dusty regions in galaxies are reddened by a very large amount, corresponding to a visual extinction well in excess of 1 mag.

The motivation for this work stems from molecular line astronomy. Dense molecular clouds, opaque at visible wavelengths because of large associated column densities of dust, are common in the Galaxy. In addition, their distribution appears to peak at a galactocentric radius of about 5 kpc (Gordon and Burton 1976). Viewed from outside of the Galaxy, this large-scale "ring" of high-density, neutral material might be expected to affect the observed colors of the starlight. In the past, *UBV* surface photometry has been used to investigate the distribution of dust in galaxies (e.g., Simkin 1976; Holmberg 1975). The present experiment is a straightforward extension of that technique applicable to regions of higher extinction.

NGC 7331 is a normal, nearly edge-on Sb I–II galaxy (Sandage and Tammann 1981). Conspicuous dust lanes associated with the spiral arms are evident (Fig. 1; Sandage 1961). At a distance of 20 Mpc (for $v_0 = 1114$ km s^{-1} [Sandage and Tammann 1981] and $H_0 = 55$ km s^{-1} Mpc^{-1}), the apparent size of NGC 7331 permits mapping with moderate spatial resolution ($5'' \equiv 490$ pc) over a region many kiloparsecs in size. The *JHK* surface photometry presented in this *Letter* shows striking spatial variations in color. These results provide important clues to the distribution of interstellar matter and the nature of the extinction law in NGC 7331, and they permit an estimate to be made of the total mass of associated interstellar molecules.

II. OBSERVATIONS

The data presented here were obtained during 1981 August at the f/35 Cassegrain focus of the 3.8 m United Kingdom Infrared Telescope at Mauna Kea, Hawaii. A dual indium antimonide detector system was constructed to permit photometry through the same focal plane aperture simultaneously at two different wavelengths; in this experiment, the wavelengths used were K(2.20 μm) and either J(1.25 μm) or H(1.65 μm). Because the most significant conclusions of this *Letter* depend upon the analysis of color variations across NGC 7331, the capability to observe simultaneously in the two passbands provides an important advantage in eliminating systematic uncertainties. The beam profiles for the three passbands were nearly identical and had full widths at half-maximum of 5''.1.

Measurements at J and K were made along north-south and east-west lines passing through the optical nucleus; the separation between points along a line was 4''.8, or approximately one beam width. Measurements at H and K were made at the same north-south positions and at 19'' W. Since the major axis of NGC 7331 is oriented at a position angle of 167° (Bosma 1981), the north-south and east-west scans are displaced slightly in orientation from the major and minor axes.

Standard infrared beam switching techniques were used for the observations. The direction of separation between the signal and reference beam positions was east-west with a spacing of 50'' for the north-south and 90'' for the east-west scans. The scan south from the

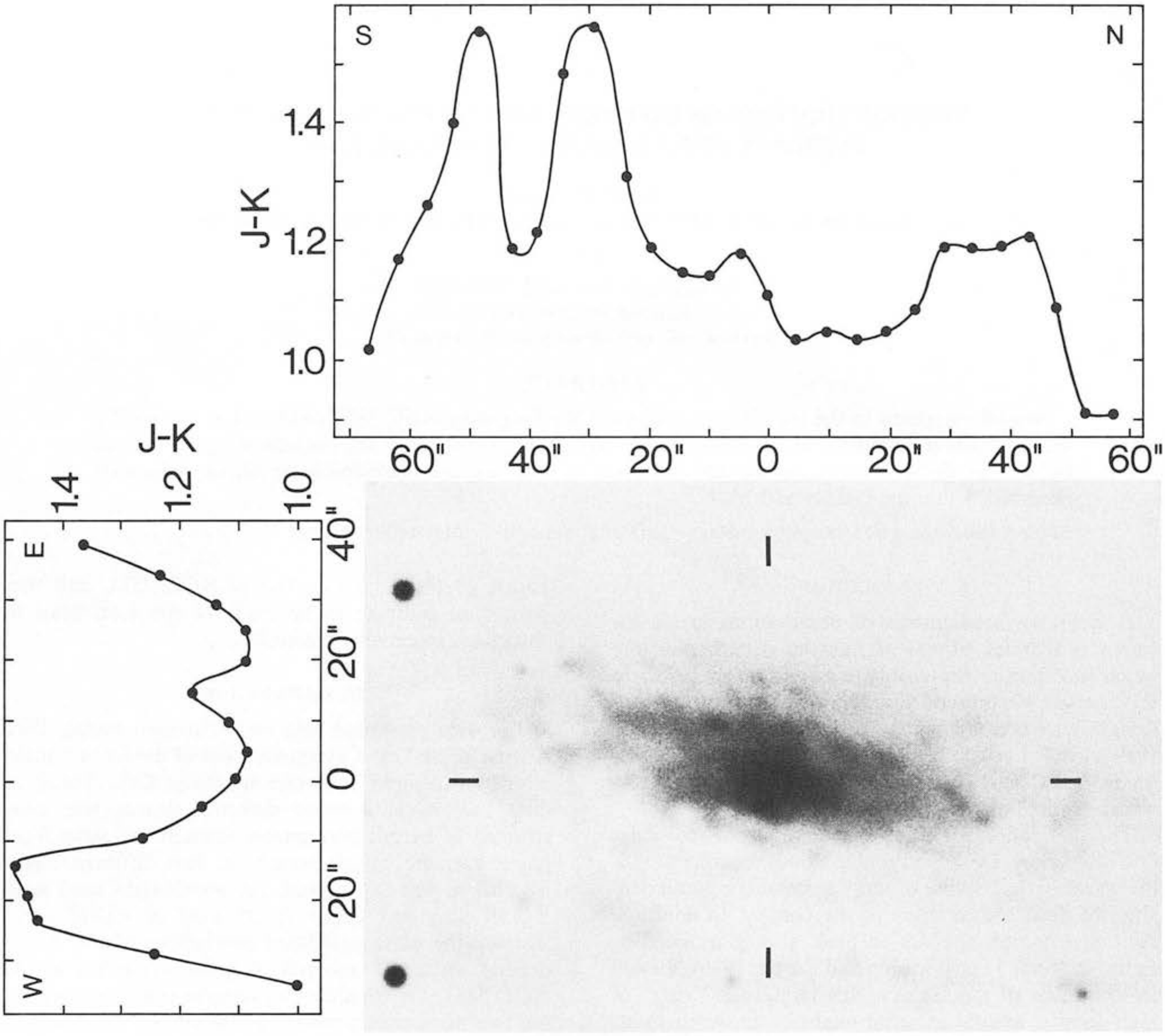


FIG. 1.—Comparison of the $J-K$ scans of NGC 7331 with a photograph of the galaxy in blue light. The scans intersected the nucleus, which is indicated by the tick marks, along lines in the north-south and east-west directions and parallel with the respective abscissae. Angular displacements along the scans are shown relative to the nucleus. The beam diameter was 5."1, approximately the separation between adjacent data points. Photograph by Dr. J. Rose.

nucleus was repeated with a beam separation of 75″ in order to demonstrate that structure evident in the scan was not due in part to the presence of field stars in the reference beam. The beam separations were chosen to ensure that negligible flux from NGC 7331 itself was present at the reference beam positions. We guided continuously on the optical nucleus with the use of an optical/infrared beam splitter and a television camera attached to a precision offset stage. The photometric standard was the star 4 Lac, for which we assumed $J = 4.33$, $H = 4.25$, and $K = 4.24$.

We observed $J-H$ and $H-K$ colors for the nucleus of NGC 7331 which are approximately 5% larger than those obtained by Aaronson (1977) using a 24″ aperture; this is sufficiently close agreement to warrant our direct use of his body of galaxy data in the comparison considered in the next section.

III. RESULTS

The results of our surface photometry of NGC 7331 along north-south and east-west lines are presented in Table 1 as the magnitude K and colors $J-K$ and $H-K$ at each position. The uncertainties are discussed in the footnotes to the table. In Figure 1 the observed $J-K$ color along the lines through the nucleus is

TABLE 1

INFRARED PHOTOMETRY[a] OF NGC 7331

Position[b]	K	$J-K$	$H-K$
Nucleus	9.23	1.11	0.31
4.8 N	10.14	1.03	0.25
9.5	10.90	1.05	0.24
14.3	11.45	1.03	0.27
19.1	11.72	1.05	0.25
23.8	11.96	1.08	0.28
28.6	12.16	1.19	0.35
33.3	12.45	1.18	0.32
38.1	12.71	1.19	0.38
42.9	13.00	1.21	0.39
47.6	13.36	1.08	0.25
52.4	13.77	0.91	0.29
57.1	14.09	0.91	0.22
4.8 S.......	10.26	1.18	0.30
9.5	10.98	1.14	0.32
14.3	11.45	1.15	0.31
19.1	11.80	1.19	0.33
23.8	12.03	1.31	0.39
28.6	12.27	1.57	0.49
33.3	12.60	1.48	0.53
38.1	12.85	1.22	0.43
42.9	13.11	1.18	0.42
47.6	13.56	1.56	0.43
52.4	13.89	1.40	0.45
57.1	13.93	1.26	0.47
61.9	14.08	1.07	0.21
66.7	14.02	1.02	0.19
4.8 E	10.84	1.08	...
9.5	11.55	1.12	...
14.3	11.98	1.18	...
19.1	12.72	1.08	...
23.8	13.29	1.08	...
28.6	13.82	1.14	...
33.3	14.00	1.23	...
38.1	14.47	1.37	...
4.8 W......	10.48	1.17	...
9.5	11.46	1.27	...
14.3	12.14	1.48	...
19.1	12.67	1.46	...
23.8	13.18	1.44	...
28.6	13.62	1.24	...
33.3	14.31	1.00	...
38.1	14.84	1.09	...

[a] In magnitudes; for all data at J and K and for H data within 30″ north and south and 20″ east and west of the nucleus, the internal accuracy is ±4%. For other points, the accuracy at H is ±8%.

[b] Displacement in arc seconds from the nucleus.

compared directly with a short-exposure photograph of the galaxy in blue light. This photograph emphasizes the bright central region of this highly inclined galaxy and shows the relationship of the inner part of the dusty disk and the bulge region and nucleus; the reader is referred to Sandage (1961) for a better indication of the overall appearance of the bulge and disk. In Figure 2, the brightness profiles at J and K measured southward from the nucleus are compared with the $J-K$ color observed at the same positions.

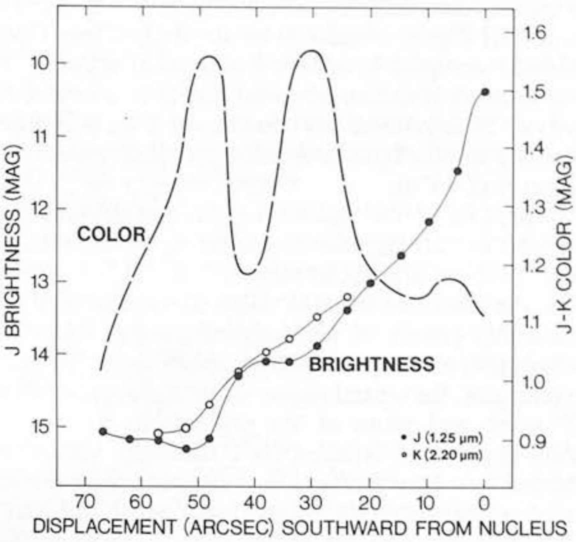


FIG. 2.—Comparison of the distributions of infrared color (*dashed line*) and relative brightness (*solid lines*) south of the nucleus. The ordinate for the $J-K$ color is at the right, that for the J flux is at the left. The flux at K has been plotted offset by 1.11 mag so that $K=J$ at the nucleus.

Figure 3 is a two-color diagram in $J-H$ and $H-K$. Results are plotted for the position of the nucleus (n), points within 20″ to the north (*open circles*) and south (*filled circles*), and three positions coincident with

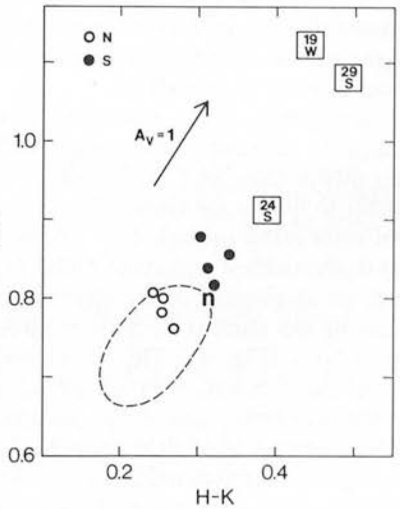


FIG. 3.—The relationship of $J-H$ and $H-K$ for NGC 7331. Open and filled circles represent colors for positions within 20″ north and south, respectively, of the nucleus (n). Open squares apply to locations therein designated in arc seconds relative to the nucleus. The region within the dashed curve is that occupied by most elliptical and spiral galaxies, as determined from multi-aperture photometry by Aaronson (1977). The extinction vector is represented by an arrow and indicates the effect on the infrared colors of the equivalent of a magnitude of visual extinction (Aaronson 1977; Frogel *et al.* 1978; Jones and Hyland 1980).

prominent dust lanes (24″ S, 29″ S, and 19″ W). The portion of Figure 3 enclosed by the dashed line is that normally occupied by elliptical and spiral galaxies; the near-infrared radiation from the bulges and central regions of these galaxies has been shown to be dominated generally by the light from red giants (Aaronson 1977; Frogel *et al.* 1978).

Inspection of the figures suggests immediately that large effects attributable to extinction by dust within NGC 7331 are present, as follows:

1. The observed infrared color is strongly and systematically redder at positions where dust lanes are conspicuous in an optical photograph (Fig. 1). At visible wavelengths, the central region extending approximately 20″ north and south of the nucleus has the smooth bright appearance expected for a relatively unobscured distribution of bulge stars, in accordance with the infrared colors observed there. Near 25″ north and south, $J - K$ becomes abruptly redder where darker material becomes apparent in the photograph. Likewise, on the west side, which is the side nearer to us (de Vaucouleurs 1958), the color reddens steeply where the optical extinction to bulge and disk stars increases sharply. $J - K$ also becomes redder in the dusty region to the east, although at larger projected radii than to the west; we discuss this asymmetry below.

2. The infrared fluxes at the shortest wavelength (J) are fainter at positions where the infrared colors are redder (Fig. 2). This effect is especially apparent at 30″ and 50″ south of the nucleus. Comparison of the southern J distribution with the bluer, northern J distribution (Table 1) shows that the flux to the north is systematically stronger and smoother than that to the south. The east-west scan shows a similar effect.

3. The locus of points in the two-color plot lies along the direction of the extinction vector (Fig. 3). This result suggests that dust within NGC 7331 follows an extinction law similar to that in the Galaxy. The magnitude of the observed color effect implies $A_v \approx 3$ (corresponding to $A_k \approx 0.3$) in the reddest regions of NGC 7331.

4. There is an asymmetry in the observed colors on opposite sides of the nucleus in both north-south and east-west directions (Fig. 1). This is at least partly attributable to the orientation at which the galaxy is viewed. On the near (west) side of the nucleus, most of the bulge stars along a line of sight are behind the dusty disk, and the light is therefore reddened. On the far side, most of the bulge stars along a line of sight are in front of the dusty disk, and the light is predominantly unreddened; at larger radii the stars in the dusty disk become relatively more important, accounting for the more gradual reddening of colors with increasing radius to the east. For these same reasons, the slight displacement of the north-south scan from the major axis can account for the average higher reddening to the south. Patchiness in the distribution of obscuring matter may also account for some of the asymmetry, especially in the north-south scan, and may result in, for example, the blue and red features seen at 40″ south and 15″ east. More importantly, although the above considerations imply that interstellar extinction is the primary mechanism responsible for the reddening, color effects due to spatial variation in the stellar population may also be present.

IV. DISCUSSION

The present observations of infrared extinction in NGC 7331 show straightforwardly that large amounts of dust are present in that galaxy. The result that the visual extinction achieves values well in excess of 1 mag satisfies the major goal of this experiment and demonstrates the utility of infrared surface photometry in the study of the interstellar medium in galaxies. It is natural to expect that surface photometry will be a superior method by which to study radial color gradients in galaxies. This is demonstrated by a comparison of the present results with the multiaperture study by Aaronson (1977); the measured difference in $J - K$ color for apertures of 24″ and 107″ centered on the nucleus of NGC 7331 is only 0.07 mag. This result is consistent with the present observations and emphasizes that even large radial color variations may not be readily evident in multiaperture work.

The simplest interpretation suggested by the data is that the bulge is encircled by a toroid of dust located in the plane of the galaxy. The inclination of the galaxy to the line of sight ($i = 15°$; Bosma 1981) produces the observed color asymmetry on opposite sides of the nucleus; the light from the bulge is reddened when seen through the disk. The mean radius of this toroid is approximately 4 kpc with a radial thickness of 2 kpc.

It is more realistic to interpret this large-scale distribution of extinction as being associated with the spiral structure seen in photographs of the galaxy; NGC 7331 shows tightly wound but well-defined spiral arms beginning immediately outside the bulge. An examination of other, more face-on, Sb galaxies strengthens the impression that the extinction toroid in NGC 7331 is associated with the spiral arms. In particular, the H I map of M81 (Rots 1975; Rots and Shane 1975) shows two spiral arms 2 kpc thick intersecting the major axis at a radius of about 4 kpc.

We can estimate column densities in NGC 7331 by noting from Figure 3 that the reddest regions correspond to $A_v \approx 3$. The relationship between visual extinction and the total hydrogen atom column density $N(\mathrm{H}) = N(\mathrm{H\ I}) + 2N(\mathrm{H_2})$ is taken to be $N(\mathrm{H})/A_v = 1.9 \times 10^{21}\ \mathrm{cm^{-2}\ mag^{-1}}$ (Bohlin, Savage, and Drake 1978), where we have assumed $A_v/E(B - V) = 3.1$. Dickman (1978) concludes that this relationship applies to molecular clouds with $A_v \lesssim 5$ and possibly to those with A_v as high as 10. This relationship implies for NGC 7331 that

$N(\mathrm{H}) = 1.5 \times 10^{21}$ cm^{-2} perpendicular to the galactic plane, which corresponds to a hydrogen surface density $\sigma(\mathrm{H}) = 11\ M_{\odot}$ pc^{-2} in the toroid.

Estimates of the surface density of interstellar gas are available for several galaxies, including our own and M81. An azimuthal average of the radial distribution of H_2 (Gordon and Burton 1976) in the Milky Way implies that much of the molecular material is located at radii between 4 and 8 kpc, and that the surface densities are 4–11 $M_{\odot}$ pc^{-2}. These values are comparable to those derived for the toroid in NGC 7331. Similar studies of H I in M81 imply values of 3 $M_{\odot}$ pc^{-2} for atomic hydrogen near a radius of 5 kpc (Rots and Shane 1975). If we assume that $\sigma(H_2) \approx 4\ \sigma(\mathrm{H\ I})$, as at radii of 4–8 kpc in the Milky Way (Gordon and Burton 1976), then the derived surface densities of molecular hydrogen in M81 and NGC 7331 are comparable.

Assuming that the inner and outer radii of the toroid are 25″ and 50″, we estimate a total gas mass of 6×10^8 $M_{\odot}$. The total mass within the central 120″ of NGC 7331, derived from the rotation curve in Rubin *et al.* (1965), assuming Keplerian motion, is 8×10^{10} $M_{\odot}$. Thus the mass of gas inferred from the extinction constitutes about 1% of the total mass in the central 12 kpc of NGC 7331. Strictly speaking, this estimate is a lower limit since some regions of the interstellar medium, in particular the cores of molecular clouds, are opaque even at infrared wavelengths.

This experiment demonstrates the feasibility of near-infrared studies of the interstellar material in external galaxies. With existing technology, this method compares favorably with molecular line observations in terms of practicality and quality of results. With the advent of large arrays of infrared detectors, we can expect great progress in this area.

We are grateful to Dr. Jim Rose for providing the photograph of NGC 7331. We benefited from comments from Drs. Jay Frogel, Tim Hawarden, and Rich Isaacman. This research was supported by NASA contract NASW 3159, the National Research Council, and the Royal Observatory, Edinburgh.

REFERENCES

Aaronson, M. 1977, Ph.D. thesis, Harvard University.

———. 1979, in *Photometry, Kinematics, and Dynamics of Galaxies*, ed. D. S. Evans (Austin: University of Texas Press), p. 147.

Bohlin, R. C., Savage, B. D., and Drake, J. F. 1978, *Ap. J.*, **224**, 132.

Bosma, A. 1981, *A.J.*, **86**, 1971.

de Vaucouleurs, G. 1958, *Ap. J.*, **127**, 487.

Dickman, R. L. 1978, *Ap. J. Suppl.*, **37**, 407.

Frogel, J. A., Persson, S. E., Aaronson, M., and Matthews, K. 1978, *Ap. J.*, **220**, 75.

Gordon, M. A., and Burton, W. B. 1976, *Ap. J.*, **208**, 346.

Holmberg, E. 1975, in *Stars and Stellar Systems*, Vol. **9**, ed. A. Sandage, M. Sandage, and J. Kristian (Chicago: University of Chicago Press), p. 123.

Jones, T. J., and Hyland, A. R. 1980, *M.N.R.A.S.*, **192**, 359.

Rots, A. H. 1975, *Astr. Ap.*, **45**, 43.

Rots, A. H., and Shane, W. W. 1975, *Astr. Ap.*, **45**, 25.

Rubin, V. C., Burbidge, E. M., Burbidge, G. R., Crampin, D. J., and Pendergast, K. H. 1965, *Ap. J.*, **141**, 759.

Sandage, A. 1961, *The Hubble Atlas of Galaxies* (Washington, D.C.: Carnegie Institution), p. 19.

Sandage, A., and Tammann, G. A. 1981, *A Revised Shapley-Ames Catalogue of Bright Galaxies* (Washington, D.C.: Carnegie Institution).

Simkin, S. M. 1976, *A.J.*, **72**, 1032.

Note added in proof.—After the observations and analyses reported in this *Letter* were completed, Young and Scoville (1982, *Ap. J.* [*Letters*], **260**, L41) detected CO outside the nucleus in NCG 7331 and provided direct evidence for the molecular ring discussed here. They have inferred a mean ring radius of ~ 45″ and a peak hydrogen surface density of ~ 25 $M_{\odot}$ pc^{-2}, comparable to our values.

IAN GATLEY and J. M. STEWART: United Kingdom Infrared Telescope Unit, 900 Leilani Street, Hilo, HI 96720

CHARLES M. TELESCO: Space Science Division, NASA Ames Research Center, Mail Stop 245-6, Moffett Field, CA 94035