A detailed study of the reflection nebula, NGC 7023

Philip Joseph Ruaraigh McDonald Moore

A DETAILED STUDY OF

THE REFLECTION NEBULA, NGC 7023

by

PHILIP JOSEPH RUARAIGH McDONALD MOORE

A thesis submitted to the University of Durham
for the Degree of Doctor of Philosophy

October, 1982

The copyright of this thesis rests with the author. No quotation from it should be published without his prior written consent and information derived from it should be acknowledged.

ABSTRACT

Polarisation and intensity maps in three broad wavebands are presented for the reflection nebula NGC7023. The data are used to investigate the structure, dust distribution and grain characteristics of the material surrounding the central illuminating star HD200775 of the reflection nebula.

Calculations have been made, using a Monte-Carlo technique, for various parameters representing the structure and content of the nebula to predict and explain the observed measurements.

The successful description of the observations puts severe restrictions on the nebular parameters. It is found that the geometry of the nebula is in the form of an extended cloud with a foreground conical cavity in which the illuminating star is situated. The dust grains are required to have a power law size distribution of the form

$$n(a) = a^{-4.05}$$

and grain material corresponding to silicates is most likely although ice cannot be excluded.

CONTENTS

Page

ABSTRACT

CONTENTS

TABLES

ILLUSTRATIONS

PREFACE

CHAPTER ONE : A GENERAL INTRODUCTION TO REFLECTION NEBULAE 1

1.1 The Varieties of Diffuse Galactic Nebulae 1

1.2 The Forms of Reflection Nebulae 2

1.3 The Importance of Reflection Nebulae 5

1.3.1 The Association with Early-type stars 5

1.3.2 Interstellar Dust 8

1.4 A Brief Historical and Technical Review 11

1.4.1 Photometry and Colorimetry 12

1.4.2 Optical Polarisation 13

1.4.3 Ultra-violet Studies 17

1.4.4 Infra-red and Radio Observations 18

CHAPTER TWO : A REVIEW OF OBSERVATIONS OF NGC 7023 20

2.1 Introduction 20

2.2 The Star HD 200775 20

2.3 The Nebula NGC 7023 26

2.4 The Optical Properties 33

2.5 Towards a Structural Understanding 39

Page

CHAPTER THREE : THE NEW OBSERVATIONS 45

3.1 The Observational Details 45

3.2 The New Data 45

3.3 Detailed Aspects of the New Observations 51

3.4 Consideration of Defects 53

3.5 Examination of the Maps 56

3.6 Comparison with Other Observations 62

3.7 Conclusions 68

CHAPTER FOUR : DUST 70

4.1 Introduction 70

4.2 Cosmic Abundances 71

4.3 Extinction 73

4.3.1 The Infra-red Regime 75

4.3.2 The Visible Region 78

4.3.3 The Ultraviolet 78

4.3.4 The Variations in Extinction 80

4.4 The Interstellar Polarisation 81

4.5 Formation, Modification and Destruction of grains. 83

4.5.1 Formation 84

4.5.2 Modification and Destruction 86

4.5.3 Mantles 88

4.6 Dust Composition : A Summary 90

4.6.1 Ice 91

4.6.2 Graphite 91

4.6.3 Silicates 92

4.6.4 Iron 92

4.6.5 Organics 93

Page

4.7 Grain Size Distributions 93

4.7.1 The OHG Function 95

4.7.2 Shattering 98

4.7.3 Empirical Determination of the Power Law Distribution 101

4.8 Mie Theory 103

4.8.1 Implementation 103

4.8.2 The Single Scattering Results 106

CHAPTER FIVE : MONTE CARLO SIMULATIONS OF NGC 7023 118

5.1 Introduction to Nebular Simulation 118

5.2 The Implementation 122

5.2.1 The General Problem 122

5.2.2 The Simulation for an Embedded Source 126

5.2.3 The Improvements to the Simulation 128

5.3 Application to NGC 7023 129

5.3.1 Basic Modelling Considerations 129

5.3.2 The Plane Slab : A Starting Point 131

5.3.3 Convex Surface : Embedded Source 136

5.3.4 Concave Surface : Exterior Source 144

5.4 Conclusions 159

REFERENCES 161

ACKNOWLEDGEMENTS 171

TABLES

		Page
1	Spectral Class versus Nebulae Expectation	6
2.1	Characteristics of HD200775	22
2.2	Surface Brightness Observations of NGC7023	34
3	Polarisation Variations with Offset Distance and Wavelength for NGC7023	69
4.1	Cosmic Abundances : Common Elements	72
4.2	Restrictions on Grain Composition	73
5.1	Conical Cavity Model Characteristics	157

ILLUSTRATIONS

		Page
Plate 1.1	Short Exposure of NGC7023	28
Plate 1.2	Blow-up of Plate 1.1	29
Plate 1.3	Medium Exposure of NGC7023	30
Plate 1.4	Long Exposure of NGC7023	31
Fig. 2.1	Complete Polarisation Map (VV-filter)	46
Fig. 2.2	Half Polarisation Map (R-filter)	47
Fig. 2.3	Half Polarisation Map (B-filter)	48
Plate 2.4	Complete Intensity Map (VV-filter)	49
Fig. 2.5	Complete Polarised Intensity Map (VV-filter)	52
Fig. 2.6	Log(I) v. Distance, central scan, all Wavelengths	63
Fig. 2.7	Polarisation v. Distance, central scan, Blue	64
Fig. 2.8	Polarisation v. Distance, central scan, Visual	65
Fig. 2.9	Polarisation v. Distance, central scan, Red	66
Fig. 4.1	Interstellar Extinction v. Wavelength	74
Fig. 4.2	Mie Calculations : Silicates, Intensity variations with Angle	108
Fig. 4.3	Mie Calculations : Ice, Intensity variations with Angle	109
Fig. 4.4	Mie Calculations : Silicates, Polarisation variations with Angle	110
Fig. 4.5	Mie Calculations : Ice, Polarisation variations with Angle	111
Fig. 4.6	Mie Calculations : Graphite, Polarisation variations with Angle	112
Fig. 4.7	Mie Calculations : Iron, Polarisation variations with Angle	113

		Page
Fig. 4.8	Mie Calculations : Organic Materials Polarisation variations with Angle	114
Fig. 4.9	Mie Calculations : Silicates, Oort-van de Hulst - Greenberg size distribution for Wavelength and Angle variations	115
Fig. 5.1	Scattering along a Line of Sight Column	123
Fig. 5.2	Scattering in a Parabolic Cavity	123
Fig. 5.3a	Plane Parallel Slab, Normal Incidence	133
Fig. 5.3b	Tilted Slab	133
Fig. 5.4a	Polarisation variations with Distance for Fig. 5.3a	135
Fig. 5.4b	Polarisation variations with Distance for Fig. 5.3b	135
Fig. 5.5	Embedded Source, Spherical Structure	137
Fig. 5.6	Embedded Source, Ovoid Structure	137
Fig. 5.7	Ovoid Geometry : Polarisation variations for all materials (North).	140
Fig. 5.8	Ovoid Geometry : Polarisation variations for Silicates.	142
Fig. 5.9	Ovoid Geometry : Brightness variations for Silicates	143
Fig. 5.10	Hemispherical Surface Geometry	145
Fig. 5.11a	Semi-Infinite Nebula, Plane Surface : External Star	145
Fig. 5.11b	Convex Boundary : External Star	146
Fig. 5.11c	Concave Boundary : External Star	146
Fig. 5.12	Conical Cavity Geometrical Details	148
Fig. 5.13	Conical Geometry : Polarisation variations for all materials (South)	150
Fig. 5.14	Conical Geometry : Intensity variations for all materials (South)	151
Fig. 5.15	Conical Geometry : Polarisation variations for Silicates and Ice	152
Fig. 5.16	Conical Geometry : Intensity variations for Silicates and Ice	153

PREFACE

The work described in this thesis was carried out in the period 1979 to 1982 while the author was a research student of the Physics Department of the University of Durham, under the supervision of Dr. S.M. Scarrott.

The computer reduction techniques used to transform the observations for this thesis were developed by the polarimetry group under Dr. Scarrott. The original Monte Carlo programs were the work of Dr. R.F. Warren-Smith. During the author's studentship at Durham he was frequently engaged in the use of the polarimeter abroad, at a number of different observatories. In addition he has processed and interpreted the new data on NGC 7023 and extended the scope of the Monte Carlo simulations to include a much larger range of possible geometries. Several fundamentally different structures have been investigated in the process of modelling NGC 7023. Many structures proposed previously have been shown to be incorrect. A new model which explains the polarisation and intensity data is presented.

CHAPTER ONE

When the lamp is shattered
The light in the dust lies dead -

Shelley

A GENERAL INTRODUCTION TO REFLECTION NEBULAE

1.1 The Varieties of Diffuse Galactic Nebulae

Reflection nebulae are a category within the more general group of diffuse galactic nebulae - luminous concentrations of gas and dust normally associated with illuminating and exciting stars. The distinction between reflection nebulae and the other varieties - HII regions, dark nebulae - depends on the nature of their illumination and excitation.

There are two mechanisms which can cause diffuse galactic nebulae to be visible: first the gas may be intrinsically luminous, as a consequence of ionisation, by hot nearby stars, secondly the dust may scatter light from associated stars in the immediate neighbourhood, or from more distant sources, e.g. the galactic nucleus.

In order to produce photo-ionisational luminosity of the local gas a star must be of at least spectral type BO (Kaplan and Pikelner, 1970), in which case the nebulae become emissive, or gaseous, nebulae. In the alternative if the illuminating star is of type B2, or cooler, then the resulting structure is a reflection nebula. Stars of type BO and B1 can produce intermediary forms between emission and reflection nebulae; these do in fact compose the majority of bright nebulae, but emission nebulae do still possess a component of scattered light.

Spectroscopically the two varieties may be distinguished comparatively easily. A reflection nebula must necessarily reproduce the spectrum of its illuminating stars; for this reason, one should expect a continuous spectrum overlaid by a series of absorption or emission features. An ionised gas would of course, produce a line spectrum, characteristic of the material composition of the bright matter. Thus there is an important distinction to be made here, for while spectra from ionised bright nebulae indicate the chemical make up of the gas, those from reflection nebulae do not. A further consequence of the different sources of illumination, is to be found in the temperatures of the two forms of nebulae. Temperatures in reflection nebulae are of the order of several tens of degrees Kelvin, while those in an HII region are typically several thousand Kelvin.

1.2 The Forms of Reflection Nebulae

Several hundred reflection nebulae have been identified. These are tabulated in the catalogues of Cederblad (1946), and Dorschner and Gurtler (1966), who searched through the Palomar Sky Survey for objects equally bright on both plates or brighter in the blue, and also for red nebulae associated with stars of spectral type G or later. Van den Bergh (1966) examined all BD and CD stars north of declination -33^{o} in which reflection nebulosity appears on both the blue and red plates of the Palomar Sky Survey. Finally Racine (1966) extended van den Bergh's catalogue through extensive photometric and spectroscopic observations of the listed stars.

Four important points emerge. First reflection nebulae are not a uniform category of objects. Secondly they are

characteristically bluer than their illuminating stars and thirdly they are extremely frequently associated with stars of spectral type B. Fourthly they are not uniformly distributed through the Galaxy. These last two points taken together have important consequences for the theory of star formation and will be examined further below.

Reflection nebulae may be divided into three sub-groups: (1) compact objects where the associated dust and gas may be intimately connected with the illuminating stars, (2) generally larger objects where the dusty matter is almost certainly part of an extensive molecular cloud and (3) objects found at high galactic latitude which have no evident illuminating stars. To dispose of this last group first: van den Bergh (1966) observed that the vast majority of clouds at intermediate and high galactic latitudes were located south of the galactic plane, with a strong concentration towards the galactic centre. The exceptions being certain red streamers, which must be shining by emission, and also some blue clouds at intermediary latitudes, which were reflecting light from the Lacerta association. K.A. Innanen (1969) concluded on further investigation that most of these objects were shining by reflection of the integrated light of the galactic nucleus.

The dark cloud reflection nebulae have been known since the second decade of this century. Their surface brightnesses, in visible light, fall between 20 and 23 magnitudes per square arcsecond and they are clearly intensity bounded rather than density limited objects. This has been known since the publication by Hubble (1922) of his seminal papers on

galactic nebulae. Hubble derived from the inverse square law a simple logarithmic relationship between the photographic magnitude of the illuminating star and the angular extent of the nebula:

$$m_* + 4.90 \log(a) = 11.02 \pm 0.10 \qquad 1.1$$

m_* and a are the stellar magnitude and nebular extent respectively, while the constants are empirically determined. Visually such "molecular cloud" reflection nebulae are frequently characterised by wispy filaments. Particularly fine examples of this structure are to be found in the Pleiades association, e.g. Merope or Maia.

Finally there are the compact nebulae. They can often have a very striking geometrical construction, note the bipolar and biconical nebulae, or they may appear turbulent and amorphous as does NGC 7023. However the definitive characteristic is localised nebulosity, usually surrounding a single star, which may itself be heavily obscured. The evolutionary status of these objects is in no way clear. Some have been proposed as the site of recent or continuing star formation. Others as the last stages in the evolution of a star. A notable example of the former is Hubble's variable nebula, NGC 2261, illuminated by R Mon, while the latter is exemplified by η Carinae which has recently been proposed as a likely candidate to supernova (Walborn and Gull, 1982) or alternatively as a Hubble-Sandage variable (Warren-Smith et al. 1979). There seems no compelling reason to presume that objects of so markedly different morphological types should in fact be considered as

representatives of a single evolutionary condition.

1.3 The Importance of Reflection Nebulae

If one were to set aside an inquiry into reflection nebulae purely in terms of their own intrinsic interest, there are two other branches of astronomy which would compel investigators to examine reflection nebulae. These are the study of the interstellar medium and the evolution of stars. Although it has been suggested that reflection nebulae are associated both with the birth and demise of stars the following summary shall concentrate only on the former.

1.3.1 The Association with Early-type Stars

It has been known from the early forties (Baade, 1944) that there is a strong connection between young early-type stars and interstellar clouds; a relationship that is particularly acute for extreme Population I stars. However, that there might be a direct generative link between such stars and reflection nebulae was heavily disputed. At the time of the publication of Cederblad's catalogue (1946) there was no known spatially significant connection between the distribution of diffuse nebulae and that of stars of any stellar type. A partial explanation for this lies in the inclusion by Cederblad of all the bright diffuse nebulae known to him. As a result of which his sample was contaminated by various old objects, for example, planetary nebulae, and also by the dispersion introduced by the low accuracy of the distances he assumed (Racine, 1966).

Far from suspecting such a link, theorists had gone a considerable way towards dismissing any such relationship as purely fortuitous. V.A. Ambartsumyan and Sh. G.

Gondeladse (1938) noted that while approximately one in three O stars excite luminosity in visible emission nebulae only ten percent of all stars in classes B2 to B8 inclusive are associated with reflection nebulae. On the assumption that nebulae could be regarded as small randomly distributed clouds at various distances from stars, which were themselves uniformly distributed in space, and neglecting the effects of absorption, Ambartsumyan produced the following table:

Table 1. Spectral Class versus Nebulae Expectation

Spectral Class:	BO	B1-B9	A	F	G	K	M
Expectation Parameter	0.2	2.2	0.6	0.5	0.1	0.45	0.05
Actual number of nebulae	23	122	28	6	1	5	2

From the close correlation between the numbers of observed reflection nebulae and his expectation parameter, Ambartsumyan was entirely correct to deduce that any association was accidental. BO stars could produce emission nebulae thus depleting their quote of reflection nebulae.

However, as more data accumulated the picture began to change. First it became clear that absorption was a very significant factor, then that nebulae frequently occur in groups, both of which observations completely undermine the analysis above.

Next came the discovery by Roberts (1957) that the B stars found in clusters and associations account for nearly all stars of this type in the solar neighbourhood; a conclusion that was extended by Ebert (1968) to nearly all early-type mainsequence stars.

Indisputable evidence for the grouping of reflection nebulae was provided by van den Bergh (1966) who noted thirteen distinct groups of reflection nebulae, and remarked further that many of these were concentrated along Gould's Belt. Furthermore the limits of such clusters were often to be found within the boundaries of recognised OB associations. This work was extended by Racine (1966) who was able to demonstrate that in only one instance was the observed connection, by van den Bergh, a result of chance association.

More recently still theories of star formation have tended towards natural mechanisms from which reflection nebulae might be expected to develop. Larson (1969a) constructed a hydrodynamical model of pre-main sequence evolution. According to this model an unstable proto-stellar cloud begins to collapse gravitationally. This increases the density in the central regions, which produces a proto-stellar core surrounded by a diffuse shell of gas and dust. Since the free-fall time varies inversely with the square root of the density, matter collapses more quickly the nearer it is to the core. The observational consequence of Larson's theory is that stars of spectral type B5 to B8 will only be found with significant circumstellar shells during their pre-main sequence lifetime.

As a natural consequence of this prediction attempts have been made to re-examine Herbig's Be and Ae stars (emission stars associated with luminosity) both by optical spectrophotometry (Strom et al., 1971) and also in the millimetre waveband (Loren et al., 1973). Loren and his co-workers found carbon monoxide emission from the nebulosity associated

with nearly all stars surveyed, while Strom et al. concluded that the majority of Herbig's stars have surface gravities appropriate to pre-main sequence stars the remainder being characteristic of zero-age main sequence. Convinced that these were some of the earliest known stars, Strom et al. proceeded to examine the regions surrounding these stars in the one micron band to test the following hypothesis. Are these stars isolated, forming from relatively small clouds or are they the first visible representatives of stellar groups? If the first proposition is correct then a search for faint red objects will prove fruitless. The contrary is true for the second hypothesis. Preliminary investigations by Strom et al. (1971) in a number of such clouds have indicated the presence of a number of red objects, which appears to confirm the second alternative.

Lastly both FU Ori and Lk H-α-190, which have infra-red excesses, (indicative of hot dust), after brightening considerably in the optical region now both illuminate small reflection nebulae (Martin, 1978). To summarise briefly: the casual relationship between reflection nebulae and the pre- and post-main sequence history of stars can no longer be doubted. An investigation of one demands a study of the other.

1.3.2 Interstellar Dust

A more detailed account of the properties of interstellar grains must wait until chapter four, nonetheless some mention of the nature of the granular properties of these particles is required at this stage. Interstellar dust forms only a minute component of all interstellar matter and accurate

estimates of the proportionality are difficult. However a broad consensus has emerged that the ratio of gas to dust is approximately 100 : 1. This figure was originally derived from the ratio in the solar neighbourhood. Using this value Oort was able to deduce the mean density of the interstellar dust from stellar motions perpendicular to the galactic plane. He then measured the combined density of stars and gas in the solar neighbourhood to be 6×10^{-23} kg m^{-3}. Given a known stellar density of 3×10^{-23} kg m^{-3} he concluded that the average density of interstellar dust as 3×10^{-25} kg m^{-3}. This figure was substantially confirmed by van de Hulst (1968) who obtained a value of 1.4×10^{-25} kg m^{-3} from 21cm observations.

Given that grains are so finely distributed - on average there is only one grain in a hundred metre cube - it is perhaps surprising that they should have any significant effect whatsoever. Yet they are of considerable importance for, following Morton's (1974) measurements of interstellar extinction, Greenberg (1974) deduced an extinction due to dust of two magnitudes per kiloparsec in the plane of the Milky Way. The interstellar reddening is the most obvious way in which the dust betrays its presence. As starlight passes through dust clouds a certain amount of energy is removed from the incident light by the grains so that the intensity of the starlight is reduced. Furthermore the absorbed energy will be re-radiated as heat while the majority is re-emitted at the original wavelength almost instantaneously (scattering). The amount of extinction is dependent on the chemical composition of the scattering materials and is thus characteristic

of those materials. Broadly speaking: in metallic particles absorption is large compared with scattering, whereas in dielectrics extinction is principally by scattering. A secondary consequence of the scattering of starlight is the polarisation, both linear and circular, of the transmitted radiation. Together all these measurements have been used in numerous attempts to identify the granular components. It is now believed that the interstellar medium is mainly dielectric.

Another problem is that the relationship between grains found in the general interstellar medium and those in reflection nebulae is poorly understood. There are excellent reasons to surmise that there should be significant variations in the composition of dust particles from different regions. It is clearly obvious that the physical environment of the grains can be markedly different; their range includes dark clouds at temperature of between 5K and 20K, to HII regions, near bright young stars, with temperatures of the order of 10,000 K. This would seem to make reflection nebulae an extremely profitable area for study, for while illuminated by nearby stars they are not of so high a temperature as to disrupt the granular structure, being typically at about 30K. However, it must be emphasised that care is necessary before the generalisation of results applicable to reflection nebulae to the entirety of interstellar dust. We do not know that the origins of these two groups of particles are in fact the same, and in some cases we have good reason to believe otherwise. Particles in reflection nebulae may condense from matter ejected from

stars, while grains in the interstellar medium are more likely to result from a slow accretion.

Yet, recent work by Greenberg and Hong (1973) indicates that small refractory grains do not accrete significant mantles. They rest their conclusion on the well known observation of the bump at 2200 Å in the ultra-violet extinction curve. Calculations of the behaviour of this feature for small grains accreting mantles of 0.005 μm thickness convince them that no significant accretion takes place. Were it to do so then the extinction curve would turn over from an upward to a downward curve at 6 μm^{-1}. Since there is no evidence for this behaviour they conclude that save for evaporation during star formation, or when very close to stars, that small grains would remain substantially unchanged during their lifetimes. Despite this evidence, the position is further complicated in that it is generally believed (McDonnell, 1978) that larger particles, perhaps the ejecta of supernovae or in very dense protostellar nebulae, might be able to act as condensation sites for mantles.

1.4 A Brief Historical and Technical Review

On account of their prominence in the visible, and as a result of the discussion above, reflection nebulae as a class have been observed in considerable detail using a wide variety of techniques. Their existance was first recognised by Slipher in 1912 who noted that the spectrum of Merope was, "a true copy" of the brighter stars in the Pleiades, but except for the original identification of the class, optical spectroscopy has played a comparatively minor

role in the continuing study of reflection nebulae. That cannot be said of the other branches of optical astronomy.

The investigation of reflection nebulae has proceeded by studies of their surface brightnesses, colours and polarisation in visible light, and now continues to expand the range of techniques involved utilising measurements in the ultra-violet, infra-red and radio wavebands.

1.4.1 Photometry and Colorimetry

Early work on reflection nebulae was bedevilled by systematic errors in equipment and techniques. A survey of the literature gives the clear impression that although there was some idea of the experimental difficulties involved there was no clear realisation of the magnitude of the errors that were systematically introduced.

Spectroscopic observations of reflection nebulae by Struve and Yerkes in 1938 gave the general impression that these nebulae were bluer than their illuminating star, but only two years earlier Keenan had published data on NGC 7023 indicating that this nebula was more luminous in red than blue light. The position became even more confused with Schalen's observations of Merope in 1948 which showed that this nebula was also redder than the illuminating star. Thus in the early fifties all that could be said with certainty about these objects was that there was general agreement with Hubble's empirical relationship and that these nebulae were seen by reflection. Even this latter statement was to be partially challenged by Aller (1958) who suggested that some contribution to the high surface brightness of these objects might be the result of flourescence.

The position began to improve in the late fifties and early sixties when Madame Martel (1958) produced data on seven nebulae showing that all were slightly bluer than their illuminating stars. Substantial confirmation was provided by Vanysek and Svatos (1964), O'Dell (1966), Elvius and Hall (1966), Roark (1967), Dahn (1967), Gehrels (1967) and Zellner (1970,1973). These observers found reflection nebulae to be generally bluer than the illuminating stars, but that the nebulae reddened with the distance offset.

The question of flourescence took rather longer to resolve. It was not entirely dismissed until the publication of the photometric and spectrophotometric observations of Rush et al. (1975) who found no evidence for flourescence in examinations of five nebulae and concluded that at most it could only account for some ten percent of the surface brightness. Their deduction was that these objects had a high albedo. Attention was concentrated on a consideration of high albedo materials since most investigators were convinced that these nebulae were optically thin. This was partly as a consequence of successful attempts at modelling. In 1976 Fitzgerald et al. were arguing for high albedo strongly forward throwing grains to explain their surface brightness studies of the Chameleon nebula. More detailed theoretical simulations now indicate that a solution of the surface brightness problem lies in multiple scattering within the nebulae.

1.4.2. Optical Polarisation

Polarisation studies of reflection nebulae began remarkably early with the observations, in blue light, of

NGC 2261 - Hubble's variable nebula - by Meyer in 1920. He found values of up to ten per cent for large scale observations. The variation of results found in colorimetry were repeated for polarimetry. Later measurement by R.C. Hall (1965) at three wavelengths produced values between 10% and 25%. Similar results maintain for all well studied nebulae. Henyey's observations in 1938 found polarisation in the Pleiades to be very small or non-existant, while those of Elvius and Hall (1965) measured values up to 15% in Merope. Even comparatively modern observations can produce wild fluctuations. Gliese and Walter examining NGC 7023, in 1951, found the polarisation to range from 0% to 56%, while Elvius and Hall, and Gehrels and Zellner all determined a much restricted range of 5% to 20%. However they disagree on the values of polarisation at particular positions in the nebula, by as much as 50%.

The explanation for the considerable variability in results must involve a number of factors. In the very early days the inherent accuracy of equipment was lower; while more recently difficulty has arisen from the use of entirely different aperture sizes by different observers. Thus they were frequently not observing areas that were coincident Still the major demon in this pantheon, which is still a problem today, is the sky subtraction. It is to be noted that Gehrels, using the technique of alternatively panning the telescope on and off the object, made his sky subtraction from a point of 2.6' from the central star. This is much closer to the star than the areas used by Elvius and Hall, so much of the discrepancy in their results may

be accounted for by this "zero point" error.

Despite the problems outlined above, the early observers succeeded in identifying many of the distinguishing advantages of polarimetry. Elvius and Hall (1966), examining NGC 2068 were able to show quite conclusively that of the two prominent stars in the reflection nebula, only one was associated with the illumination of the nebula, in any significant manner. Their technique used the centro-symmetric pattern that is characteristic of polarised light from a single source illuminating dust by reflection. Very much the same procedure has been applied to the bipolar nebula S106 (Perkins, King and Scarrott, 1981) to identify a completely obscured optical source with a prominent infra-red source.

Other general polarisation features of importance are:

(i) the variation of the magnitude of the polarisation with distance from the illuminating sources,

(ii) the behaviour as a function of wave-length, and

(iii) the identification of aligned grains in magnetic fields and thereby the position and orientation of the magnetic fields themselves.

Taking these in order: many nebulae exhibit a general increase in polarisation with distance. This is easily explained in qualitive terms from an understanding of the functional dependence of the polarisation upon scattering angle. Some objects, at greater distances, exhibit a slow decline from a peak in the polarisation. The interpretation of this behaviour is rather more complex since it could result from a number of processes: depolarisation on trans-

mission through a dense medium, the addition of a component of unpolarised light or as a consequence of some particular geometry which limits the selection of scattering angles. The wavelength dependence has usually been found to resemble the interstellar polarisation - approximately a linear decline with inverse wavelength. This has been considered as supportive of van de Hulst's equilibrium condition's for grain production and destruction. For many years one of the puzzles in explaining polarisation (Gehrels, 1974) was the success of overly simplistic models requiring only a single scattering; this in complexes where the surface brightness indicated optical depths greater than unity. This is now understood in terms of 'strongly forward throwing' grains.

Magnetic fields were first identified by Elvius and Hall (1964) who explained anomalous polarisation vectors in terms of aligned non-spherical grains in the filamentary structures of NGC 2068. Comparatively small magnetic fields are required to induce sugh alignment. Fields of the order of 10^{-6} gauss are sufficient. It was once thought that the galactic magnetic field was the responsible agent, but it is now clear that in certain objects the field is intrinsic to the object itself. A fine example of this is to be found in the central regions of R. Mon (Gething et al., 1982) where it is believed that a toroidal magnetic field is associated with the central disc.

The discussion above has been devoted to linear polarisation. Circular polarisation too can be produced and has been detected, but the studies are still embryonic as a result of the minute signals involved. Serkowski (1973)

has found -0.1% circular polarisation in VY C.Ma., but it is not yet known whether it results from linear to circular conversion or from depolarisation.

In resume: polarisation studies give information on nebular structures, grain size and composition, the relative proportions of the different sizes and can give an insight into the presence and form of magnetic fields. To a lesser extent they are indicative of density (very high percentage polarisations are only to be expected in optically thin nebulae) and the origins of the granules.

1.4.3 Ultra-violet Studies

Ultra-violet studies of reflection nebulae are extremely scarce. It is only recently that satellites have become available for observations and before that the blocking effect of the atmosphere ruled out such studies. Only three general conclusions can be relied upon to date: the interstellar hump at 2200Å is also to be found in the dust surrounding reflection nebulae (Bless and Savage, 1972). This was confirmed from Orbiting Astronomical Observatory measurements due to Witt and Lillie (1978) though the prominence of the feature can vary. Secondly the scattering efficiency in the ultra-violet is rather higher than that in the visible, and thirdly the grains appear to scatter more isotropically i.e. are less forward throwing, and thus more Rayleigh-like than in the visible. The difficulties involved in ultra-violet work are illustrated by the declaration by Viotti (1976) that the observations from the S2/68 experiment on the TD1 satellite, made it unlikely that there was any significant nebular contribution from

NGC 7023, the cloud surrounding HD200775. Thus he ruled out considerations of such a contribution from his analysis of the stellar spectrum. However re-examination of his results, in the light of new optical photometry, led Witt and Cottrell (1980) to conclude that up to 40% of the U.V. intensity might have originated in scattering from the cloud!

Of particular relevance to the study of grain properties are the observations by Witt and Cottrell (1978) of the reflection nebulosity in Orion, which indicates the similarity between granules in the interstellar medium and those in reflection nebulae. Also Andriesse et al. (1977) found that ultra-violet photometry of Merope formed a consistent set with ground based optical work which generally confirmed the relative blueness of the nebula with respect to the illuminating source.

1.4.4 Infra-red and Radio Observations

The regions of, or surrounding, numerous reflection nebulae have been investigated in both the radio and the infra-red. However, it is obvious that in both wavebands that it is not the reflection nebulosity as such that is being probed, but quite distinct features of the structures. Much of the work in the radio consists of spectral surveys. A search for carbon monoxide lines, which are used as a density probe, has been carried out in a number of nebulae (Loren, 1973). These lines are known to be asccoaited with molecular clouds near very young stars. The stars themselves have also been examined. Lepine and Nguyen-Quang-Rien (1974) have found OH emission from M1-92 and HD200775, the illuminating star of NGC 7023. More recently still carbon recombination

lines have been found in cold dense CII regions among early type B stars (cool Stromgren spheres). Examples are the ρ Oph, NGC 2023 and M78 complexes. These point to small HII regions in all these objects (Knapp et al., 1975).

Infra-red observations overlap the radio; they too reveal circumstellar dust shells, in reflection nebulae, primarily around Ae and Be stars. The additional evidence that I.R. work can bring to the problem of star formation in cloud complexes rests on the capability to detect obscured sources. Grasdalen et al. (1973) have discovered over 40 I.R. point sources in the Ophiuchus dark cloud, while Strom et al. (1975) have detected a number of sources in the cloud containing M78. These measurements are believed to show that these are large regions of star formation: the reflection nebulae are merely the most obvious visible examples. Both techniques will continue to add to our knowledge of the general structure of reflection nebulae.

CHAPTER TWO

Hamlet: Do you see yonder cloud that's almost in shape of a camel?

Polonius: By the mass, and 'tis like a camel indeed.

Hamlet: Me thinks it is like a weasel.

Polonius: It is backed like a weasel.

Hamlet: Or like a whale?

Polonius: Very like a whale.

A REVIEW OF OBSERVATIONS OF NGC7023

2.1 Introduction

NCG7023 is a large diffuse nebula located in Cepheus. On account of its high surface brightness it is one of the most intensively observed; yet it is still poorly understood. On the Palomar Observatory Sky Survey red prints it shows a bright central core, approximately 2' x 3' with tenuous filaments visible up to 7 arc minutes from the star, HD200775, the source of the illumination. Less heavily exposed photographs indicate that the core itself is composed of knots and filaments, while the surrounding area is devoid of other nebulosity within a radius of about 5^{o}. Observations have frequently been proven contradictory, even "simple" matters, such as the determination of distance, give cause for conflict. No convincing model of the structure has been advanced.

2.2 The Star HD200775

HD200775, which is also catalogued as $BD67^{o}$ 1283 or MWC361, is the light source for the reflection nebula. Its

1950 co-ordinates are: R.A. = $21^h00^m59^s.7$, Dec. = 67^o $57'55''$. Herbig (1960) classified it amongst his Ae and Be stars, that is bright stars that have just reached the main sequence or were pre-main sequence, and were also associated with nebulosity. As such it might be expected to possess a circumstellar shell. It has been observed most extensively in the visible region but considerable work has been carried out in the IR, and now the ultraviolet is being examined. The distance to the star has been variously estimated as 350pc (Witt and Cottrell, 1980), 440pc (Viotti, 1969), 540pc (Mendoza, 1958) and 600pc (Aveni and Hunter, 1967). The trend of all modern direct observations of HD200775 favours the less remote evaluations of its distance; for this reason a distance of 350pc will be used in the calculations which follow, chapter five. However, Cohen et al. (1981) have once again suggested a more distant position, 500pc, from observations of PV Cephei, which they believe associated with NGC7023. The effect is merely to influence scaling factors internal to the calculations which should not be material to the conclusions. Coupled with the uncertainty in the distance there are various different identifications of the spectral type, these vary from B5e to B2IV, Mendoza (1958), van den Bergh (1966), Aveni (1967), Guetter (1968), Viotti (1969), Strom (1972). The span of measurements is given in Table 2.1.

The fundamental difficulty with the classification of HD200775 is that it has a variable spectrum (Rush, 1975). Usually it exhibits a spectrum in which H-α and H-β appear

Table 2.1 Characteristics of HD200775

Spectral Type:	B2Ve - B5e	I:	6.55
T effective:	16000 - 19000 K	K:	4.56 - 4.59
E(B-V):	+0.58	H:	5.42 - 5.48
U:	7.5 - 7.35 *	L:	3.59 - 3.41
B:	7.8 - 7.7 *	M:	2.5
V:	6.8 - 7.4	N:	1.7
R:	6.86		

* Same observers 3 nights apart. Milkey and Dyck, 1973.

in emission, and the upper Balmer lines are almost completely filled in, however this pattern may be interrupted for periods of a few days. Then the upper Balmer lines are seen with the absorbtion profile typical for a B2IV spectral type. Most observers do not seem to be aware of this variability, despite the wide range of spectral identifications.

Strom et al. (1971) have observed hydrogen emission lines in HD 200775 from H-α and H-ε. They found a double peak in H-β and H-γ through to H-ε show the P Cygni profile characteristic of an expanding shell. V sin(i) was measured to be less than 80 kms^{-1}. Given that HD200775 is a shell star it is reasonable to expect an HII region restricted to small distances from the star.

The first reported ohservation came from Martel (1958). Kleinman et al. (1979) confirmed the presence of the HII region during their re-examination of the AFGL catalogue (HD200775 is listed as AFGL 2695), but the measurements of the size have varied greatly. Witt and Cottrell (1980b), from a study of the colour differences in the surrounding

nebula, deduced that the ionisation was of the order of 100 arcseconds around the star. Even intuitively this appears improbable, for were it so extensive it should certainly have been better documented by previous observers. However Ney et al. (1980) observe that the brightness is the same for a 4 arcsecond diaphragm as it is for a 26 arcsecond apperture, from which they concluded that the ionisation region is confirmed to within a radius of 2 arcseconds.

To examine the realism of this measurement as an estimate of the size of the HII region a crude approximation has been used to estimate the size of the Stromgren sphere, for a star of effective temperature 17000 K. For the purposes of the calculation it is assumed that no stellar photon leaves the solid angle into which it was emitted and the Planck distribution is used to calculate the number of ionising photons. It is true that the shape of the black-body spectrum is not a good approximation to the radiation law for stars in the ultraviolet, however it still gives an acceptable estimate to the total number of photons (Osterbrock, 1974). Given the uncertainty in the stellar parameters the calculation was performed for a selection of the values listed above, but under the assumptions noted the variability was only by a factor of two or three-fold. Obviously different values of the bolometric magnitude are associated with different displacements for the star, thus some of the variation tends to cancel out.

The amount of radiant energy crossing a unit area in a unit solid angle in a unit frequency range in a unit time is:

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(h\nu/kt) - 1} \qquad (2.1)$$

Then the number of ionising photons per second is simply the integral of this quantity divided by the energy per quantum, where the lower bound for the intergration is the Lymann limit (91.2 nm). Therefore in the range considered we obtain:

$$\frac{dN}{dt} = \int_{Ly}^{\infty} \frac{B\nu \, d\nu}{h\nu} \approx \frac{2}{c^2} \int_{Ly}^{\infty} \frac{\nu^2 \, d\nu}{\exp(h\nu/kT)} \qquad (2.2)$$ per unit area

the approximation being valid in the integration range considered. To obtain the total number of photons per second it is necessary to multiply by a geometric factor.

The equilibrium radius of the Stromgren sphere is given by:

$$R_s^3 = \frac{3}{4\pi n_i n_e \alpha} \frac{dN_i}{dt} \qquad (2.3)$$

(Osterbrock, 1974).

where α, the recombination factor, dependent on the temperature of the gas, is approximately $4 \times 10^{-7} m^3 s^{-1}$. The electron and ion density, n_e and n_i respectively are usually presumed to be equal. These can be estimated from the E (B-V) measurements, combined with the column density measurements of Bohlin, Savage and Drake (1978). They found from a study of interstellar HI the following relationship:

$$N_H/E(B-V) = 5.8 \times 10^{21} \text{ atoms } cm^{-2} mag^{-1} \qquad (2.4)$$

where N_H is the column density. Witt and Cottrell (1980) and

Viotti (1976) have variously measured E (B-V) as 0.44 and 0.29, which corresponds to ion densities of 245 cm^{-3} and 160 cm^{-3} respectively. Using the smaller value since this maximises the Stromgren radius:

$$R_S = 7.2 \times 10^{13} m \qquad (2.5)$$

At a distance of 350pc this corresponds to a radius of 1.4 arcseconds, i.e. an aperture of 2.8 arcseconds. Now the boundary region between ionised and neutral media for ionisation by radiation, is sharp, typically ten per cent of the Stromgren radius. Therefore an apperture of 3 arcseconds should be sufficient to enclose the HII region. This calculation is extremely crude, but it should place an upper limit on the radius of the Stromgren sphere. Thus Ney et al. seem to be close, but slightly generous, in their experimental determination. As a consequence of such a calculation it is necessary to estimate the stellar radius, which is found to be approximately 3 or 4 solar radii.

Garrison (1978) in spectrophotometric observations has noted a smaller Balmer discontinuity than that for main sequence stars, which is now to be associated with Herbig's Ae and Be stars, and also an infrared excess. This latter observation confirms that of Gillet and Stein (1971). Garrison's models of reddening for Herbig's emission stars indicate that the IR excess cannot be explained by the free-free emission alone, as suggested by Milkey and Dyck (1973), thus giving further weight to the idea of a circumstellar shell. He does not rule out the possibility that thermal bremsstrahlung may contribute to the IR

emission but it is not the main effect. On the assumption of a stellar radius equal to 3.8 solar radii Garrison calculates the mass loss rate to be between 10^{-7} and 10^{-6} $M_\odot$ /yr (the variation being dependent on the assumed densities for an expansion velocity of 70 km s^{-1}).

At longer wavelengths Loren et al. (1973) have found carbon monoxide emission, the peak being coincident with HD200775, while Lepine and Nguyen-Quang-Rien (1974) suggest that the expansion is restricted to within a few tens of stellar radii, from their OH measurements.

In the ultra-violet Walker et al. (1980) have discovered that the stellar spectrum is anomalous. The principal difference is that the 2175 Angstrom feature is much weaker than the norm. It most resembles, but is not so extreme as θ Ori. They suggest that the grains close to the star have been processed in a different manner from normal interstellar dust.

In summary, HD200775 is a variable Be star surrounded by an expanding dust shell, and also a small HII region.

2.3 The Nebula NGC7023

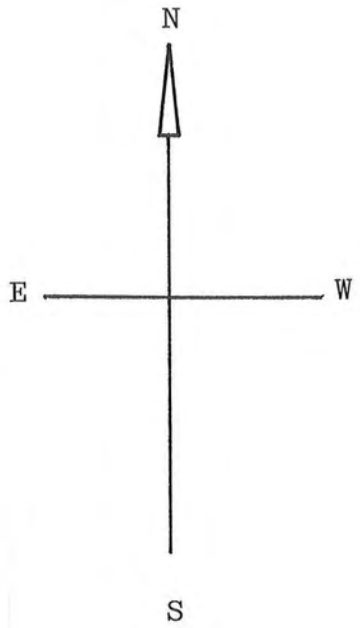
The nebula NGC7023 is a reflection nebula in a much larger molecular cloud. It has been observed extensively in the optical and the infrared, Gehrels (1967), Zellner (1973), Whitcomb et al. (1981), etc., but the size and shape have not been delineated. The nebula is situated towards the northern limit of a molecular cloud which is approximately rectangular and 30' by 60' in extent. About 5' west of HD200775 there is a region of increased star counts: by a

factor of three or so. Weston (1953) was able to show that the stars in this region, brighter than the nineteenth magnitude had no common proper motion and so must be an unrelated background group, seen through a "window" through the nebula. To the east, at about 7', there is a highly obscured region. The apparent shape and size of the reflection nebula is highly dependent on the integration time used. Plates 1.1 to 1.4, courtesy of Dr. S.M. Scarrott and Dr. R.F. Warren-Smith, were taken at the Wise Observatory, in Israel, with the McMullen electronographic camera in August 1981, and demonstrate clearly both the variation in shape and the "growth" of the reflection nebulosity with integration time. Quite obviously NGC7023 is an intensity limited rather than a density limited object. The alteration of the apparent shape with integration time and generally smoother appearance towards longer wavelengths have permitted entirely contradictory statements to be made about the structure. Zellner (1974) described it as, "thick and chaotic" while Whitcomb et al. (1981) chose to study it, "because of its simplicity". To be fair to Whitcomb and his collaborators the infrared structure does seem somewhat less complex than that in the visible, but they do stress the similarity between the far infrared brightness and the optical form of the nebula.

Elmegreen (1980) has suggested that the whole complex may be the site of low mass star formation within a comparatively small molecular cloud, of the order of a thousand solar masses. If this were so, she hypothesises,

PLATE 1.1

NGC 7023 BLUE LIGHT



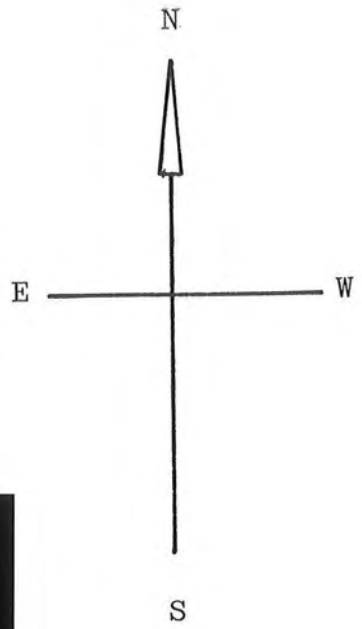
SCALE : 1mm : 9.5"

Simulated Exposure Time 15 Minutes

Reflection of Television Screen in South-East Corner

PLATE 1.2

NGC 7023 BLUE LIGHT



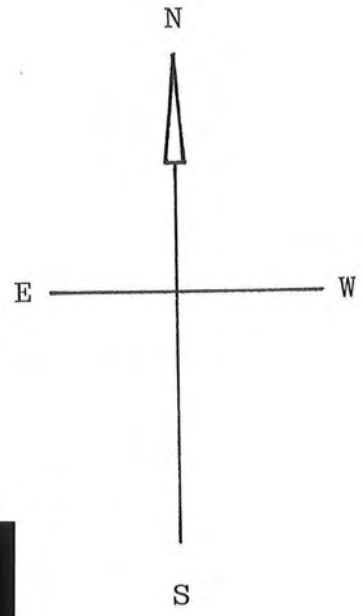
SCALE : 1mm : 4.8"

Blow-Up of Central Region of Plate 1.1

Reflection of Television Screen in South-East Corner

PLATE 1.3

NGC 7023 BLUE LIGHT

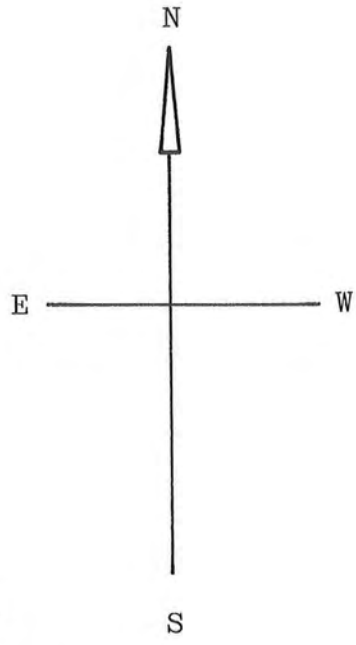


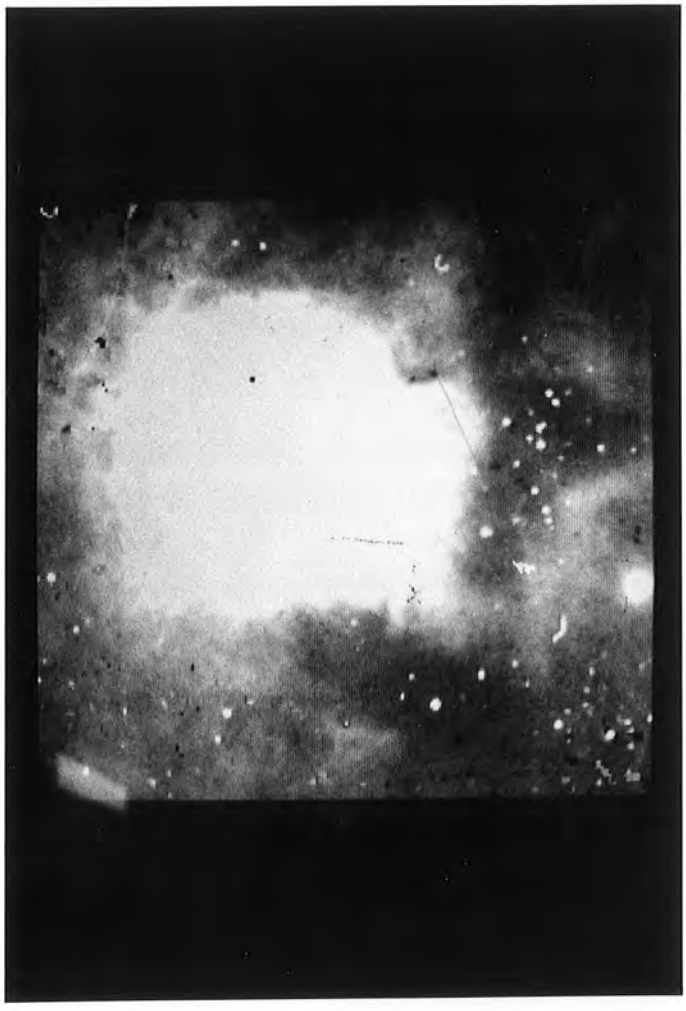
SCALE : 1mm : 9.5"

Simulated Exposure Time 30 Minutes

PLATE 1.4

NGC 7023 BLUE LIGHT





SCALE : 1mm : 9.5"

Simulated Exposure Time 45 Minutes

that the star formation might be occurring within the entire complex, rather than being restricted to the boundaries of the molecular cloud. This would distinguish star formation in small molecular clouds from that in the giant molecular clouds, where the models of Israel (1978) suggest only edge-formation. There is considerable observational evidence to support this suggestion. Strom. et al. (1972) have placed HD200775 slightly above the main sequence and so classified it as a pre-main sequence star. As early as 1953 Weston had discovered a cluster of irregular variable stars in the vicinity of HD200775. These resembled RW Aurigae or T. Tauri type stars, some exhibiting appreciable light variation during intervals of only a few hours. Furthermore this central cluster had a common proper motion. Weston concluded that they formed a T-association and so identified the region as the site of recent star formation. Continuing work by Rosino and Romano (1962) has identified fourteen T-Tauri variables in the field of NGC7023, and of these three (Romano, 1975) are found in binary systems.

Strom et al. (1972) mapped a 75 square arc minute area about HD200775 in a search for other embedded stars. Since shell emission would be maximal at approximately 5 μm they scanned the area at 1 μm. Although this wavelength was not quite optimal, it permitted a much faster scan of the area. Only about one star of the colour index found is to be expected in such a restricted field (Allen, 1963) so they could rule out M-giants or M-dwarfs, when they discovered several such sources. Price and Walker (1976) claim to have identified the variable HZ Cep, in the eastern

part of the reflection nebula with an infrared source seen at 11 μm and 20 μm. Elmegreen and Elmegreen (1978) found a CO peak at the position of a 6cm continuum source identified by Pankonin and Walmsley (1978) at R.A. = 21^h02^m Dec. = $67^o58'$ and presumed that this was yet another embedded object. This however appears unlikely as it has a non-thermal spectrum. Haslam et al. (1978) believe it to be 4C67.34, a background source, which just happens to coincide with the position of the molecular maximum.

2.4 The Optical Properties

The visual surface brightness of the nebula in the central region is approximately 20 magnitudes per square arcsecond. However the brightness distribution is irregular and does not exhibit a smooth monotonic decrease in intensity with increasing angular distance from the star. Table 2.2 summarises the photometric observations, the variations are due to the highly irregular brightness distribution, the different parts of the nebula sampled and by the variation and improvement in equipment with time.

The most detailed and most recent observations of surface brightness have been reported by Witt and Cottrell (1980). They scanned across the nebula in the four cardinal directions, sampling 38 points in the range 30" to 300" from HD200775. Unfortunately they normalise their brightness measurements in terms of the brightness of the central star without stating what magnitude they take for HD200775. Given the variability of the star it is not possible to present their results in magnitudes per square arcsecond.

Table 2.2 Surface Brightness Observations of NGC7023

Author	Mag /arcsec^2	Notes	Year
Keenan	19.5	mean value 0.5' from star	1936
	21.7	mean value 3.2' from star	
Collins	21.78	overall mean	1938
Martel	20.58	0.8' south of star	1958
Johnson	23.2	photoelectric 3' north	1958
	22.91	photoelectric 1.5' north east	
Grygar	22.9	photoelectric 1.1' south	1959
	22.79	photoelectric 1.6' east	
Vanysek	20.77	photoelectric 1.0' north	1963
Zellner	19.45	photoelectric 0.5' mean	1969

Most observers recognise the complex structure of the nebula, Witt and Cottrell note a bright ridge at about 175" east of HD200775 and a prominant brightness enhancement 125" and 200" North. This structure is most apparent in the ultra-violet. They observe further a brightness plateau around an average offset distance of 125 arcseconds, in all directions save to the west. Given the difficulties in categorising the structure directly authors tend to examine the general behaviour with increasing radial distance. All note a general fall off in intensity with increasing distance from the star. Witt and Cottrell (1980) express this as a power law relationship:

$$s \propto r^p \qquad (2.6)$$

where S is the surface brightness of the nebula, r the radial offset distance and p some empirically determined constant.

Correcting for the size of their apperture - depending on the steepness of the surface brightness fall off, the effective distance from the star will be less than the pointing distance - they discover that a value of p = -1.44 fits the available data, averaging over the entirety of the nebula observed in the waveband 5500 Å to 3500 Å.

This is in complete contrast to their results for an examination of the Merope nebula (Witt, 1977), where they found p to vary monotonically from -0.41 to -0.84, for the same wavelength interval.

Vanysek (1969) during an analytical investigation of light scattering in a homogeneous spherical nebula, found that a value of p = -0.1 should be expected from such conditions. Thus the claim by several authors (Svatos 1964; Vanysek 1969) that NGC7023 is a uniform spherical nebula is no longer tenable.

The colour variation with distance has been investigated by many observers; Keenan (1936) was the first using photography and found that the nebula was very nearly white but slightly redder than the star, in some directions. Schalen (1945) interpreted this behaviour in terms of scattering from an embedded star by metallic particles of diameter between 0.1 and 0.05 microns. However no other observer has found this result. Rather the consensus is that the nebula is bluer than the star (Greenstein, 1948) with the definite exception of a region approximately 100 arcseconds North, which is redder than the star over the region 4700 Å to 3500 Å. Beyond this agreement is poor. Measurements by

Elvius and Hall (1966) are usually in agreement with Martel (1958) but Vanysek and Svatos (1964) are frequently substantially redder, by 0.3 magnitudes or more. Witt and Cottrell (1980) who have investigated the colour at more positions than any other observers have noted a general trend. The nebula is only slightly bluer than the star near (less than 30 arcseconds) to the star, but it becomes increasingly bluer to a maximum between 100 and 150 arcseconds, depending on direction. Beyond this point the colour behaviour varies with wavelength. Colour differences between 4100 Å and 4700 Å remain blue both east and west of the nebula, but between 3500 Å and 4100 Å the colour tends to become redder, with maximum red value at about 300 arcseconds. In the positions of observational overlap they found broad agreement with the observations of Elvius and Hall (1966). Witt and Cottrell remarked that the position of divergence in colour behaviour occurred at approximately the same radial distance, 100 arcseconds, as the plateau in intensity they had noticed. They suggested that this might indicate the presence of a cavity of angular radius 100 arcseconds about HD200775, yet at the same time were convinced that the star was embedded. It is difficult to see how they reconciled these concepts.

Optical polarisation was first observed in NGC7023 by Henyey (1936) who found a mean value of 12 percent, but this value was calculated from the mean of only three points. Gliese and Walter (1951) observing in blue light claimed to detect values between 0 percent and 56 percent,

with a mean value over the nebula of 16 percent. Their very high values have never been repeated and we assume therefore that they were erroneous. In the following year Weston (1953) discovered marked variation in the polarisation from feature to feature within the nebula but found no confirmation of the highest values of Gliese and Walter. Also in blue light, he did not discover any polarisation in excess of 30 percent. Martel (1958), again in the blue, found values between 10 percent and 20 percent, but also in a third of the areas she examined that the polarisation was significantly non-radial. Elvius and Hall (1966) measured polarisation at three wavelengths: 3750 Å, 4500 Å and 5600 Å. Contrary to another finding of Gliese and Walter, that the degree of polarisation was independent of radius, Elvius and Hall discovered a trend towards increasing polarisation with radial distance, which then peaked and declined. The most significant feature was the very strong radial pattern, which proved unequivocably that HD200775 was the light source of the nebula. They also investigated the wavelength dependence of the polarisation. This differed throughout the nebula, some parts indicating no significant variation while others showed an almost linear rise with inverse wavelength. Gehrels (1967) has also observed NGC7023 at 3600 Å, 5600 Å and 7400 Å, while Zellner (1970) has done so at 4500 Å. Zellner however observed only one region - 38"NE of HD200775. Both Gehrels and Zellner found systematically higher values than Elvius and Hall, but Zellner

should not be considered as an independent observer, since he collaborated with Gehrels. However in the single region that they all considered they measured the same gradient of polarisation versus inverse wavelength. Gehrels was convinced that this demonstrated that the grains were exhibiting the same wavelength dependence as interstellar grains. This is not in accordance with the more extensive sample due to Elvius and Hall, who found markedly different behaviour at different points in the nebula. If it is presumed that the large difference between their absolute determinations of percentage polarisation results from the different appertures used and the differing values of the 'sky background' subtracted, and therefore that the relative differences are accurate, then one must deduce a variation in grain properties throughout the nebula. Density fluctuations alone would not be sufficient to account for this variability. Perhaps the only simple explanation would be a variation in the grain size distributions. The most recent investigation of the polarisation (at 6500 $\mathring{A}$) by Ney et al. (1980) found small scale variations in the structure, particularly a region of 35% ± 5% polarisation along a, "forked tongue of nebulosity" to the south west of HD200775, but added nothing to the general impression already developed. Thus the most significant features of the polarisation are: the strongly centro-symmetric pattern around HD200775, a rise and fall in polarisation with distance from the star, and a variation with wavelength, which is itself position dependent, throughout the nebula.

2.5 Towards a Structural Understanding

One fundamental difficulty in understanding any, but the closest, of extensive astronomical objects is that the structure is observed only in projection and no other angle of view is possible. In the case of NGC7023 this has permitted considerable freedom to suggestions for the structure. Until recently NGC7023 has been regarded as, "one of the finest examples of a reflection nebula with a deeply embedded illuminating star" (Witt and Cottrell, 1980a) and as such has attracted theoretical interest primarily in attempts to represent it as either a plane-parallel slab (Witt and Cottrell, 1980b) or as a spherical cloud (van Houten, 1961; Vanysek and Svatos 1964; Vanysek, 1969). However as Witt and Cottrell (1980b) concede, "the detailed geometry of the HD200775/NGC7023 system is not known".

The most recent evidence of all, from an infrared survey by Whitcomb et al. (1981), shows that the structure of the optical depth in the infrared is inconsistent with either a homogenous spherical cloud or plane-parallel slab. They tend towards the 'blister' structure proposed for ionisation, caused by massive early type stars, by Habing and Israel (1979): the present author considers the "bubble" or "blow-hole" models of Cohen et al. (1975) as more appropriate, on the grounds of stellar type.

Even the optical depth of the nebula has been disputed: Gehrels (1967) concluded that NGC7023 was optically thin from early surface brightness observations. This contention received support from the non-geometry dependent polarisation models of Zellner, which showed moderate success in reproducing

spot values of the polarisation. As Zellner's models required an optically thin nebula - they considered only single scattering - they seemed to confirm Gehrels. However it is now realised that optically thick nebulae, with highly "forward-throwing" scattering particles (i.e. most of the flux is scattered at 0^{o} to the angle of incidence) can produce substantially the same intensity structures. Additional photometric observations have now resolved the issue in favour of an optically thick nebula, Witt and Cottrell (1980a).

Unfortunately the simple question as to the relative positions of HD200775 and the cloud boundary is far from solution. It would seem a trivial exercise to deduce, from the reddening of the star, whether or not it is embedded within the cloud. It has not proved so. Estimates of the total reddening are in approximate quantative agreement. Elvius and Hall (1966) measured E (B-V) to be 0.6 mag. while Duke (1951) ahd found 0.56 mag. Values as low as 0.54 mag. and as high as 0.7 mag. have been reported, but these are the extremes, (Strom et al., 1972; Viotti, 1976). Just what fraction of this extinction is inherent to the nebula itself and how much is interstellar is in no way obvious, being dependent on the assumed spectral type of HD200775 which does vary.

Elmegreen and Elmegreen (1978) from a study of star counts found a ratio, r, of approximately three times as many stars in the hole, in the central portion of the cloud, as in an equivalent area of the cloud itself. A typical region beyond the boundaries of the entire molecular cloud

produced the same result. Following Dickman (1976) they used $A_V = 4 \log_{10}(r) = 1.9$, and so deduced that the totality of the extinction was interstellar (N.B. the factor 4 is a function of direction in the Galaxy and was extracted from the tables of van Rhijn, 1929). This was also the conclusion reached by Walker, Yang and Fahlman (1980) from the ultraviolet extinction curve, from IUE spectra between 1250 Å and 3200 Å. The difficulty here is the choice of the standard star. Comparison was made with θ Ori, but a good spectral match is difficult to achieve on account of the variability of HD200775.

It is normally possible to obtain an independent estimate of A_V by extrapolating photometric observations to the infrared, where extinction tends to zero. However HD200775 is centred on an infrared source, i.e. it heats the dust in its vicinity. The thermal emission generated vitiates any attempt to estimate the stellar brightness, by this means.

New estimates of the intrinsic reddening are in considerable contrast to earlier optical or ultraviolet measurements. Racine (1966) reported E (B-V) = 0.38, intrinsic to the star, and Viotti (1976) measured the colour excess to be between 0.28 and 0.41, depending on the method of measurement used. Witt and Cottrell (1980b) have cast doubt on some of the initial assumptions used by Viotti to interpret the TD1 satellite data, most importantly his proposition that the 2200 Å interstellar absorption feature is normal for HD200775. Here they are in agreement with Walker et al. (1980) in suggesting that the extinction hump

is abnormally weak. Coupling this interpretation of Viotti 's work with the new surface brightness measurements in the optical they estimate the colour excess to be between 0.13 and 0.26, which is much more in line with the Be stars. It should be noted that these estimates are based on a spectral identification of HD200775 as B3 IV - V. As it has been catalogued as early as B2 and as late as B5 these must be considered as conservative limits. Thus the best modern measurements taken together permit a variation in the intrinsic colour excess from 0.0 to 0.26. Clearly it is not possible to deduce the relative positions of the star and the cloud from these observations.

Nonetheless considerable effort has been applied to the modelling of NGC7023. To date all models have assumed an embedded star, but then most work was analytical, and at a time when the intrinsic reddening was considered rather high. Both van Houten (1961) and Vanysek (1969) considered spherical structures, the latter also modelling annular shells. The best evidence for a spherical geometry was provided by the photometry of Grygar (1959) who found a radial power law fall off with index -0.1. This is characteristic of a spherical structure illuminated by a central star, which is all that can be handled with facility analytically. Vanysek also considered the colour distributions that should be produced in such a structure, but given the disagreements in the measurements, was unable to reach any conclusion.

More recent evaluations of the index to the radial power law (Witt and Cottrell, 1980a) do not reproduce Grygar's value, so we must now doubt a spherical structure. Witt and Cottrell proposed a plane slab, but the modelling was so inconclusive that they were able to suggest from the same colour data that there might be a cavity around HD200775 of the order of 100 arcseconds, in radius.

Whitcomb et al. (1981) have carried out a survey of NGC7023 in the infrared which involved mapping the temperature and optical depth near HD200775. Both peak near, but not precisely at, the position of the star. It was immediately clear that the density was highest to the northwest. The simplest interpretation of the infrared data was that HD200775 was close to the edge of a molecular cloud and surrounded by a circumstellar shell. Calculations by Icke, Gatley and Israel (1980) for the infrared appearance of blister HII regions closely resemble that of NGC7023, and even predict the slight displacement of the optical star from the position of the IR maximum.

Clearly there are difficulties in applying the "blister" or "champagne" explosion models for HII regions around bright stars as all calculations to date, Icke (1979), Whitworth (1979), Tenorio-Tagle (1979), Bodenheimer et al. (1979), Israel (1978), Kandel and Sibille (1978) are for O stars on very early type B stars (Castor, McCray and Weaver 1975) and HD200775 is much less energetic. Yet the possibility that there may be a cavity in the molecular cloud containing NGC7023 certainly bears further investigation. Consequences of such a structure, from theoretical calculations, can be

axial symmetry. This was reported in nearly all the models noted above, and also instability in circumstellar envelopes or cocoons (Kahn, 1974).

In conclusion, opinions of the structure of NGC7023 are fluid and changing; reddening data is unable to resolve the problem of the relative geometry of the star and nebula. Spherical, plane-parallel and blister models have been considered. None has yet shown itself to be conclusive but the last now appears the strongest hypothesis.

CHAPTER THREE

Sensus, non aetas, invenit sapientiam -
Observation, not old age, brings wisdom.

Publilius Syrus

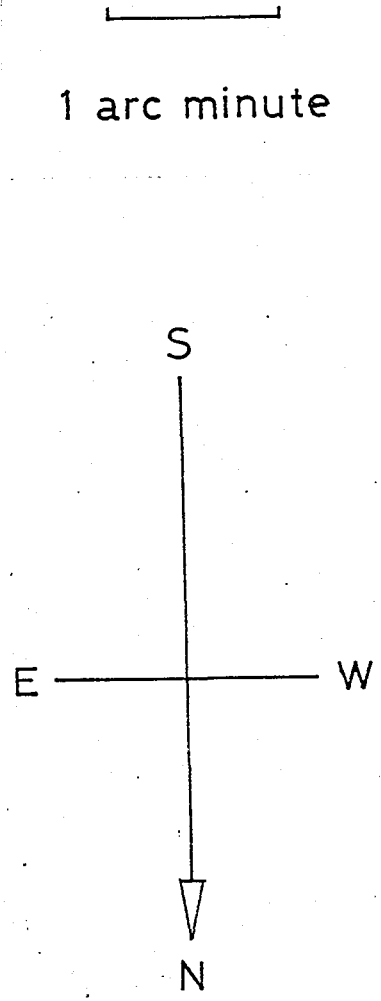
3.1 The Observational Details

New observations of NGC7023 were obtained in polarised light, in July 1980, at 4400 Å, 5200 Å and 6500 Å. The observers were Drs. S.M. Scarrott, P. Jorden and H.G. Perkins and the telescope, the Wise one-metre of the University of Tel Aviv. The acquisition system comprised the McMullen electronographic camera and the Durham polarimeter. Complete descriptions of these devices and the associated reduction programs are given by Pallister (1976), Axon (1977) and Warren-Smith (1979). The imaged area, centred on HD200775, was a rectangle 7.5 x 5.5 arc minutes squared.

3.2 The New Data

One idiosyncrasy of the Durham polarimeter is that images are measured in strips, containing only half the full field. Thus complete polarisation and intensity maps were produced for the broad V-filter, Fig. 2.1 and plate 2.4 respectively, and half maps of polarisation in red and blue, Figs. 2.2 and 2.3. The integration areas for the pixels plotted are 21 square arcseconds. Since the complete visual map is the product of the merger of two distinct half-maps, taken in alternating strips, it was necessary to align the separate images. This was accomplished using two different methods. First stars that had already been identified in the field of the two images were used; these had been needed

Fig. 2.1 — 60% POLARISATION



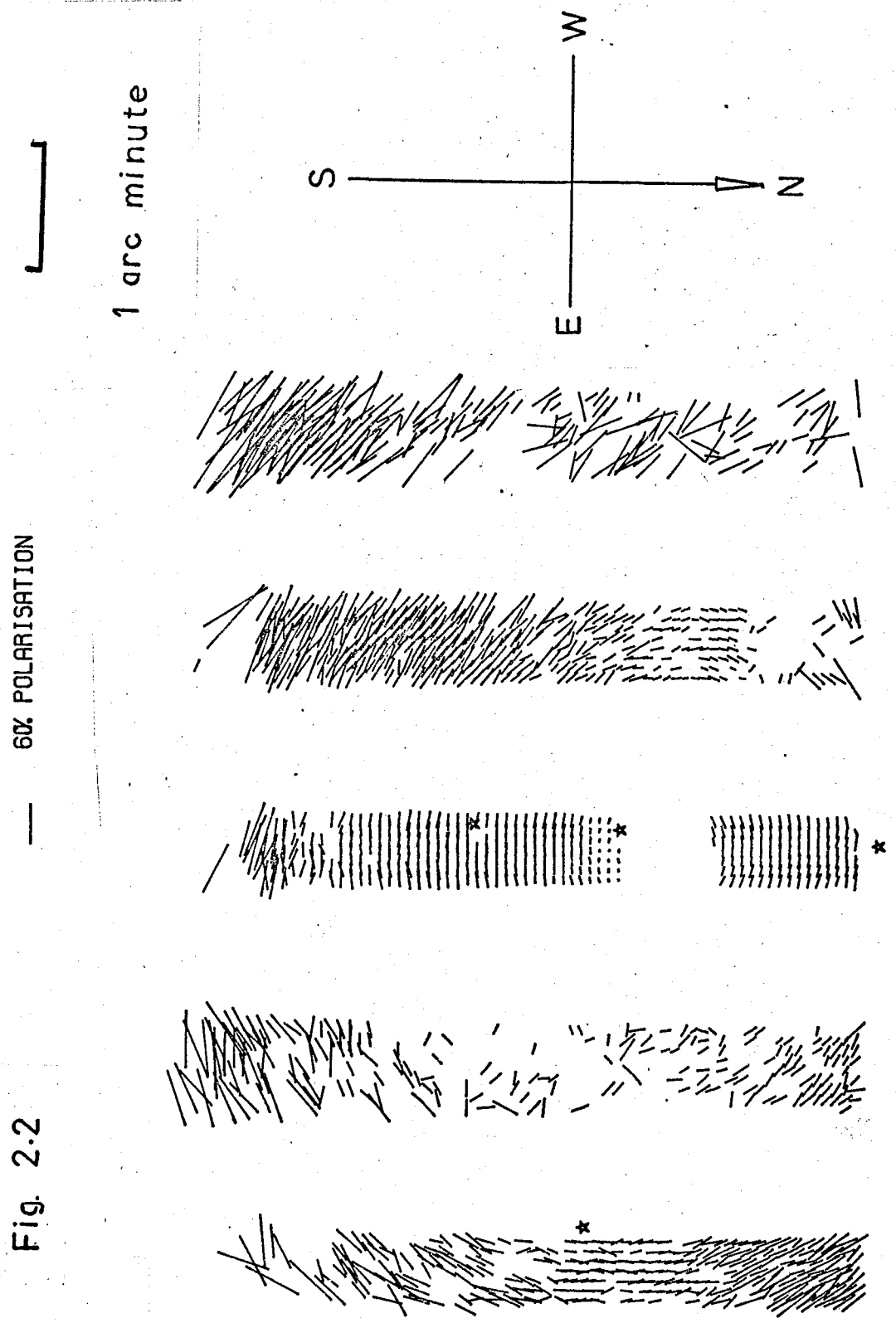


Fig. 2.2

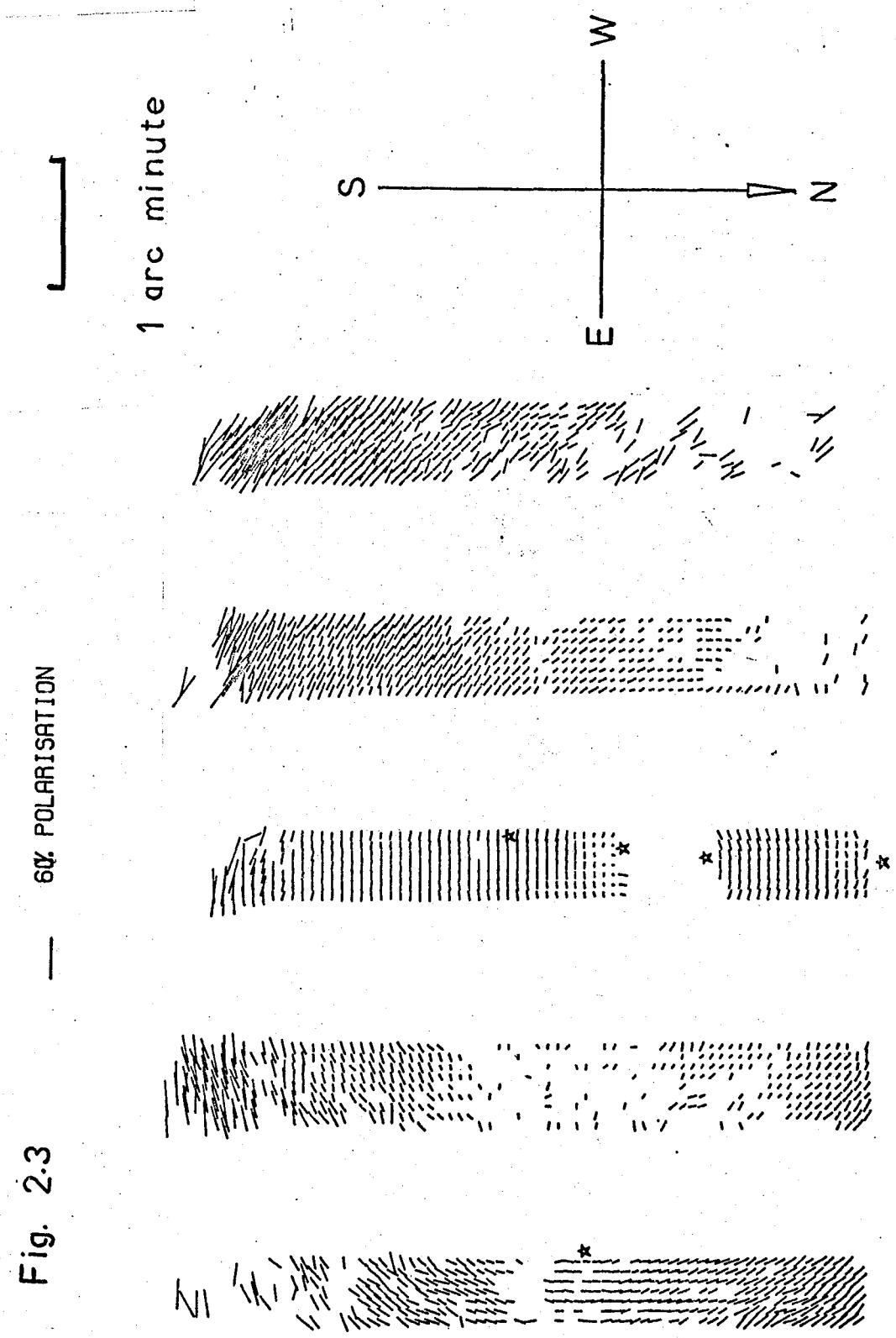
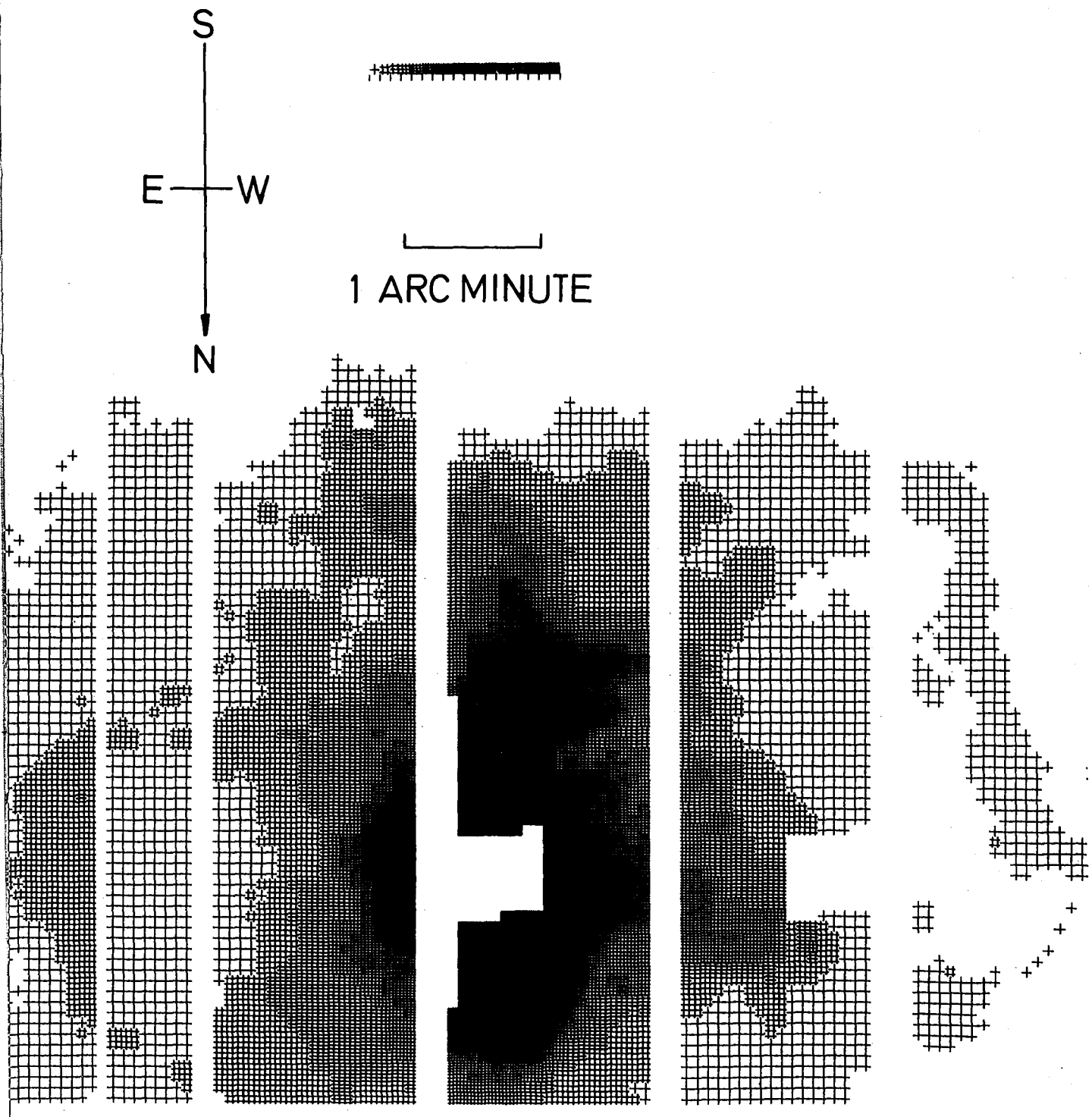


Fig. 2.3


S
E
W
N
1 ARC MINUTE

anyway to superimpose the electronographic plates. Taking these stars and using knowledge of their positions with an approximate knowledge of the orientations of the two half-maps it was possible to refine the approximation by comparing the distances between stars on the two separate images with those apparent on the published photographs of Keenan (1936) and Scheur and Gehrels (1967). The second independent method was to use the form of the polarisation pattern and to assume continuity of vector behaviour across grids - an excellent assumption in the brightest parts of the images. The pattern closest to HD200775 is quite unmistakably centro-symmetric. This is apparent on all the maps, and is entirely to be expected, both from previous less complete observations (Martel, 1958; Elvius and Hall, 1966) and from the very simple analytical prediction that the pattern of polarisation caused by dust scattering, around a single illuminating source will be centro-symmetric. On the grounds of symmetry alone it is obvious that such a pattern must exist, but symmetry by itself could not indicate whether the pattern could be radial or in the form of concentric rings. These patterns are invariably ring-like, as detailed calculations predict. Knowing that close to the star the vectors are circularly symmetric it is a trivial matter to extend the perpendicular bisectors of the vectors to their point of intersection at the centre. The two separate halves of the map are moved until their centres are co-incident. The author is grateful to M.R. Gething for the use of his program to accomplish this. The inevitable complication is that there will be some degree of error on the values of the angles

of the polarisation vectors, so that even the perpendiculars on the same map will not quite intersect at the centre. To avoid this problem an interactive procedure employing a least squares fitting method, which rejects particularily bad vectors was used. Knowing the position of the true centre it is a simple matter to shift the origin of one half of the map to bring both of them together.

3.3 Detailed Aspects of the New Observations

The new observations are plotted in map form; for clarity each has a lower level intensity cut off, that reduces contamination from low intensity vectors, which are frequently random, chaotic and not indicative of the underlying pattern. Plate 2.4 is a greyscale intensity map, of rather poorer quality than the photograph-like illustrations in chapter two, but is very useful for superimposing polarisation and intensity and thus understanding the variations in polarisation. Fig. 2.5 shows a polarised intensity map - the product of percentage polarisation and intensity - for the broad V-filter. Here the vector length corresponds to the strength of the polarised intensity, but the position angles of the vectors are those of the percentage polarisation. The almost square hole at the position of the central star, HD200775, is caused by an opaque mask which blocks direct illumination by the star. In the absence of such a screen damage would result to the photocathode of the camera. There is also another similar hole in the data on the VV-images, three grids - 145 arcseconds - west of HD200775. This gap is also 45 x 45 arcseconds square. The normal practice during

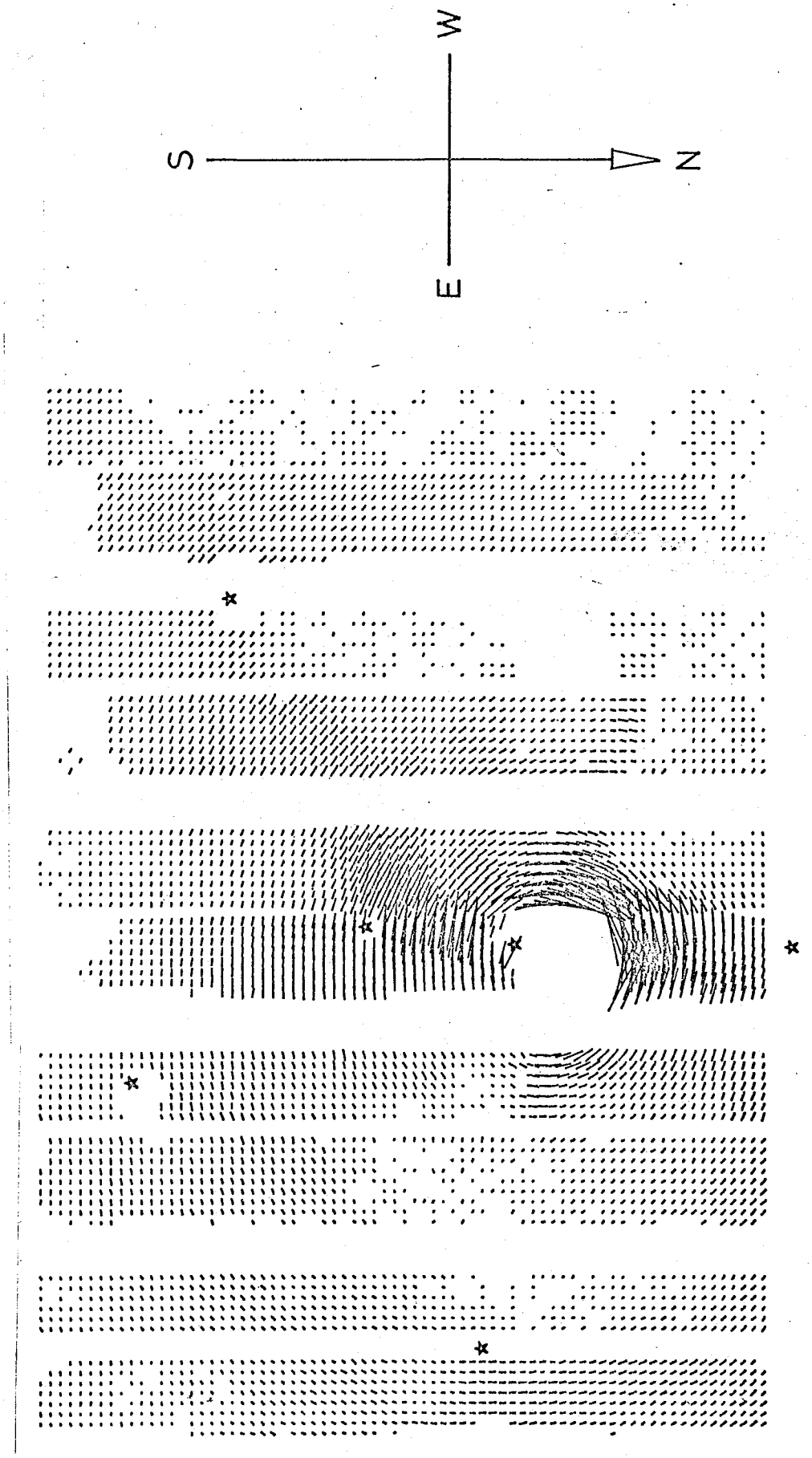


Fig. 2.5

observations is to take electronographs of half the image, then alter the position of the movable grids. However, on this occasion, the observers varied the tracking position of the telescope, guiding approximately three grids to the west of the star for the second half of the image. As a consequence a portion of the image was obscured unnecessarily by the opaque mask which was used to block the light from the central star; this is the second blanked position. This movement is also the explanation for the consistent alteration in the size of the grid gaps: wide, narrow, wide, narrow, etc. The smaller scale irregular variation is however, an artefact of the slight variability in the size of the grids. Furthermore, since data corruption, perhaps due to imperfect alignment during observation, is particularly common on edge boundaries, such data is often rejected during the reduction sequence; this tends to produce a marginal widening of the spaces between grids. Marked on all the maps are the star positions used to align the images, but with the notable exceptions of the one, or two, "stars" marked on the edges of the central intensity peak, i.e. the masked region, which are not in fact faint stars in the field at all, but rather the sharp edges of the central intensity peak, cut off on either side of the image by the opaque mask. To the reduction programs these two prominences are indistinguishable from stars, and so may be so used for the image alignment.

3.4 Consideration of Defects

The major source of error in all modern polarimetric studies is the subtraction of the background sky signal. The magnitude of this error has quite obviously been under-

estimated in the past. Both Zellner (1970) and Gehrels (1967) express considerable surprise at the large departure between the values they recorded and those measurements of NGC7023 made by Elvius and Hall (1966). Earlier observations had not begun to indicate any such problem since none were sufficiently complete to permit point by point comparison. The main exception to that statement were the observations of Gliese and Walter (1951) which, even averaged over the entire nebula, were vastly in excess of those of any other observers and are generally regarded as the product of some experimental failure.

The solution is to discover some area of sky, in the vicinity, which is representative of the sky-background uncontaminated by the object. The standard methods either beam switch between sky and object, or pan off the object and integrate on a separate region of sky. This has the consequence that different areas of sky are used to normalise the several parts of the object examined. The inherent inconsistency is manifestly obvious. The Durham group attempt to negotiate this difficulty by another method. Since extensive images can be observed it is possible to select areas of low intensity within the field of view and to use these as the sky regions. The advantages of this method are: (a) that there is no need to expend observing time on sky sources, (b) that a choice of sky regions are available after the observations are made, thus giving a better informed selection, and (c) greater consistency in the sky subtraction, since the same sky area is used to reduce the entire image. The primary difficulty - little or no sky, in

the field of view, only appears for large objects, or when the field is artificially restricted. With lengthy exposure times, of course, many objects appear vastly larger than might be anticipated from shorter exposure time plates.

A secondary problem frequently encountered (Gething et al., 1982) is that of variable apperture size. As an apperture increases in diameter the area measured will also grow. In observations of NGC7023 the appertures of diaphragms or baffles have varied considerably. Elvius and Hall (1966) used diaphragms corresponding to 29" and 4", while Martel (1958) used typically 35", Gehrels (1965, 1967) and Zellner (1970) used 60", and Witt and Cottrell (1980a) 50". Other observers have not indicated the appertures that they used. Now depending on the steepness of the slope of the surface brightness fall off, in all directions, the integrated intensities will differ, with apperture size. In general only a flat intensity gradient could be expected to produce consistent results for varying diameters. Since polarisation is a second order effect, the apperture size is even more crucial. With these new measurements we are better placed than ever before to make comparisons, since not only do we have extensive coverage of most of the brighter parts of the nebula and a consistent sky subtraction, but also we are able to add our picture elements together and thus reproduce the integrated effect of much larger appertures.

3.5 Examination of the Maps

Plate 2.4 is a greyscale intensity map taken through a broad V-filter, processed by a plotting routine adapted by M. Gething from the greyscale routine due to Dr. W.S. Pallister, with a minor alteration by the author. It draws a series of crosses for each pixel, the total number of crossed lines being proportional to the intensity of the image at that point, but with some constant intensity step length decided by the program. Thus a "negative" image is developed, but with upper and lower level intensity cut offs which may be chosen by the user. Different features of the data may be accentuated by varying the intensity step length or changing the cut offs. Plate 2.4 has been deliberately selected to display the heterogeneity of the nebular structure with reference to the polarisation maps.

Considering the overall appearance of the map it can be remarked that the sky subtraction from each half was fairly consistent between the two. One slight defect, not apparent on the intensity maps, will be discussed further below. As the two half maps were obtained and reduced separately, care is required to ensure that the zero point of the intensity scale for one of the maps is adjusted to that of the other. Slight differences in observing time, or, emulsion quality, may alter both the zero level and the step length. Another point to note is that different areas of "sky" must be used for the two halves of the map; this should aid in the identification of faulty sky subtraction. The plotting program was instructed to interpolate between nearby contiguous grids. It was permitted to extend a

grid edge by as much as two picture elements. This interpolation is not perfect, some discontinuities in intensity as a result can be detected, but the general validity is well illustrated by the most westerly grid. It is quite impossible to discern directly that this is composed from two separate grids. In total the figure shows ten grids, although only six bars are obvious. Gaps in the data, or points below the lower intensity cut off are left blank. The sharp edge to the north, on nearly all grids, results from the slight northern displacement of the centre of the field, marking the edge of the imaging system. The total area is approximately 470" by 330" in the field of view.

The intensity behaviour is clearly irregular and asymmetric. On the grossest scale, there is a general fall off in intensity with distance from the centre, but there are many shorter scale variations. North to south, within 90" of HD200775, the intensity fall off is rather smoother than that east to west, but approximately 55" south from the star a (bright) filamentary structure running south-westward is apparent. East-west the brightness falls quickly to 100" to 120" from the centre but rises again at about 200" in both cases. Thus the nebula has a curved bright filamentary structure which gives it almost a "two-handled" though slightly skew, appearance; the easterly "handle" is clearly smaller and more intense than the westerly structure. The central regions of the nebula have an elongated appearance. There are many local inhomogeneities: 125" SSE there is a small dim area, while about 130" WNW there is a brighter forked tongue projecting into a lower intensity area.

The polarisation maps require examination with some care. HD200775 is seen to be the only important source of illumination for NGC7023. A consideration of those sections of the grey-scale map (plate 2.4) most distant from HD200775 indicate that these are areas of lower illumination and thus most liable to photon counting and systematic errors. The chief systematic error is, as always, the sky subtraction. During the processing of the electronographic images dim areas, remote from the central star, were selected to provide the sky background. However, for such an extensive object as NGC7023 there are difficulties in finding such a region, or regions, within the field. The present difficulties are illustrated by the need to make four separate attempts at sky subtraction, in this case, before the resulting maps were considered satisfactory. Even these are not quite perfect.

An examination of the most easterly and westerly grids in Figs. 2.1, 2.2 and 2.3 and also the most southerly parts of all grids, shows a concentration of random vectors in these areas. Since these are also the lowest intensity areas it is probable that the sky subtraction is not sufficiently accurate in these regions. For the main part however, continuity, across several grids, of the centro-symmetric pattern is obvious, and strong, and gives cause for confidence in the data in those areas. In the central three grids of Fig. 2.1, the pattern is almost perfectly centro-symmetric. Thus from the polarisation pattern too, there can be no doubt that this pattern arises from illumination from a single source - HD200775. The obvious

mechanism is scattering either from spherical particles, or randomly orientated non-spherical particles. Comparisons with other observations indicate that we are within the Mie-scattering domain. The perceived increase in noise with wavelength is also characteristic and to be expected under this interpretation. The red map, Fig. 2.2, is noticibly noiser than Fig. 2.3 the blue image.

The general polarisation behaviour is noted identically on all three figures, thus there is no structural variation, only apparent at one wavelength. There are however, two parts of all three maps where the pattern is not anticipated. First consider an area between east and southeast of HD200775, about 100" to 120" from the star. On all three images there is at least partial breakdown of the centro-symmetry. As a defect on one plate could not be expected to effect the others it must be considered as real. A search for a secondary source of illumination revealed nothing neither is there the pattern that should result from a contribution from such a source. A second prominent star would produce a small circular pattern around its own position with a larger "turn over" area as its influence became less important than that of HD200775, at greater distances. This is not seen. On the other hand we have already noted, plate 2.4, intensity variations, which must be indicative of structural or density fluctuations within the nebula. This partially depolarised region is contiguous with one such structure, therefore the most probable solution is that there is another such inhomogeneity revealed by the polarisation disruption in this region. No other observations to clarify

this or the following point are presented in the literature.

The second region of note is sited from due west of HD200775 to north west, at approximately 90" to 120", from the star. This region differs from that described above. First the behaviour is not generally chaotic, but often rather well ordered. Secondly, although it appears virtually identically on both Figs. 2.1 and 2.3, the red plate, Fig. 2.2 is significantly different. Thirdly, on Fig. 2.1, there is an indication of a breakdown in the vector continuity across grids in the most northerly part of this region.

Moving south to north along the affected grids in Figs. 2.1 and 2.3, the pattern may be described as (a) the normal centro-symmetric behaviour, (b) maginally north of due west, a continuation of the vector sense from (a) but this is now in contradiction to the anticipated circular pattern. Thus the position angle of the polarisation vectors remains almost constant for a distance of approximately two arc minutes. Next, region (c), there is a sharp discontinuity in the vector angles. On all three images the change of direction is in a sense corrective to the centro-symmetry, but only in the red does it seem sufficient. This region is only 30" south to north before we reach (d) a region of chaotic vectors in the red, and a partial reversal to section (b), though much noiser, for the B and VV-filters. On the full VV-map (Fig. 2.1) this area exhibits discontinuity of the vectors across grids, which is a firm indication of error. The form of the error becomes clearer on examination of the corresponding intensity

and polarised intensity images (Figs. 2.4 and 2.5). Fig. 2.5 heightens the contrast between those regions of high and low intensity. The first region indicated above, which had consistent circular behaviour at all wavelengths, clearly shows its nature as a low intensity area. Likewise the second area now discussed shows variations in the intensity which match the shifts in position angle. The vectors in section (c) which are "correct", or almost so, depending on wavelength, are seen to be on a higher intensity prominence than those neighbouring vectors which are chaotic or run counter to centro-symmetry.

The disruptions now appear comprehensible. As we scan through regions (a) to (d) we pass across areas of different intensity. Below a certain threshold value we fall beneath the accuracy of the sky subtraction, thus permitting the imposition of an artificial pattern onto the underlying structure. At a lower intensity still there is no existing order so that the result of the real and false patterns still appears random. In the intermediary area (c), there is a bright limb which is sufficiently bright, in contrast to the sky background to overcome the "pollution" of the field in red, but is not quite able to do so in the blue or the visual, although there is a partial correction of the pattern. As reflection nebulae are generally brighter in the blue than the red, this behaviour seems consistent, since we are proposing an overly intense "sky".

It is important to note that this is the first study with sufficient spatial continuity to be able to detect such a sky background defect. Had any earlier observer

experienced the same difficulties, it is unlikely that he would be aware of it.

Traces across the map are also presented. Fig. 2.6 shows intensity plotted logarithmically against distance. The central blanked region marks the position of the mask covering HD200775. There is obvious consistency in the behaviour at all three wavelengths: a gradual fall off in intensity with distance. The northerly gradient is appreciably more inclined than that to the south. Also presented are scans in polarisation : Fig. 2.7 (blue), 2.8 (visual) and 2.9 (red). Notable features are: the greater error margins marked on the red trace - this is to he expected as the electronographic camera is blue sensitive. Secondly the similarity in the pattern at all wavelenths, we are indeed probing the same structural regions. Thirdly the general form of the curves: a rise and fall on both sides of the star, but not quite a symmetrical one, and fourthly the very similar magnitudes of the percentage polarisations recorded at the different frequencies. These last two points are particularly relevant to any understanding of the structure of the nebula.

3.6 Comparison with other observations

Direct point comparison with earlier data is difficult, for a number of reasons. Some observers have failed to make clear precisely which sections of the nebula they were observing. Additionally a number of observers have used a variety of apperture sizes, and some do not distinguish these observations corresponding to a particular apperture. Yet another problem results from the form of the experimental

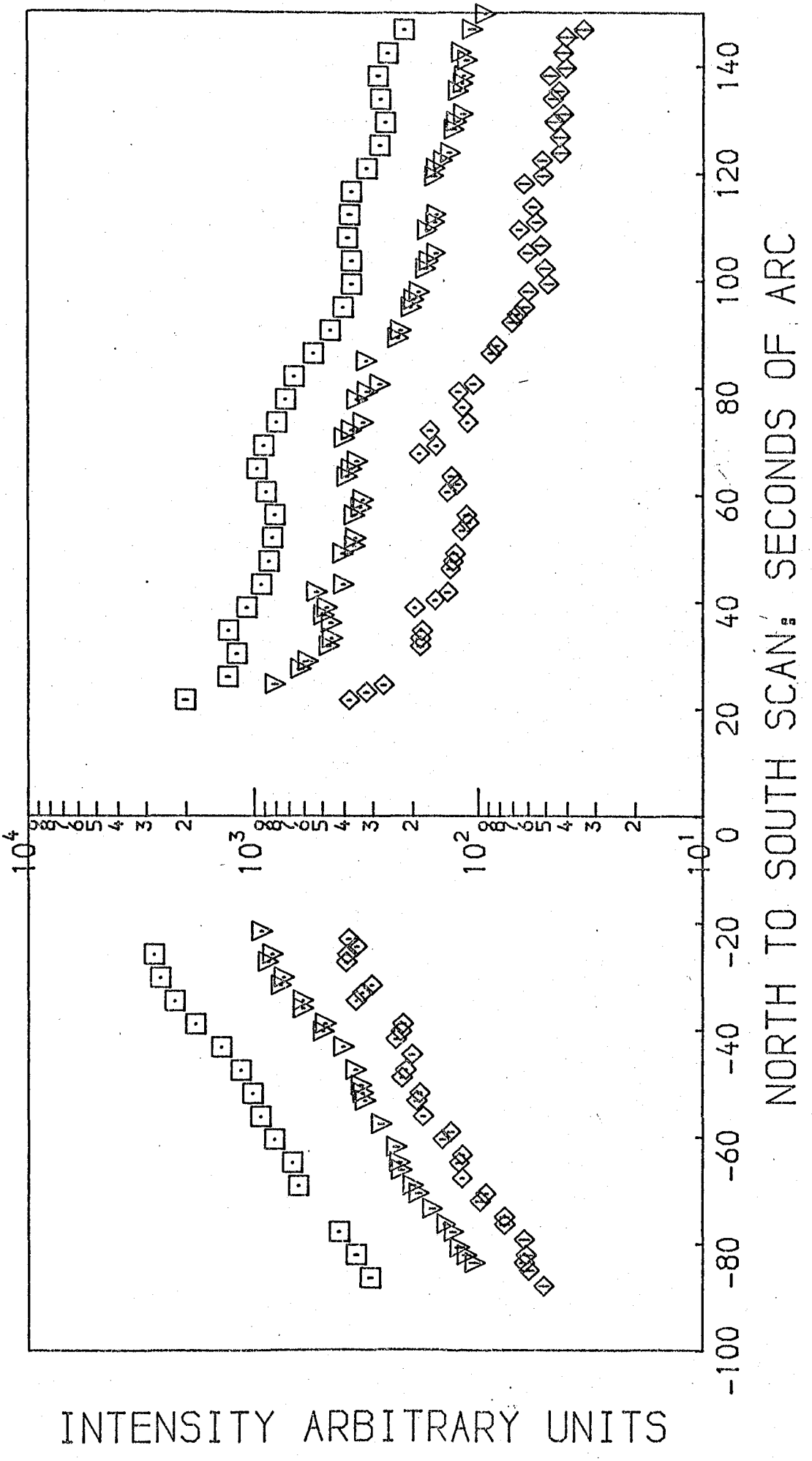

BLUE ▽ VISUAL □ RED ◇
INTENSITY ARBITRARY UNITS
NORTH TO SOUTH SCAN: SECONDS OF ARC
10^4
10^3
10^2
10^1
0
20
40
60
80
100
120
140
-20
-40
-60
-80
-100

Fig. 2.6

Fig. 2.7 NGC7023 B-FILTER

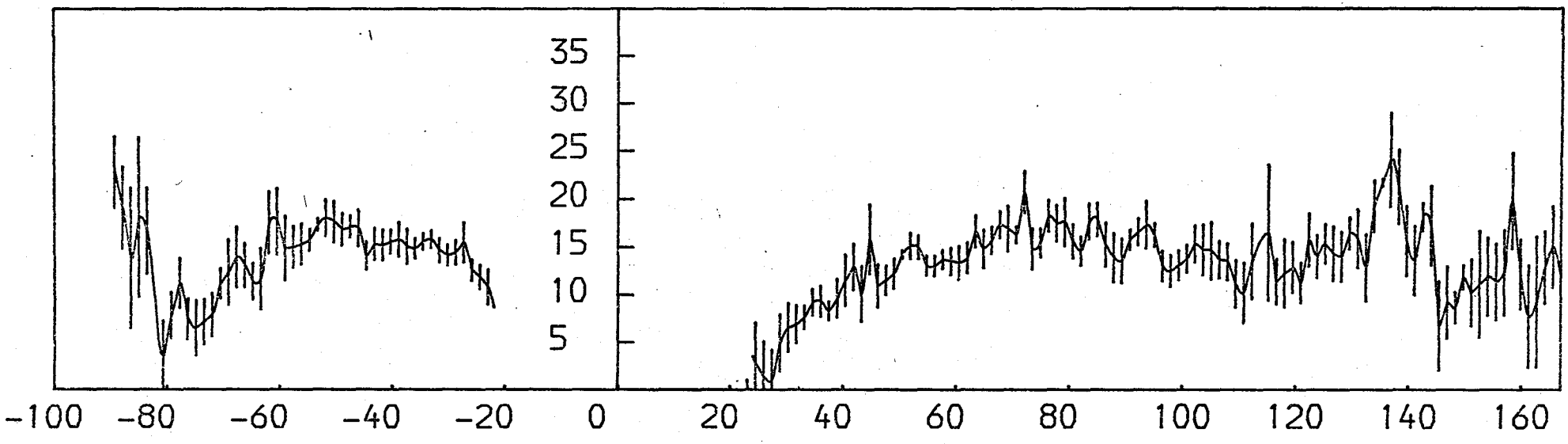


Fig. 2.8 NGC7023 VV-FILTER

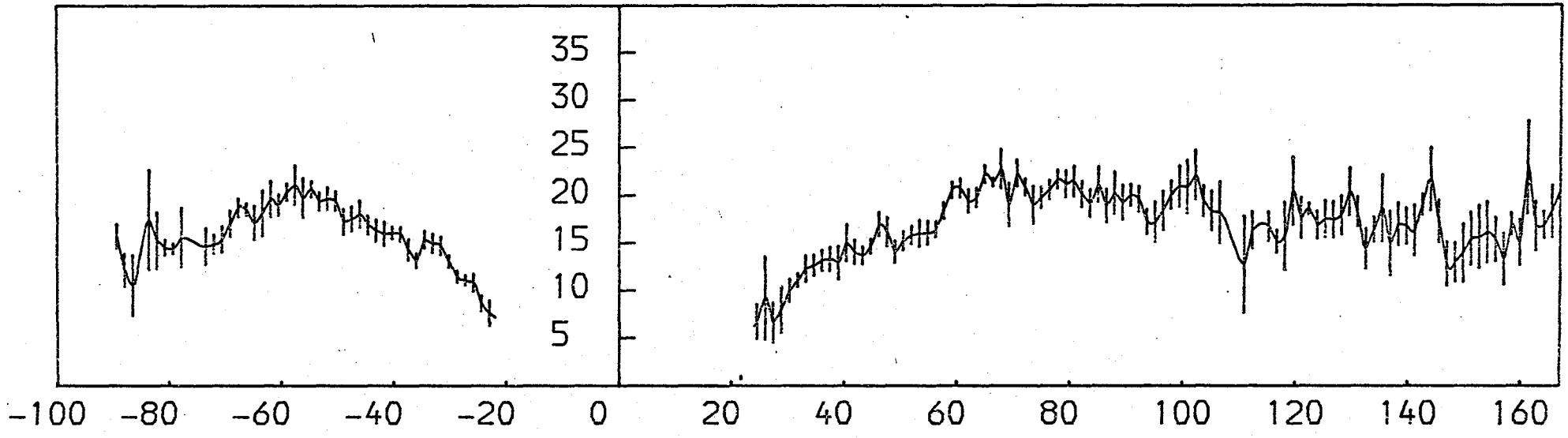
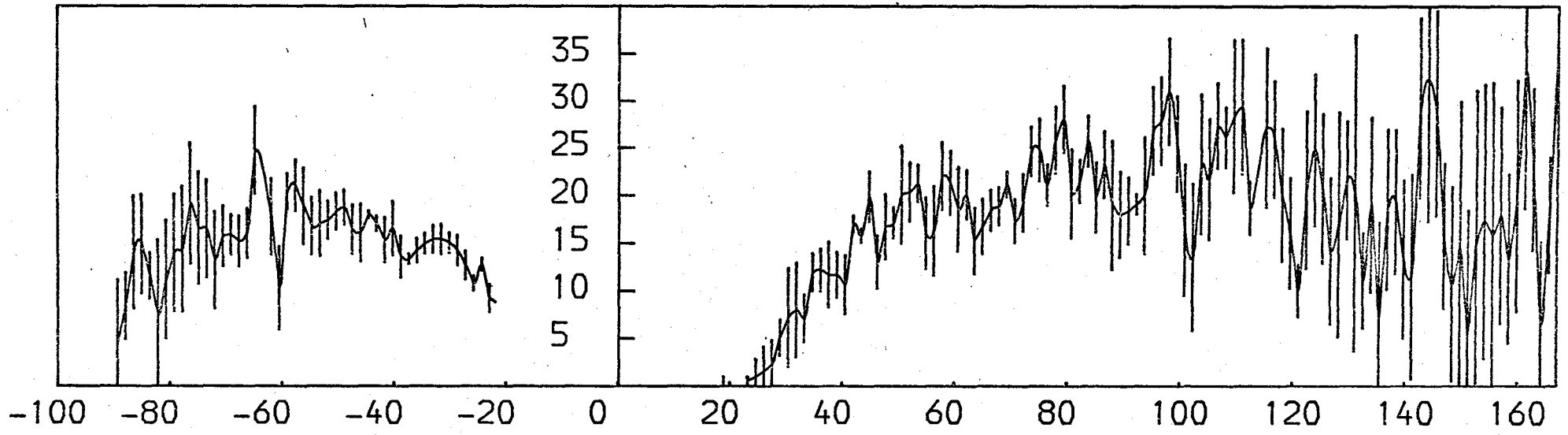


Fig. 2.9 NGC7023 R-FILTER



technique presented here. It is an inevitable consequence of observing in "strips" that some data will be lost on the grid boundaries. Since some previous observers have used very large appertures it is the ineluctable result that some portion of the area included in their measurements should fall on one of the sections between grids thus making comparison impossible.

Still comparisons and contrasts can be made with the results following. First there is no evidence at all for the very high polarisations reported by Gliese and Walter (1951) - up to 56% - within 3 arcminutes of the star. Rather the data agree broadly with the best modern observations by Elvius and Hall (1966), Gehrels (1967), Zellner (1970) and Ney et al. (1980), that indicate polarisations generally between 5% and 25%, but with notable local variations across the nebula. The strong centro-symmetry first reported by Martel (1958) is very obviously confirmed, as is the local variation from it. Gehrels and Zellner had both observed in the same two positions; neither of their offsets could be reconstructed since they used large appertures, up to 60", which extend over grid boundaries into our blank regions. However, Elvius and Hall made considerably more observations employing appertures typically of 29" diameter and so it was possible to calculate the equivalent polarisation measurements for a number of regions. These observations, with the new equivalent observations, formed by a summation of our picture elements to construct an equally sized diaphragm are reproduced in Table 3. Points 67" south and 186" east display excellent agreement, the others rather

less so. However it was noted that the mean difference between the two sets of measurements was rather less than that between Elvius and Hall and either Gehrels or Zellner. N.B. Gehrels and Zellner produced very similar results, but they were using the same equipment and experimental techniques; it would be incorrect to consider them as truly independent observers.

Both sets of observations underline the variability in the structure of the nebula with offset distance. The most probable source of error is the sky subtraction - inconsistency in the sky subtraction is an unavoidable consequence of the technique used by Elvius and Hall. No portion of the area of the new data which is thought to suffer from sky subtraction problems was used in comparison. It could not have been, since no other observer has reported measurements in that region.

3.7 Conclusions

The present work confirms that HD200775 is the only important light source in the nebula NGC7023. There is a general fall off in intensity with distance from the star; this is slightly asymmetrical north-south. Other irregular variations in brightness are noted, some of which correspond to changes in the polarisation pattern. This last is fundamentally centro-symmetric, displaying a rise and fall with offset from HD200775 at all wavelengths: the form and magnitude being very similar for all the wavelengths observed. There is a marginal tendency toward higher polarisations at longer wavelengths, with evidence that this behaviour is not uniform across the nebula. At distances remote from

HD200775 it appears most pronounced. Broad agreement with previous observations is indicated.

Table 3 Percentage polarisation variation with offset and wavelength:29" apperture, or equivalent.

Offset Distance Arcseconds	Elvius & Hall 0.375μ	New Data 0.44μ	Elvius & Hall 0.45μ	New Data 0.52μ	Elvius & Hall 0.56μ	New Data 0.65μ
33"S	4.2	6.75	8.2	11.2	12.9	9.92
67"S	11.2	15.35	15.8	19.1	19.1	18.6
133"S	3.5	14.1	9.0	17.1	13.3	17.9
152"SSE	9.9	14.3	8.9	16.9	10.1	
67"E	4.3	3.7	6.3	7.2	11.0	9.1
186"E	8.1	14.8	13.1	17.85	20.3	21.0
94"NE	14.0	14.0	3.9	15.8	5.8	16.8

CHAPTER FOUR

What is this quintessence of dust? -

Hamlet

DUST

4.1 Introduction

The form, composition, evolution and structure of dust in the interstellar medium, or in the vicinity of stellar associations, is one of the most complex and intractible of all the active problems in astronomy. Few investigations have been so catholic in the range of observational or theoretical techniques applied.

It is essential to consider the interstellar extinction and polarisation, cosmic abundances, the evolution of massive stars (viz. supernovae) and interstellar shocks; the effects of chemical and physical processes in domains so varied as HII regions and cold interstellar clouds. Transitory or transitional effects such as the condensation or evaporation of grain mantles, or surface chemical reactions cannot be ignored. Caution must be applied to the consideration of the effects of grain evolution - formation, growth, fragmentation and destruction -, temperature, the possible restrictions of: the diffuse interstellar absorption lines, size distributions, particle type and chemical composition or the surface texture. In all a torrent of research: often confirmatory, sometimes contradictory, perhaps confused by coincidental similarities, constantly evolving.

The field is so wide-ranging and inter-connected that it is scarcely possible to present a clear modular exposition. The sections following will attempt to draw out the most important constraints for the present investigation.

4.2 Cosmic Abundances

The broad uniformity of certain features of the interstellar dust, important examples being the interstellar extinction law (Johnson, 1968) and Serkowski's relationship between polarisation and wavelength (normalised by the wavelength of maximum polarisation) (Serkowski, 1971), indicate that a consideration of cosmic abundance ratios could place significant restrictions on the general composition of the dust. Even in apparently anomalous regions, e.g. the Rho Ophiuchi dark cloud (Carrasco et al., 1973) these considerations should still apply.

Cosmic abundances are obtained from studies of solar system abundances and from thermally unprocessed chondritic meteorites (Cameron, 1973; Knacke, 1980). Elements other than hydrogen and helium total approximately two percent by mass, with respect to hydrogen. A significant fraction of this percentage is expected to comprise the grains. Table 4.1, selected from Greenberg (1968), lists the most important elements (normalised to the hydrogen fraction) in terms of the relative numbers of atoms.

This, the most crude, limitation on grain composition can be refined by a consideration of the Kramers-Kronig dispersion relations (Purcell, 1969; Carroff et al., 1973), connecting the real and imaginary parts of the dielectric susceptibility, as an integral over frequency:

$$\chi'(w_o) = \frac{2}{\pi} \int_o^\infty \frac{w\chi''(w)dw}{w^2 - w_o^2} \qquad (4.1)$$

Table 4.1 Cosmic Abundances : Common Elements

Element	Relative Number
H	1
He	1.20×10^{-1}
C	3.70×10^{-4}
N	1.17×10^{-4}
O	6.76×10^{-4}
Mg	0.34×10^{-4}
Si	0.32×10^{-4}
Fe	0.26×10^{-4}

Chiao et al. (1973) have integrated over the interstellar extinction curve from the infra-red to the ultraviolet, in the special case of $w_o = o$, to restrict the elemental composition of grains. The extinction A (magnitudes cm^{-1}) is given by:

$$0.921\ A(w) = \frac{4\pi}{c} w\ \chi''(w) \qquad (4.2)$$

The calculations are but slightly dependent on particle shape and thus the resulting dust to hydrogen density ratios, r_d/r_h, are fairly insensitive to grain composition:

$$r_d/r_h = 3.0 \times 10^{-3} g \int_o^\infty \frac{A(w)dw}{A_v w^2} \qquad (4.3)$$

where g is a factor dependent on grain density and shape.

Table 4.2, adapted from Martin (1978), tabulates the required elemental abundance to cosmic abundance ratios for

certain chemical species.

Table 4.2 Restrictions on Grain Composition

Material	Formula	rd/rh	O	C	Si	(MgFe)	Fe
Dirty Ice	$O_{0.58}C_{0.32}N_{0.1}H_{2.8}$	5.8×10^{-3}	0.27	0.27			
Orthopyroxene	$(Mg,Fe)SiO_3$	1.1×10^{-2}	0.42		2.9	3.1	
Olivine	$(Mg,Fe)SiO_4$	1.0×10^{-2}	0.34		1.8	3.9	
Magnetite	Fe_3O_4	8.8×10^{-3}	0.23				4.4
Iron	Fe	9.3×10^{-3}					6.4
Silicon Carbide	SiC	4.7×10^{-3}		0.33	3.9		
Graphite	C	2.7×10^{-3}		0.61			

The important point to note from these studies is that iron and silicon, strongly advocated on other grounds, see below, are underabundant.

4.3 Extinction

The single most investigated aspect of interstellar grains is without doubt the interstellar extinction. Defined as the difference in magnitude between a reddened star and an equivalent unreddened star, at the same wavelength:

$$A(\lambda^{-1}) = \{m(\lambda) - m(\lambda_V)\}/\{m(\lambda_B) - m(\lambda_V)\} \qquad (4.4)$$

the normalised extinction has been observed from the far infra-red to the ultraviolet. Figure 4.1 shows a composite smoothed curve, in these regions, after Schultz and Wiemar (1975), Nandy et al. (1975) and Yorke et al. (1973), each group examining the I.R. optical and ultraviolet spectrum respectively. Extinction curves have been measured for many stars of varying spectral type, in different regions of the

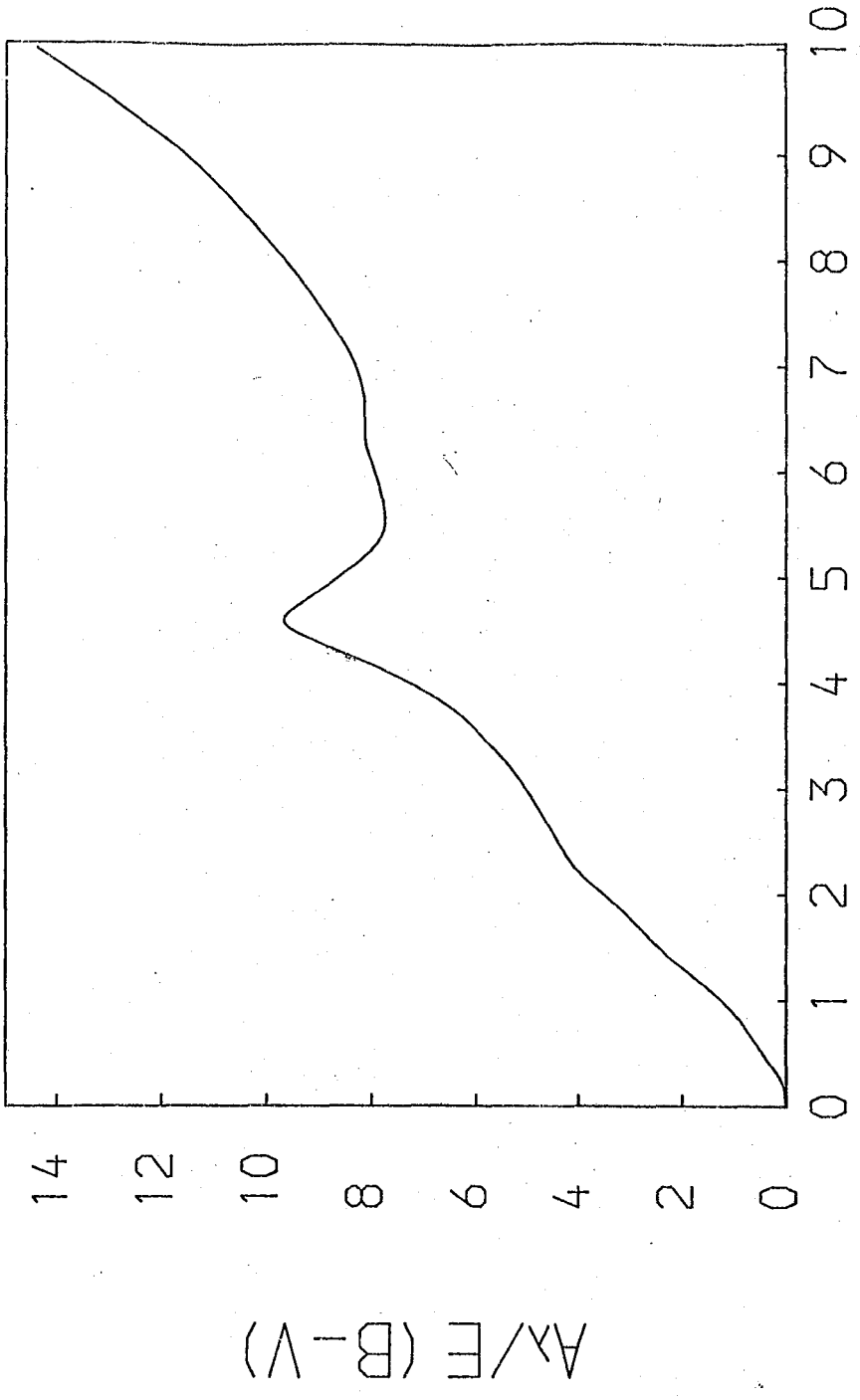

INTERSTELLAR EXTINCTION
$A_\lambda/E\ (B-V)$
0
2
4
6
8
10
12
14
0 1 2 3 4 5 6 7 8 9 10
$1/\lambda\ (\mu m^{-1})$

Fig. 4.1

of the galaxy. There is an overwhelming uniformity, notable exceptions being the U.V. behaviour of θ Ori and ξ Ophi (Bless and Savage, 1972). There is no systematic galactic variation, but Savage (1973) has noted that five out of six of the OAO-2 stars that displayed abnormal ultraviolet extinction were associated with nebulosity - the probability of specialised processing in these regions is high. The extinction is characterised by three notable features: first the very prominent peak at 4.6 μm^{-1} (2200Å), secondly a general increase for the far U.V. and thirdly, rather less marked, a "knee" at 2.3 μm^{-1}. Attempts have been made to understand these features in terms of a single chemical composition, and for uniquely sized grains; the conception is now that such models are too simplistic, and multi-component dust particles with a distribution of sizes, dominated perhaps by different elements at varying wavelengths, are now considered to be a more accurate summary of the position.

4.3.1 The Infra-red Regime

Extinction measurements in the infrared are plagued by large intrinsic errors. The major source of difficulty is in the stellar comparison - many stars emit in the I.R. from circumstellar dust or by free-free processes.

Investigations concentrate on a search for band structures. The earliest of these was that for the 3.1 μm feature of the O-H stretching mode in water ice (Danielson et al. 1965, Knacke et al.,1969a) which was thought to deposit by accretion in the interstellar medium. From comparison with laboratory spectra Knacke et al. (1969b) were able to place an upper limit of 10% on the icy component

of grains. Subsequently, however, the ice band has been discovered in dense molecular clouds (Merrill et al., 1976). Gillet et al. (1975) have discovered a significant variation in the strength of the ice feature, between that in the BN source in Orion and VI Cyg. No. 12. It seems very likely that the ice is being shielded in the former.

Silicon Carbide has been identified as the source of the 11.5 μm emission in the circumstellar emission features of carbon stars (Gilra, 1973), while much of the broadband emission in the 8 to 13 μm range is attributed to graphite (Martin, 1978). However, by far the most significant identification in the region is the 9.7 μm "silicate" feature, found both in absorption and emission spectra, (Merrill and Stein, 1976). If this structure is due to silicates (it is believed to result from the Si-O stretching mode) then a second feature should be detectable at about 20 μm, due to the bending vibrations. There is. At this point the problem becomes more complicated. The 9.7 μm feature is found in all silicates, but the particular variety should be deducible from the precise wavelength and its profile, which could be expected to be a function of the crystaline (or other) structure. Both small particle transmission studies and Kramers-Kronig, i.e. bulk, analyses have been performed. A considerable number of candidate materials have been proposed. Terrestrial silicates examined include: dunnite, serpentine, chlorite, murchison, vigarano, cold bokkveld, murray, orgueil and ngoya (Zaikowski and Knacke, 1975; Penman, 1976; Day, 1974; Friedmann et al., 1979). Both moon rocks and meteoritic silicates have also been

tested. All these materials suffer from one of two defects, either the peaks are slightly misplaced or they exhibit a more detailed structure than that found in the interstellar observations. This has led to a very wide investigation of more esoteric silicate materials. Day (1974, 1976) has prepared amorphous magnesium silicates in the laboratory and Kratschner and Huffman (1979) amorphous olivine. Duley and McCullough (1977) have examined some amorphous iron silicates, while Zaikowski et al. (975) have investigated layered hydrous silicates - phyllosilicates. These last are very similar to the hydrous silicates found in the chondricic, Cl (largely unprocessed thermally) meteorites studied by Knacke (1980).

Millar and Duley (1978) have pointed out that mixtures of silicon and magnesium oxides also display similar characteristics at the same wavelengths. These are perhaps the only serious threat to the silicate identification.

As an aside, it has been remarked by Wickramasinghe et al. (1977) that a number of organic compounds including various polymers and polysacharides could account for the bands. Knacke (1980) has indicated the probable coincidental nature of this observation, since these features occur in the spectra of oxygen rich stars, where organic molecules are not observed, but the 9.7 micron band does not appear in the spectra of carbon stars where such organic compounds might be expected to be present.

The general conclusion is that a mixture of amorphous or hydrated silicates can explain this feature well. Crystaline forms appear to be precluded.

4.3.2 The Visible Region

The visible region has two notable features: the approximately linear behaviour of the extinction with inverse wavelength - this is to be anticipated for any material so long as the grains are small (Wickramasinghe, 1963). The implication is λ^{-4} scattering, with grains roughly comparable in size to the wavelength. The second aspect is the change in gradient at about 2.3 μm^{-1}. The cause of this variation is unknown. Hayes et al. (1973) propose a suitable size distribution of grains, while Greenberg (1978) suggests a variation in particle compositon which is masked by the 2175Å ultraviolet hump, but which tends to obscure the underlying change in reddening at optical wavelengths.

4.3.3 The Ultraviolet

Early U.V. measurements were collected from rocket and balloon measurements (Stecker, 1969; Bless et al., 1968; Carruthers, 1970, 1971; Navach and Lehman, 1971) and now a number of orbiting satellites have produced more accurate measurements: OAO-2 (Bless and Savage, 1972; Savage, 1975; Code et al., 1976), Copernicus (Yorke et al., 1973), TD-1 (Nandy et al., 1975) and the ANS (van Duinen et al., 1976).

Although there are marked variations for some stars, especially in the far ultraviolet, the overall pattern is clear. There is a very prominent fairly symmetrical peak at 2175Å ± 30Å (i.e. 4.6 μm^{-1}, see Figure 4.1); the full width at half maximum has been variously estimated as 480Å (Savage 1975) and as 360Å (Nandy et al., 1975), the precise value depending on the baseline chosen. This feature is well correlated with E(B-V) indicating either that the inter-

stellar grains giving rise to it are the same as those involved in optical extinction, or that they are well mixed with the component important in the optical (Dorschner et al., 1977). Following the 2175Å hump is a shallow depression from approximately 5.5 μm^{-1} to 7.5 μm^{-1}, after which there is a general rise from this minimum towards shorter wavelengths. The explanation of the 2175Å peak is eagerly sought. Some of the earlier suggestions included absorption edges of silicates, absorption by hydrocarbons and even colour centres in quartz (Martin, 1978). These have now been abandoned, the most favoured contender being graphite, or mixtures containing it. It seems probable that this band is due to double bonds of carbon involving π-electrons which are strongly coupled to the lattice (Andriesse and de Vries, 1974). Excitation of the coupled π-electrons induces absorption and infrared emission. More circumstantial evidence is provided by carbon depletion (Field, 1974); graphite could form in the atmospheres of late-type carbon stars and then be expelled into the interstellar medium. Further support for the graphite identification comes paradoxically from its weakness in certain stars. Carrasco et al. (1973) have shown that some grains in the Rho Oph cloud have accreted mantles, probably ice, which shield and therefore weaken the 2175Å hump. Gilra (1972) has indicated that even a thin coating would be sufficient to drastically reduce the size of the peak, so the graphite identification appears strong.

Further conformation comes from the Mie calculations of

Wickramasinghe and Nandy (1971) who use a graphite-iron-silicate mixture to fit the U.V. extinction. Caution should be applied to their calculations, which advocated small uncoated particles, since they used constant values for the refractive indices of iron and silicates, which does not seem justified in this region. Subsequent calculations indicate that variations in shape, orientation and size distributions are permissible (Mathis et al., 1977), so long as there is a "topping up" of smaller graphite particles to reduce the steepness of the peak (Hong and Greenberg, 1978).

Despite the strength of the above arguments in favour of graphite, the position is not yet secure. Millar and Duley (1980) argue that small (50Å) particles of certain metalic oxides have surface exitons close to the peak, which might be shifted to coincidence by suitable mixtures and through a size distribution.

The far ultraviolet rise is generally attributed to small particles (Witt, 1979), the identification of which is very insecure, but Andriesse and de Vries (1973) have suggested that Platt particles - interstellar free radicals, basically unsaturated molecules, of roughly 10Å - could also explain the behaviour.

4.3.4 The Variations in Extinction

It has already been noted that Savage (1972) had found anomalous near ultraviolet extinction associated, in five out of six cases, with nebulosity. Clearly this is very relevant to considerations of the applicability of the general results to the particular case under study, since NGC 7023 is a

reflection nebula. It is also a point which seems to have been largely overlooked by Witt (1979) in his investigation of the size distribution in reflection nebula, in the ultraviolet. Local variations in extinction characteristics must always be recalled. The best example is the behaviour in the central regions of the Orion nebula where Lee (1968) found values of the extinction ratio as high as 5.5, almost twice the normal value. Generally we must be mindful of three modifying processes: growth (e.g. by accretion in dark clouds), modified composition and variations in sizes (e.g. by sputtering in shocks on HII regions, or by ageing).

4.4 The Interstellar Polarisation

The discovery of the interstellar polarisation by Hall and Hiltner (1949) has identified large-scale, galactic alignment, of elongated particles. Numerous mechanisms have been propared to explain it, e.g. Davis-Greenstein and the Purcell pin wheel effects.

The linear polarisation, measured for several hundred stars, have led to the construction of an empirical relationship between the maximum polarisation and the wavelength at which it occurs (Serkowski, 1971; Coyne et al., 1974):

$$P(\lambda)/P_{max} = \exp(-1.15\ \ln^2\ (\lambda/\lambda_{max})) \qquad (4.5)$$

This function is remarkably general, indicating that the optical constants of the material giving rise to the polarisation do not change markedly, as a function of wavelength, in the visual and the near infra-red (Martin, 1974). This behaviour has been modelled using the Mie theory applicable

to dielectric cylinders of "dirty ice" in the size distribution due to Oort, van de Hulst and Greenberg (c.f. 4.7 below), by Coyne et al. (1974), and by Gehrels for pure and dirty ices, silicates, metalic grains and graphite. Interestingly, of these graphite was the least acceptable. Shapiro (1975) has conducted a similar anslysis for platelike magnetitite (Fe_3O_4) grains, aligned by Purcell's mechanism (1974), but only for single sized distributions. Mathis has also modelled the polarisation for several materials using a power law distribution. The interstellar linear polarisation seems an undiscriminating measurement, rather the latitude to identification has become almost baroque. It is fundamentally limited in that essentially only one geometry, illumation at zero degrees, is available for examination. Fortunately this is a hanidcap which is not present in studies of reflection nebulae.

Two other points of note are the connection between the extinction ratio and the frequency of maximum polarisation (Serkowski, 1975):

$$R = (5.8 \pm 0.4)\ \lambda_{max} \qquad (4.6)$$

where λ_{max} is in microns, this is of obvious application in extinction studies. Secondly the wavelength of maximum polarisation is generally very close to 5450Å, but there are some regions where it is significantly different: Orion and Ophiuchus, dense clouds with larger grain sizes, have higher λ_{max}. While the Cygnus OB2 association has $\lambda_{max} = 0.41$ μm (Savage and Mathis, 1979).

Circular polarisation is also detected, though as it is a second order effect, following as a consequence to linear

polarisation, it is much weaker. Martin (1974) has shown that for a dielectric material the circular polarisation will go through zero at the wavelength maximum for linear polarisation. This is not so for metallic solids. As it is the former behaviour that is observed (Kemp and Wolstencroft, 1972; Mathis et al., 1972); it is now possible to exclude metals as the major responsible agents for polarisation, and also graphite (Martin and Angel, 1976).

4.5 Formation, Modification and Destruction of Grains

The formation of grains by the direct condensation of nuclei form the atoms, or molecules, of the interstellar gas is so slow a process that efforts have been turned towards other mechanisms. The perceived wisdom (Wickramasinghe and Nandy, 1972; Aannestad and Purcell, 1973) is that nuclei are formed in other denser environments and are then expelled into interstellar space. Three radically different processes have been suggested, that are still supported. Hoyle and Wickramasinghe (1962) have proposed formation of grains, or grain cores, in the atmospheres of cool stars, followed by a subsequent expulsion by radiation pressure. Later they advanced a second mechanism (Hoyle and Wickramasinghe, 1970), formation in supernovae. More recently Burke and Silk (1976) have suggested a process which includes both the formation and modification of grains in massive photostellar shells. Grains may be further modified by the accretion of mantles (Barlow, 1978c) or they may be totally destroyed in interstellar shocks or by erosion or chemisputtering (Draine, 1979). Ironically, the destructive interstellar shocks include those produced by supernova remnants (Barlow, 1978a).

4.5.1 Formation

First consider formation by condensation in cool stellar atmospheres. This proceeds by two different pathways depending on the nature of the stars. In carbon stars, those with a carbon to oxygen ratio greater than one and surface temperatures in the region 1500K to 3000K, most of our ideas of chemical kinetics may be considered to apply (Donn et al., 1968). In fact some modifications are required (Salpeter, 1974a) since classical nucleation theory was derived from considerations of liquid droplets maintained by weak polar forces, whereas in stellar atmospheres we must consider small crystals held together by valence bonding, with the result that the condensation process for carbon will be dominated by acetylene $C_2 H_2$, as a mediator, rather than pure carbon. These stars cycle through pulsations of period, approximately, one year, during which time the temperature will range through the span indicated above. In the expansion phase grains may be expelled under radiation pressure and escape (Tabek, Hirth, Heyrick and Roark, 1975). The carbon fraction of the atmosphere will not be significantly depleted as it can be replenished from the stellar core.

A similar scheme operates for giant oxygen-rich stars (type M), these have a C/O ratio less than one. In this case the primary condensate is believed to be a mixture of silicate materials (Gilman, 1969), primarily magnesium silicate, deposited originally in an onion layer fashion, though such a structure is unlikely to persist. Salpeter has suggested a complex "servomechanism" connecting grain size and opacity

with velocity and temperature and thus nucleation rates. He believes that such a process might lead to fairly large silicate grains, of the order of one micron.

The other "pure" formation process is that involving supernovae explosions, or less dramatically, but still qualatively the same, stars undergoing spontaneous mass loss, not connected to grains. From the abundances of Arnott (1969) Wickramasinghe and Nandy (1972) argue that as much as 10% of the ejected matter could be condensible material, and that this source alone could explain most of the dust in the galaxy. Observational evidence for such a mechanism, in type II supernovae, has been provided by Gallagher (1977) who has examined fifteen explosions and one, Nova FH Ser 1970, in substantial detail. He concludes that the diminution in light output 50 to 70 days after the outburst can be explained by the formation of dust, which absorbs the optical radiation to reradiate at longer wavelengths. More precisely he gives the timescale, in, days, for the obscuration by dust in these supernovae as:

$$t_{dust} = (320/V)\ (L/L_{\odot})^{\frac{1}{2}} \qquad (4.7)$$

where v in kms^{-1} is the speed of expansion and L the luminosity of the supernova. Falk, Lattimer and Margolis (1977) have conducted a similar analysis from which they deduced that the major factor influencing grain growth was monomer depletion and not reduced density, and further that the resulting size distributions are relatively flat, being centred about 0.1μm to 0.3μm. The importance of these processes, in the present study, lies in their effects on grain composition and the size spectra.

4.5.2 Modification and Destruction

The processes which modify grains are chiefly those which destroy them and so it is difficult to consider these effects in a compartementalised manner. The important factors are evaporation (perhaps by chemical agency or collision), sputtering, shattering and photo desorption.

One of the most interesting developments in the modification of grains results from an extension of Larson's (1969, 1972) models for protostellar collapse (Burke and Silk, 1976). They consider the formation of a central protostellar core from an interstellar cloud. On account of the vast increase in density grains embedded in the gas; falling in towards the proto-star, will grow by accretion. The cloud collapse proceeds until a circumstellar shell is formed, supported on the inside by radiation pressure and compressed from without by the ram pressure of the infall. For protostars exceeding approximately five solar masses the radiation pressure is sufficient to expell the shell containing the grains, which will become unstable and break up. For our purposes only those processes affecting the grain size distributions are important. Grain kinetic energies and collision times are sufficient to play a dominant role in the resulting size distributions, through the mechanism of shattering. As shattering is also relevant in the "equilibrium" conditions for the interstellar medium considered by Oort and van de Hulst (1946) we shall put off a detailed examination to section 4.7, and merely state the conclusions here. Investigations of shattering by Hartman (1969) has shown that the resultant size spectrum tends towards a power law. This is

expected to be the case here also, with the implication that the total number of grains can be increased for the same given mass of grains. Even more importantly the new size spectrum is found to be bimodal and thus a possible theoretical explanation of the observations by Witt (1973) in the far ultraviolet.

In recent time the action of sputtering as a prominent destruction mechanism has been hotly debated. Defined as a process by which lattice particles in a solid may be ejected by atomic or ionic bombardment, the precise effects depend crucially on the collisional energies and the chemical species involved. Three distinct regimes exist: the hard sphere collision, the screened Coulombic and the pure Coulombic. The first investigators seem to have been Burke and Silk (1974a) who examined the processing of graphite in a hot gas, deriving particle lifetimes. Barlow (1978a) went much further, using sputtering yield estimates taken from ion bombardment experiments by Stuart and Wehner (1962), Cuderman and Brady (1968) and Kenknight and Wehner (1964) to investigate interstellar shocks from cloud-cloud collisions and from supernovae remnants. From these he concluded that the lifetimes for icy grains would be 2×10^8 yrs and for refractory grains 5 to 20×10^8 yrs. This gives a pre-eminence to sputtering as the principle destruction mechanism for grains, especially as Barlow (1978b) has gone on to show that the classical cloud-collisional destruction process of Oort and van de Hulst is ineffective, since detailed studies of the shock interface indicate that only on the interface itself will the relative velocities of grains be sufficient

for collisional disruption. Barlow does not consider shattering in this analysis so that it may be rather more complex, nonetheless from Barlow's scenario a grain lifetime to shock destruction, by cloud-cloud interaction of 1.4×10^{10} yrs is developed. This is about the age of the Galaxy and is much longer than other destruction processes.

Chemical reactions on grain surfaces have been considered as destruction agents (Bar-Nun, 1975). The consensus is that they are ineffective except for the desctruction of graphite by chemisorbed hydrogen and oxygen, for grain temperatures greater than 65K (Barlow, 1978b; Barlow and Silk, 1977). Barlow proposes this mechanism as the explanation for the weakness of the 2175Å feature in a number of objects. Draine (1979) has reconsidered the reaction yields of the destruction processes and concludes that they are much lower than the values adopted for the earlier calculations. The new values render chemical destruction of graphite unimportant save in extremely compact HII regions.

Finally photodesorption - the process by which an absorbed particle is desorbed from a surface by a quantum process - has also been considered. Barlow has calculated the lifetime of a 0.1 μm ice grain against destruction in the average stellar radiation field to be 5×10^4 yrs. Icy grains are no longer expected. For other materials, most particularly metals, photodesorption is not effective.

4.5.3 Mantles

The accretion of volatile coatings to form mantles around existing cores was first proposed by Wickramasinghe (1965), who estimated the deposition of ice onto graphite cores. He concluded that the formation of a mono-layer was

relatively easy, proceeding through the chemisorbtion of oxygen, but after the satisfaction of the surface valence bonds that further accretion, by van de Waal's attraction was much less efficient requiring very low temperatures - less than 5K. Impurities however might alter some of these considerations. The observed weakness of the ice band at 3.1 μm seems to confirm Wickramasinghe's ideas. This is contrary to the view prevalent in the literature that he was a strong proponent of ice mantles.

Wickramasinghe and Nandy (1968) have examined the extinction curves for graphite coated by solid diatomic hydrogen and Krishna Swamy and Wickramasinghe (1968) quartz covered by dirty ice. Neither showed agreement with the data. In confirmation of the lack of H_2 mantles, Graham and Duley (1971) have found that spectral variations in the ultraviolet were absent. The universal opinion is that any H_2 formed will be rapidly expelled from the grains (Duley, 1970). Barlow (1978c) has gone further to conclude that in unshielded regions all non-metal atoms will be ejected from silicate grains as mono-hydrides and from graphite and iron as saturated hydrides, due to photodesorption and by recombination desorption. After expulsion the hydrides should dissociate rapidly by the action of the interstellar ultraviolet, followed by a repetition of the entire cycle. This continuous process of hydride formation, ejection and destruction should be sufficient to prevent the formation of such mantles in unshielded regions. For metals, the position is more complex; it seems that some metals could experience direct depletion from the medium. Another process

relevant to grain size distributions is the inhibition of mantle growth for very small grains - 0.005 μm - due to rapid transitory fluctuations in the grain temperature caused by the absorption of a series of single photons (Purcell, 1976), but the importance is equivocable.

In shielded regions saturated molecules can exist in free space, thus exothermic molecular recombination is no longer an inhibitory factor and photodesorption will also be reduced, with the result that mantles can form. Direct evidence for coatings has been provided by a number of observers: Strom et al. (1975) who noted increases in R, the ratio of total to selective extinction, and an increase in the wavelength of the linear polarisation maximum; Carrasco et al. (1973) detected R variations in ρ Oph; Morton (1974) found elemental depletion with respect to cosmic abundances, from ultraviolet (Copernicus) observations towards ξ Oph.

Thus it can be fairly concluded that mantles will form so long as they are adequately protected. However upon re-exposure to direct U.V. they will be rapidly evaporated due to photo-desorption. This process is extremely rapid: an ice grain of radius 0.1 μm would be destroyed in 5×10^4 yrs by the mean interstellar radiation field (Barlow 1978c) or in an HII region in only 130 yrs, at 2 parsecs from an O star.

4.6 Dust Composition : A Summary

A generous selection of materials have been proposed as candidates for the substance of the interstellar dust: metals, ices, silicates, silicon carbide, organic compounds,

graphite, quartz, metallic oxides, etc. In the following chapter the more prominent contenders are examined to test for their occurrence in NGC7023. A brief resume is given of the evidence leading to their inclusion. In all cases two considerations are borne in mind: cosmic abundance ratios and a desire to span the range of possible refractive indices. The latter is the physical parameter characteristic of each grain species as input into the Mie calculations. So the distinctness of the index of refraction for some element compound is the first vital factor in identifying it uniquely.

4.6.1 Ice

Ices, mixtures principally of water ice, frozen methane and ammonia, refractive index 1.33 (Greenberg, 1968) were originally proposed by Lindblad in 1935. The belief was that direct condensation from an accumulation of diatomic molecules, through condensation nuclei, to grains, was possible. Such a hypothesis is now discredited, on the probabilistic grounds of extremely slow reaction rates. Furthermore, although fits to the visible extinction have been produced they fail for both the U.V. extinction hump and for the far ultraviolet (Krishna Swamy and O'Dell, 1967). However, the possibility of grain mantles in shielded regions is still considered highly probable.

4.6.2 Graphite

Graphite, refractive index 2.5 - 1.3i (Wickramasinghe, 1973) is strongly supported for two reasons. First there is an obvious viable formation mechanism, in cool stellar atmospheres; secondly the presence of the strong absorption

peak in the ultraviolet at 2175Å. This identification believed to be caused by small carbon particles, (0.02 μm) (Gilra, 1972) may not be entirely secure since the plasma oscillations of the electrons are not independent of size and shape. Thus variations of the parameters might be able to shift the peak significantly. However, it is undoubtedly true that graphite is by far the best known candidate for this feature.

4.6.3 Silicates

Silicates, refractive index 1.63 - 0.05i (Hanner, 1969) are also considered probable, on similar grounds to those for graphite. As before there is an obvious formation mechanism, again nucleation in cool stellar atmospheres, M-type stars in this case, followed by expulsion. Many silicates have been found in chondritic meteorites and most importantly there are the infra-red features at 9.8 μm and 20 μm. Once again there are alternatives, principally metallic oxides to explain the 9.8 μm peak, but silicates are clearly the strongest contender to explain this feature.

4.6.4 Iron

Metallic grains of which iron is the most abundant were advocated early by Schalen, 1936, to explain the interstellar extinction. Although it is now regarded as underabundant to perform this task, it could still condense onto existing silicate cores in the later stages of outflow from stars, as cooling proceeds. It is also possible that originally totally dielectric grains may have their metallic content enriched by the vaporisation of more volatile elements by sputtering. So rather less favoured than either graphite or

silicates it may still be important. Its relatively high cosmic abundance of silicon ensures that it cannot be ignored. The refractive index is 3.55 - 3.0i (Greenberg, 1968).

4.6.5 Organics

A range of organic and even biological materials have been suggested to explain various of the infra-red features, and also the general lack of structure in these bands. Polycyclic aromatic molecules were proposed by Donn (1968) and more recently terrestrial-like bacteria by Hoyle and Wickramasinghe (1979). Their value of m = 1.16 has been adopted to examine this hypothesis.

In conclusion, it seems increasingly probable that dust grains are composed of several species, perhaps each dominating at a different region in the spectrum; the cosmic abundance ratios alone exclude almost any other contention for physically realistic compounds. The position is further complicated by modifications to any such mixture by processing in stellar environs or by the growth of mantles in dark clouds.

4.7 Grain Size Distributions

Grain size distributions have already been mentioned several times. Their importance in defining or modifying many aspects of the radiation scattering functions is crucial. In certain regimes, Rayleigh-Gans scattering, size dominates totally over compositional content, and in the Mie region is of approximately equal weight. Some knowledge of the sizes of the scattering particles is therefore essential, leading many authors to consider a variety of possible functions (Oort and van de Hulst, 1946; Greenberg,

1960; Wickramasinghe and Nandy 1971; Mathis et al., 1977).

Proposed distributions, where n(a) gives the number of grains of radius a in the grain population, include:

Single size: $n(a) = \delta(a - a_o)$ (4.8)

constant: $n(a) = k$; for $a < a_{max}$ (4.9)

$n(a) = o$; $a > a_{max}$

exponential: $n(a) = k.\exp(-a/a_o)$ (4.10)

gaussian: $n(a) = k.\exp(-a^2/a_o^2)$ (4.11)

Deirmendjian's function:

$$n(a) = k.a^{3/2}.\exp(-\tfrac{1}{2}(a/a_o)^3) \quad (4.12)$$

Oort-van de Hulst-Greenberg function:

$$n(a) = k.\exp(-5(a/a_o)^3) \quad (4.13)$$

Power Law: $n(a) = k.a^{-q}$ (4.14)

Single sizes were the first to be considered, presumably on the grounds of simplicity and also because of the limitations of the computational resources available, when these calculations first began. Two major criticisms can be applied to the use of singly sized grains. First, the optical cross-sections for such distributions exhibit strong size dependent resonances at particular wavelengths. These are not found in the observations and are clearly physically unrealistic. The second criticism, which also applies to all the other distributions, save for equations 4.13 and 4.14, is that there is no theoretical justification for these size spectra. Equation 4.9 was used in successful attempts to avoid the resonance problems of equation 4.8. While 4.10 and 4.11 were suggested to include some form of growth effect, leading to size broadening in the Gaussian case.

In any attempt to justify a particular size distribution two observational constraints are applicable. First the function must be able to model the interstellar extinction. Most can do this with varying degrees of success. Secondly the function should explain the general uniformity of the interstellar extinction and polarisation. It is really quite remarkable that these factors are so lacking in variability throughout the galaxy, save for certain regions of star formation and a number of dark clouds already noted. The deduction that equilibrium condition pertains is compelling. For this reason, and also because Hong and Greenberg (1978) have shown equation 4.13 to be a better fit to the extinction then any of the others so far considered we will confine further discussion to equations 4.13 and 4.14. Equation 4.12 was used by Wickramasinghe and Nandy (1971); it derives from Deirmendjians (1969) studies of terrestrial clouds. That this is a regime of pronounced disequilibrium is self-evident.

4.7.1 The OHG Function

The Oort-van de Hulst-Greenberg, hereafter OHG, distribution has gained almost universal favour in recent years. It derives from a modification by Greenberg (1960) of the initial function due to Oort and van de Hulst (1946) which is the equilibrium solution to a simple grain growth and dstruction model. Consideration of the derivation is instructive, for so dominating has the influence of this function become that many authors ignore the implicit assumptions and limitations.

Oort and van de Hulst examined the equilibrium conditions which result from the mutual destruction of grains by

stochastic impacts, during random collisions between interstellar clouds. At that time ice was the favoured chemical constituent and so they considered the steady growth of an icy particle punctuated by total destruction, through evaporation, on impacting with another grain.

Clearly there must be some minimum value of the relative velocity required for complete evaporation, V_{min}. Of importance is the penetration distance before a grain is slowed below this value: for ice 2.8 kms^{-1}, $MgSiO_3$ 5.8 kms^{-1}, iron 7.6 kms^{-1} and graphite 21.5 kms^{-1} (Barlow, 1978b). This distance was calculated from the equation of motion of the dust particle:

$$Mg \frac{dv}{dt} = - \pi a^2 \rho v^2 \qquad (4.15)$$

where Mg: grain mass; v: grain velocity, a: grain radius, ρ: density of the cloud. Then the penetration distance, d, becomes:

$$d = \frac{4a\rho_g}{3\rho} \left(\frac{V_o}{V_{min}}\right) \qquad (4.16)$$

where ρ_g: grain density, V_o : initial grain velocity. On typical assumptions for the required values d equals 0.3 pc and thus gives a destruction probability P:

$$P = n_g \, 4\pi a^2 \, d \qquad (4.17)$$

where n_g is the number density of grains, approximately equal to $7.6 \times 10^{-5} m^{-3}$, giving P=0.87. With a mean collision time of 9×10^6yr between clouds, Oort and van de Hulst obtained a lifetime against destruction of 10^8 yr, which exactly balances grain growth to 0.1 μm during that time. Now the equilibrium condition is given by:

$$\frac{da}{dt} \cdot \frac{dn(a)}{da} + P(a)\, n(a) = o \qquad (4.18)$$

where $\frac{da}{dt}$ is the rate of growth by accretion. Then for P proportional to a^2 solutions of the form of equation 4.13 are obtained. Since many destruction processes are nearly dependent, this appears quite satisfactory.

There are at least three difficulties with this analysis; the first we shall just note in passing. Erosion by sputtering, which for larger particles is arealy dependent, will become a volume regulated function when the sputtering depth becomes comparable with the particle size; thus small particles could be expected to display an anamalous destruction behaviour. Secondly complete vaporisation on impact is not the only obvious modifier of the size distribution. A simpler mechanism, and one which operates to lower relative velocities is shattering. Since the details of this mechanism are complex and have, so far, been neglected in the literature we shall discuss them in depth. Finally, cloud-cloud collisions have been reinvestigated (Stone, 1970; Aannestad, 1973). Barlow (1978b) extending their results concludes that grain lifetimes against collision, in this regime, were considerably underestimated by Oort and van de Hulst. The important alteration to the calculation is that a shock front is set up on the impacting interface alone with the capacity to destroy the grains: with the result that grains are only capable of destruction by this method in times less than 2×10^4yrs. The mean lifetime is then estimated to increase to 1.4×10^{10}yrs, completely invalidating cloud-cloud

collisions as a destruction mechanism. This in itself would not matter too greatly to the form of the size function so long as the main destruction processes were dependent on a^2, which apart from our first point they frequently are (caveat temperature fluctuations in small particles), were it not for shattering.

4.7.2 Shattering

Shattering is a universal phenomenon acting over many scales of length with astonishing regularity in the ultimate product. From aerosols in the Earth's upper atmosphere as small as a hundredth of a micron (Junge, 1963), through terrestrial and lunar dust and rocks - ten microns to ten metres (Hartmann, 1969) - to meteorites and asteroids up to hundreds of metres across (Hellyer, 1970), shattering continues to exhibit characteristic bulk behaviour which can best be described by a power law distribution. A slight complication results from the different presentation of comparable data on size spectra in the different areas of study. We are always concerned with size distributions in the range a to a + da. Geological and atmospheric observers usually present their data as mass dependent cumulative logarithmic functions. To adjust for this alteration in variable and scale it is merely necessary to subtract one from, for example, Hartmann's indices and multiply by three. The latter modification being on the assumption that the mass is proportional to volume. Hartmann (1969) has reviewed terrestrial rock fragmentation, lunar and terrestrial secondary impact cratering, ejecta from hypervelocity blasts,

gravels, clays, volcanic steam explosions, the debris in the vicinity of six of the seven Surveyor probes and also telescopic asteroid measurements. He made two discoveries: in every case a power law developes, the precise index being dependent on the degree of multiple fragmentation, and secondly that a general function for the incremental mass distirbution of fragments mass, m, resulting from the desctruction of an initial body mass m_o, was given by:

$$Z\ (m_o,m) = \frac{A\ m_o^{\ x}}{m^{x+1}} \qquad (4.19)$$

where A and x are constants.

Hellyer (1970) has used this result to find the equilibrium fragmentation function, under the assumptions that no particles are lost or gained to the system, and that fragmentation alone modifies the mass distribution. This requires solution of the differential equation:

$$\frac{\delta n(m,t)}{\delta t} = -n(m,t)p(m,t) + \int_m^{m_{max}} n(m_o,t)p(m_o,t)z(m_o,m)dm \qquad (4.20)$$

for $m > o$, $t > o$ and where $n(m,t)$ is the number of particles in the range m to m + dm at time, t, and $p(m,t)$ is the probability of a particle of mass, m, being destroyed between times t and t + dt. Under equilibrium conditions $n(m,t)$ must deconvolve to the separable functions: $n(m,t) = a(m)b(t)$, since by definition we are entering a steady state. Using these assumptions Hellyer develops an equilibrium solution for n(m):

$$n(m) = k.m^{-y} \qquad (4.21)$$

where y is approximately 1.8 and k is some constant. For a size distribution this becomes simply:

$$n(a) = c.a^{-5.4} \qquad (4.22)$$

Thus, there are both observational and theoretical reasons to believe that shattering produces power law size spectra. We have already suggested that shattering needed to be considered in the calculation of Oort and van de Hulst, but if the grain lifetime against collision is as long as that calculated by Barlow, shattering by this mechanism may not be important.

There is however, another situation in which it could play a part. Burke and Silk (1976) have considered the processes operative during proto-stellar collapse. Throughout the free-fall time the grains are strongly coupled to the gas, which should effectively shield them from collisions, although they may grow by the normal process of accretion. However, once a radiation driven shell forms, along the lines of the model due to Larson, for stars greater than five solar masses, the situation alters. Convective turbulent motion in the expanding shell will produce eddies. Grains from one eddy can collide with those from another; thus the relative velocity of grains will equal that of the convective eddies, resulting in a grain collision time much shorter than the evolutionary time of the shell. Presumably some fraction of the shattered particles will be evaporated; Burke and Silk have attempted to calculate the additional effects of this loss and also subsequent accretion. They find that the final spectrum is strongly dependent on the

evaporated fraction, f. This is to be expected: if f is high we should revert to an OHG distribution. This was also their conclusion, but for low or intermediate values of f a power law results. The only effect of f then is to translate the function linearly. Attempts to include the additional possible effect of nucleation by particles that fall below some minimum shattering size - if one exists in the size range we are considering - indicate that bimodality is possible. Given that this modification process might effect as much as half the total population of grains (Burke and Silk, 1976) makes consideration of the effects essential. What is urgently required is some empirical derivation of the size distribution. Fortunately Mathis, Rumpl and Nordsieck (1977), unaware of the effects of shattering have attempted just that.

4.7.3 Empirical Determination of the Power Law Distribution

Mathis et al. (1977) have sought to apply quadratic programming to the interstellar extinction data, for a variety of chemical species. Quadratic programming is very similar to least squares fitting requiring the minimisation of χ^2:

$$\chi^2 = \sum_i (C_{pi}-C_{oi})^2/(\sigma_i C_{oi})^2 \qquad (4.23)$$

where C_{pi} is the predicted quantity, C_{oi} the corresponding measured value and σ_i the relative error. The difference from least squares fitting is in the inclusion of inequalities. In this case for example, the number of chemical species must exceed zero. The observations used were the

extinction values as a function of wavelength between 0.1 μm and 1.0 μm. The materials considered included: graphite, enstatite, olivine, silicon carbide, iron and magnetite. Trial and error modelling, binning the relative numbers of particles according to size produced three results: a substantial range of sizes, a rapid decrease in population with increasing size and unphysical empty bins. Mathis and his co-workers were struck by the similarity in their histograms of grain size distribution by the resemblance to a power law. This led them to repeat their calculations using power law functions. Good fits were obtained for indices of about -3.5 for mixtures involving graphite and silicates, or graphite and magnetite, or graphite and silicon carbide. Graphite was necessary to explain the 2175Å feature. Martin (1978) has pointed out that although silicates were not particularly favoured by these tests, that this is perhaps because the strength of the silicate band is unknown and as such could not be included in the calculations. A low band strength would demand a higher fraction of silicates.

Knowing that shattering tends towards a power index of -5.4 it is encouraging that Mathis et al.'s values do not exceed this factor.

Weighing these considerations, and also a further test to be outlined in the next section, we shall use a power law function for our subsequent calculations across size spectra.

4.8 Mie Theory

The scattering of electromagnetic radiation by particles of any shape demands the solution of Maxwell's equations both within and without the particle and also on the surface of the particle. The obvious complexity of the problem ensures that few analytical solutions exist. At the asymptotic limits for small and large particles solutions exist, but for an arbitrarily shaped particle of size comparable to the wavelength, the Mie regime, there are only a few specialised solutions, for conditions of high symmetry. The first such solution, for homogeneous spheres of unrestricted size, was discovered independently by Mie (1908) and Debye (1909). The only other complete solution is for infinite cylinders (Greenberg, 1968), but a number of approximate experimental determinations for spheroidal particles have been evaluated: Chylek et al. (1976) in the optical, Zerull and Giese (1974) for microwave analogue experiments. These last authors - Chylek with certain reservations regarding the correct adjustments to the Mie theory for the surface wave components of non-smooth spheres - have shown that the scattering behaviour of irregular randomly orientated particles is approximated with good agreement to the experimental determinations, by spherical particles. The present work proceeds on that basis, using spherical solutions for various materials noted in the previous section.

4.8.1 Implementation

An account of this solution is presented by van de Hulst (1957), and a computational scheme for its implementation is given by Wickramasinghe (1973). The scattering problem is

completely defined by the wavelength of the radiation, λ, the radius of the scattering particle, a, its shape and the refractive index of the particle, m = n-ik. Van de Hulst introduced the dimensionless quantity, x:

$$x = 2\pi a/\lambda \tag{4.24}$$

From this the resulting scattering amplitudes at large distances from the scattering particles are given by:

$$S_1(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left[a_n \pi_n(\cos\theta) + b_n \tau_n(\cos\theta)\right] \tag{4.25}$$

$$S_2(\theta) = \sum_{n=1}^{\infty} \frac{2n+1}{n(n+1)} \left[b_n \pi_n(\cos\theta) + a_n t_n(\cos\theta)\right] \tag{4.26}$$

where $$\pi_n(\cos\theta) = \frac{1}{\sin\theta} P'_n(\cos\theta) \tag{4.27}$$

and $$\tau_n(\cos\theta) = \frac{d}{d\theta} P'_n(\cos\theta) \tag{4.28}$$

P_n represents the Legendre polynomial functions, a_n and b_n are spherical Riccati-Bessel functions and θ is the angle to the beam direction.

$S_1(\theta)$ is the scattering amplitude perpendicular to the scattering plane, and $S_2(\theta)$ represents the amplitude in the scattering plane, perpendicular to the direction of propagation. The scattering function for a single particle is defined as:

$$F(\theta) = \tfrac{1}{2}\left[|S_1(\theta)|^2 + |S_2(\theta)|^2\right] \tag{4.29}$$

In the following calculations we will be examining grain size distributions, n(a), where n(a) is given by equation 4.14:

$$n(a) = k.a^{-q} \tag{4.14}$$

where normalisation is applied, between the limits of the maximum and minimum grain size:

$$\int_{a_{max}}^{a_{max}} n(a)da = 1 \qquad (4.30)$$

The average scattering functions then become:

$$F_i(\theta) = \int_{a_{min}}^{a_{max}} |S_i(\theta)|^2 \cdot n(a) \cdot da \qquad (4.31)$$

for i = 1,2.

The final intensity is then given by:

$$I(\theta) = F_1(\theta) + F_2(\theta) \qquad (4.32)$$

and the percentage polarisation by:

$$P(\theta) = \left| \frac{F_1(\theta) - F_2(\theta)}{F_1(\theta) + F_2(\theta)} \right| \times 100 \qquad (4.33)$$

The applications of these relationships utilises the computer implementation of Warren-Smith (1979) with some minor alterations.

New calculations are performed with the grain size distribution indicated. This has several advantages; to those presented above should be added the smoothing effect of a spectrum of sizes. Single particles pass through resonance conditions (Wickramasinghe, 1973) which are computationally problematic and cause unphysical "ripples" in the calculated intensity and polarisation functions. Various methods have been employed to evaluate these integrals, from the error-prone five-point Simpson Rule of Roark (1965), through the adaptive Simpson Rule of Hanner (1969) to the 255-point adaptive gaussian quadrature method, due to Patterson (1968) employed by Warren-Smith (1979). The present calculations use an improved method, from the NAG subroutine library, that elaborates on Patterson's

procedure. Patterson's algorithm is employed as a subset of the integrating process. This is improved by the addition of new rules under which the rate of convergence is monitored, decisions taken thereby re subdivision of the integral, and singularities are detected. If sharp end-point peaks on singularities do occur appropriate transformations are applied to render the function integrable. Finally random transformations of the integrand are applied to improve reliability in specific cases (NAGFLIB, Mk8, 1981).

4.8.2 The Single Scattering Results

Previous calculations of Mie scattering for various materials and grain sizes (Hanner, 1969; Zellner, 1970; Greenberg, 1973; Hong and Greenberg, 1978) characterise their ensemble of scattering particles by some mean size, $\bar{a}$. This cannot be done in the case of a power law distribution, since this function has no single mean or modal radius, independent of the integration limits. The obvious consequence of this distinction is that while all other size distributions considered (section 4.7) are approximately flat in their wavelength response, i.e. measurements at different wavelengths will produce roughly the same value for the mean size, the grain sizes of a power law function are effectively selected by the wavelength of the scattered light, in a linear manner. The dioristic parameter in this instance is then the index of the power law, q, equation 4.14. A further consequence is the need to integrate within size limits: a_{min} and a_{max}, respectively the smallest and largest grain sizes considered. In practice this introduces no additional alteration into a computational system.

The demonstration of both these contentions is quite simple. Consider an integral of the form:

$$I = \int_c^d x^{-k} \; ; \quad k > 1 \tag{4.34}$$

the solution of which is given simply by:

$$I = \frac{1}{(1-k)} \; x^{(1-k)} \Big|_c^d \tag{4.44}$$

This function becomes aasymptotically indefinite. Thus the mean grain size for a power law function is:

$$\bar{a} = \int_{a_{min}}^{a_{max}} n(a).a.da \; / \int_{a_{min}}^{a_{max}} n(a)da \tag{4.45}$$

$$\bar{a} = \int_{a_{min}}^{a_{max}} a^{1-q}.da \; / \int_{a_{min}}^{a_{max}} a^{-q}da \tag{4.46}$$

Ergo:

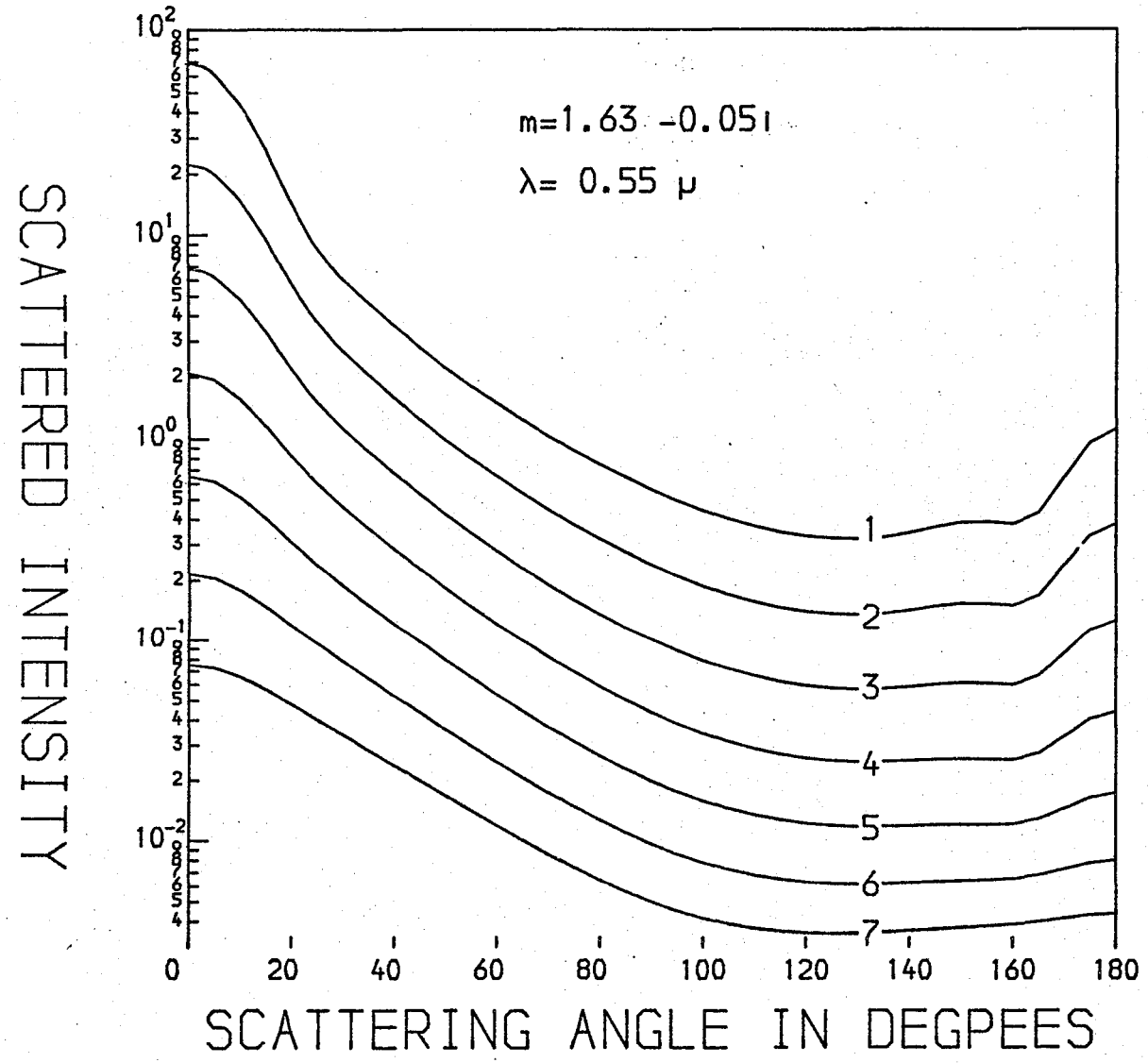
$$\bar{a} = \left[\frac{1}{(2-q)} \cdot a^{2-a} \; / \; \frac{1}{(1-q)} \; a^{1-\rho} \right]_{a_{min}}^{a_{max}} \tag{4.47}$$

$$\bar{a} = \left(\frac{1-q}{2-q}\right)a \Big|_{a_{min}}^{a_{max}} \tag{4.48}$$

Thus it is apparent that there is no unique size and that the size measured depends on the frequency of the radiation used, since this will decide a_{min} and a_{max}.

Integrated polarisation and intensity scattering functions for power law distributions of singly scattering spheres are presented, for a selection of materials, figures 4.2 to 4.8, for a wide range of indices. The most salient features are as follows: first, figures 4.2 and 4.3 showing the intensity

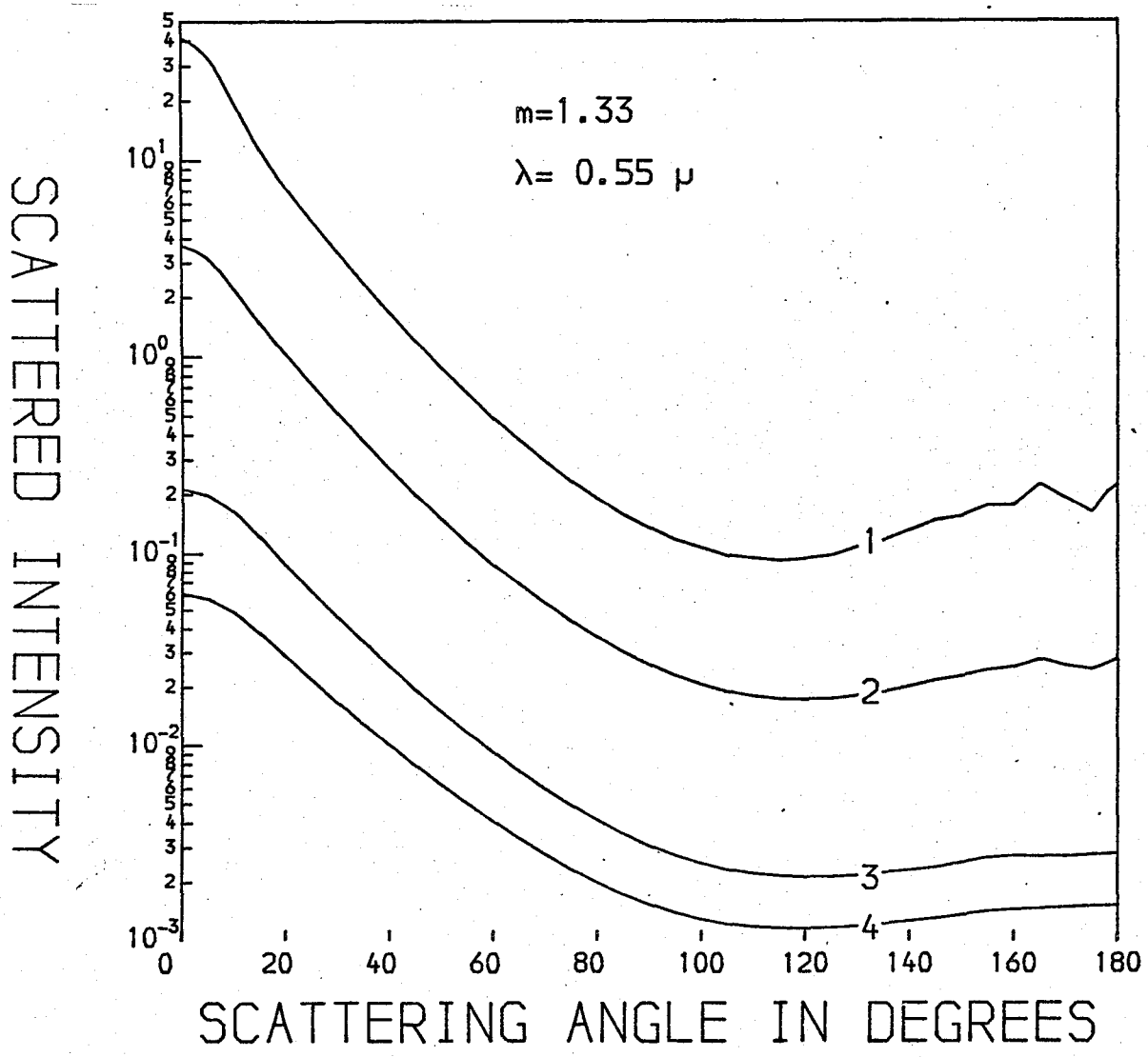
Fig. 4.2



SILICATES: $n(a) = k \cdot a^{-q}$

curve no.	q-values
1	2.0
2	2.5
3	3.0
4	3.5
5	4.0
6	4.5
7	5.0

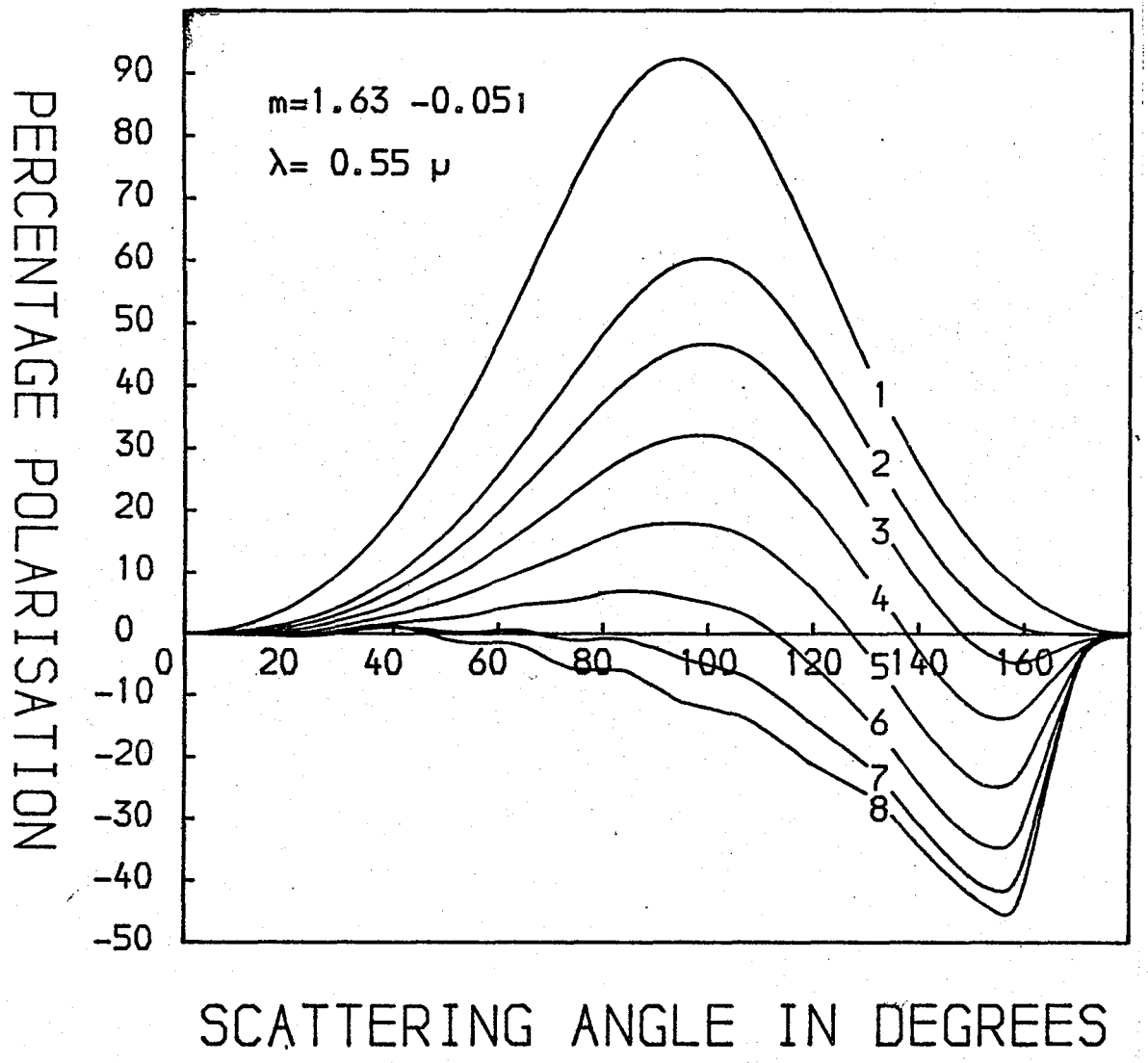
Fig. 4.3



ICE: $n(a) = k \cdot a^{-q}$

curve no.	q-value
1	5.0
2	4.5
3	4.0
4	3.5

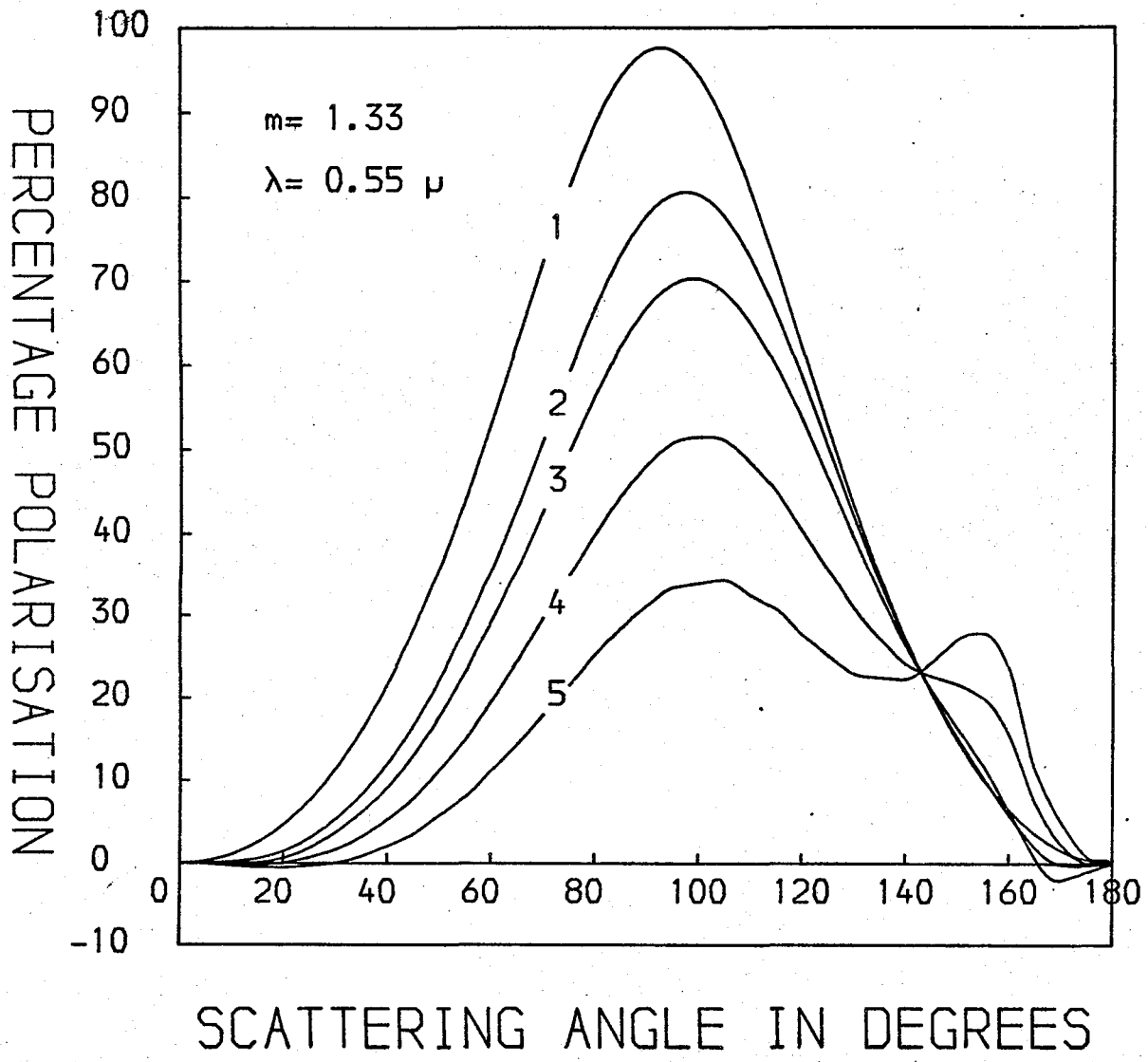
Fig. 4.4



SILICATES: $n(a) = k \cdot a^{-q}$

curve no.	q-values
1	2.0
2	2.5
3	3.0
4	3.5
5	4.0
6	4.5
7	5.0
8	7.0

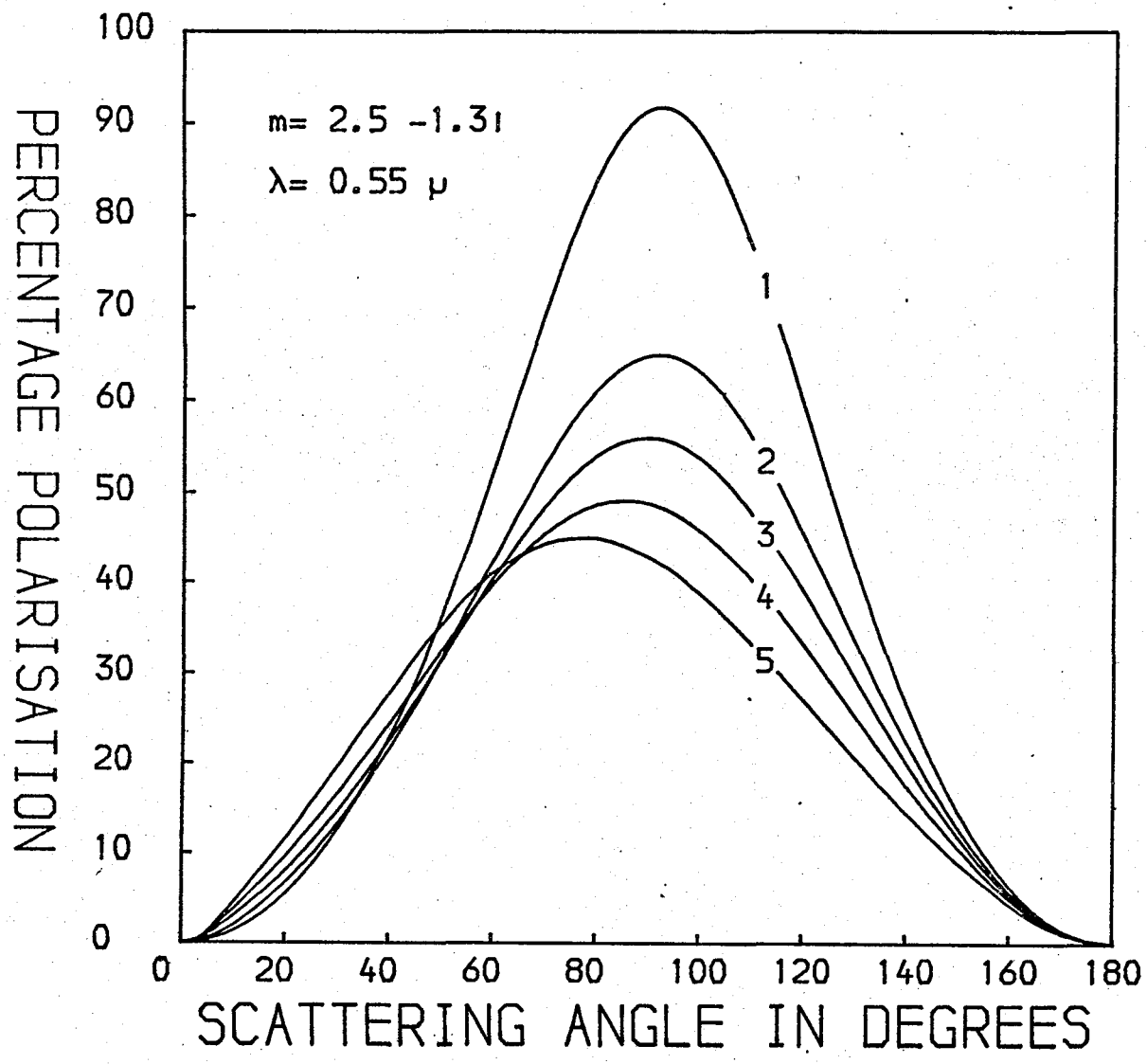
Fig. 4.5



ICE: $n(a) = k.a^{-q}$

curve no.	q-value
1	7.0
2	5.0
3	4.5
4	4.0
5	3.5

Fig. 4.6

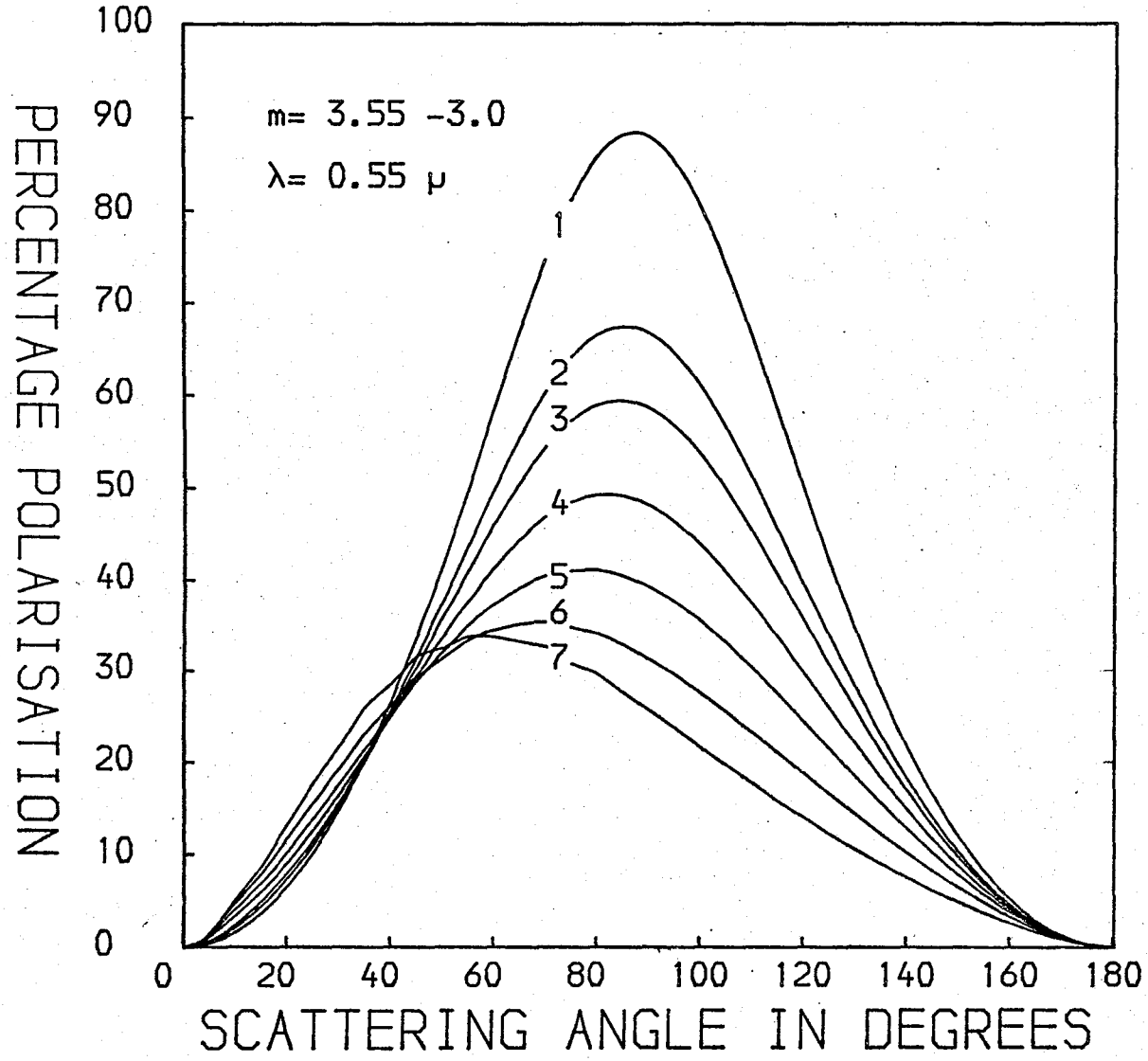


Graphite:

$n(a) = k \cdot a^{-q}$

curve no.	q-value
1	7.0
2	4.5
3	4.0
4	3.0
5	3.5

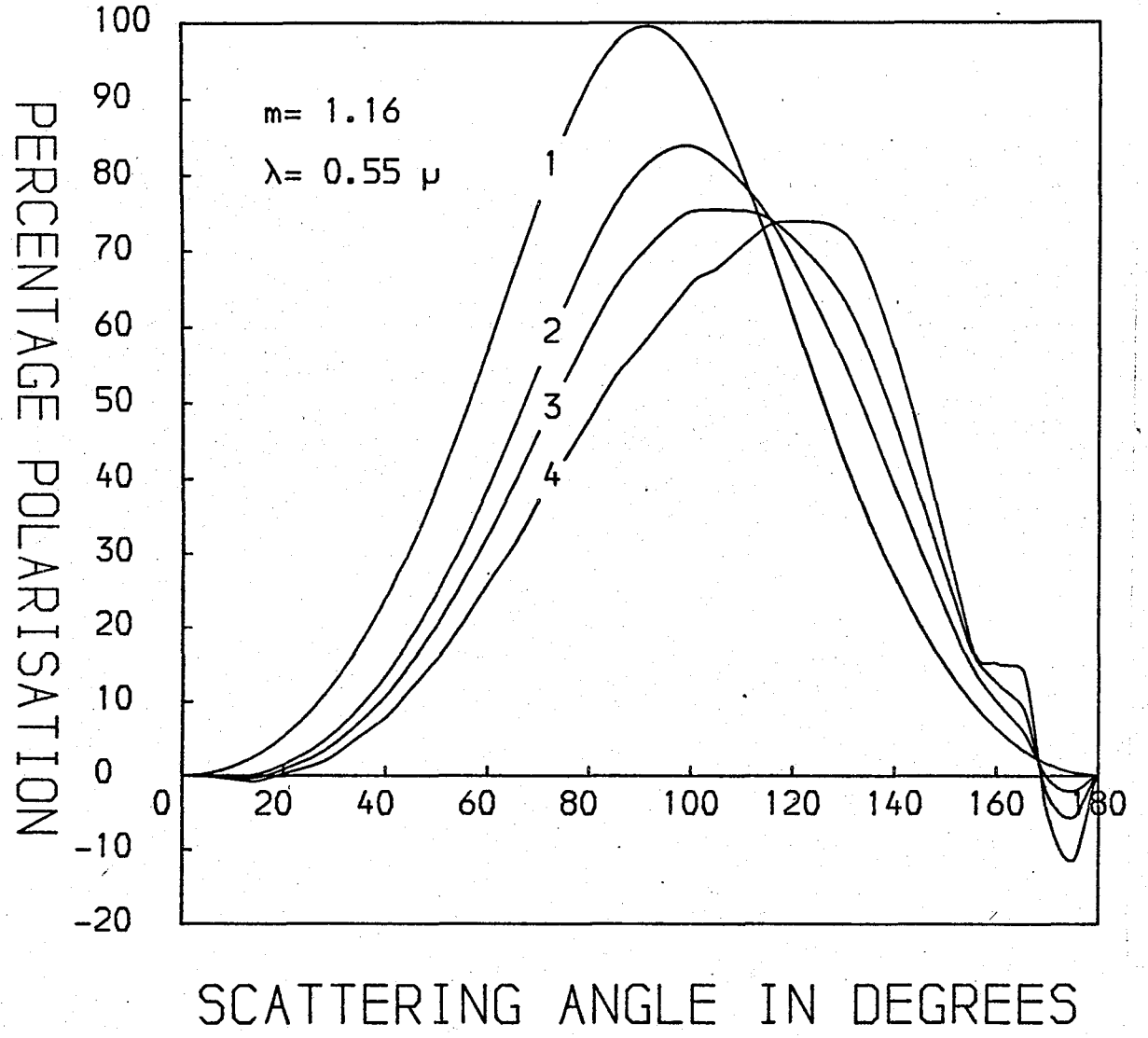
Fig. 4.7



IRON: $n(a) = k.a^{-q}$

curve no.	q-value
1	7.0
2	5.0
3	4.5
4	4.0
5	3.5
6	3.0
7	2.5

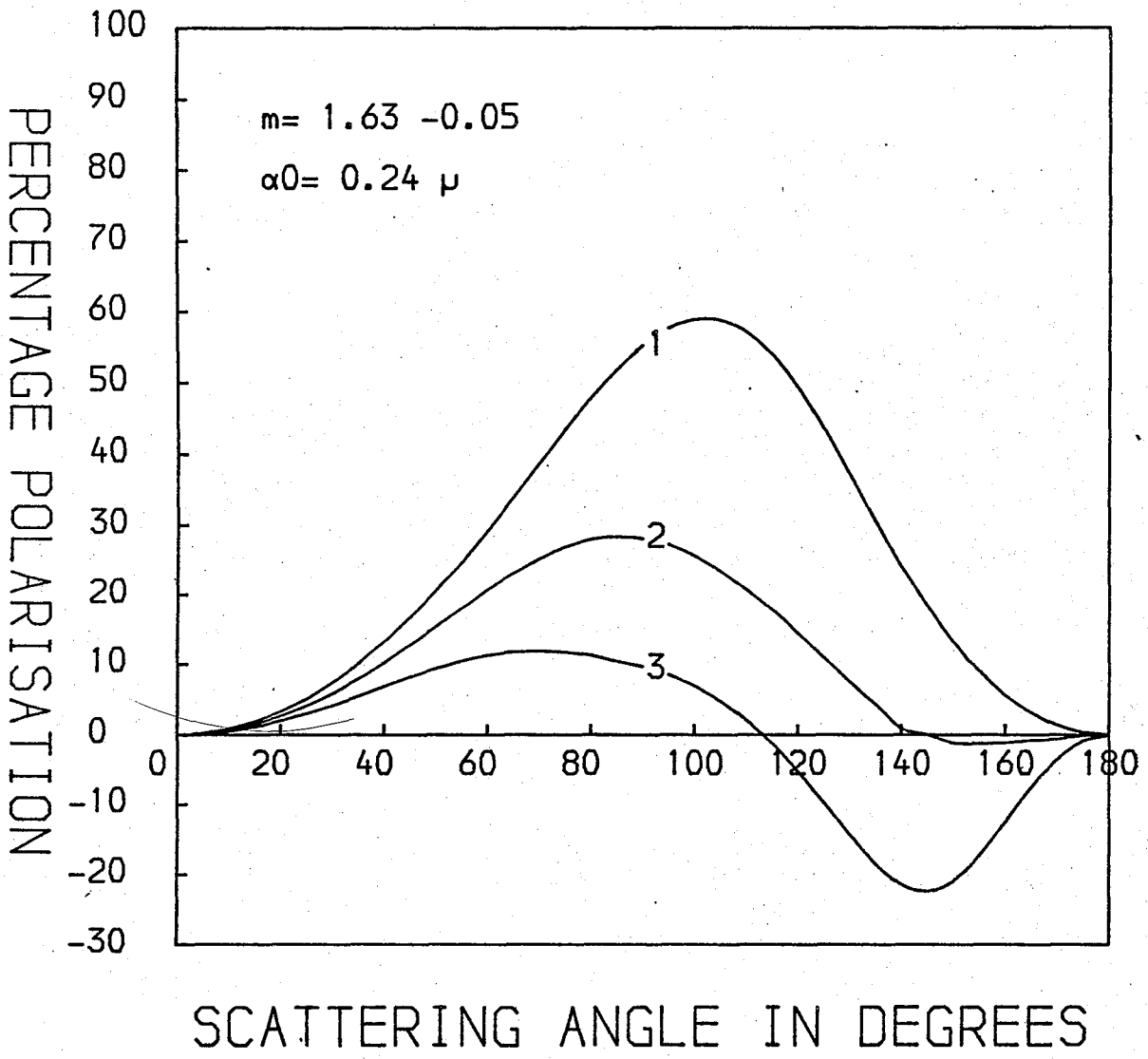
Fig. 4.8



ORGANIC: $n(a) = k \cdot a^{-q}$

curve no.	q-value
1	7.0
2	4.0
3	3.5
4	3.0

Fig. 4·9



SILICATES:

$n(a) = k \cdot \exp(-5(a/a_o)^3)$

curve no.	λ (μm)
1	0·65
2	0·55
3	0·44

of the scattered light as a function of angle. These graphs are plotted on a logarithmic scale. Only Silicates and Ice are plotted since the forms of the intensity functions are so similar from material to material that little advantage accrues in presenting more.

Intensity is characterised, for all materials, by a strongly forward throwing nature, that is nearly all the radiation is scattered at or near to zero degrees. Many materials have a slight peak at 180 degrees. The depth of the trough between nought and 180 depebds on material and power index. Polarisation behaviour is much more complex. Obviously by symmetry, polarisation at nought and 180 degrees must be zero in all cases. All materials exhibit a peak in the vicinity of ninety degrees. Further for high values of q, i.e. large negative indices, typically q = 7, polarisation tends towards one hundred percent. In this region the irregular behaviour of the functions become much more symmetric about 90 degrees, as the position of the peak tends towards coincidence with ninety. Interestingly negative polarisations are not produced by all materials and most substances have a minimum peak polarisation, e.g. graphite figure 4.6. Negative polarisation, where it is found, always results from backscattering.

In a separate calculation, figure 4.9, the wavelength behaviour for an OHG distribution of spheres size $\bar{a}$ = 0.24 μm was investigated. The corresponding power law calculation has all the curves exactly superimposed for any wavelength, but here the massive change in polarisation with wavelength is obvious. As no such behaviour is noted in NGC 7023, but

rather as the polarisation is almost wavelength independent, no further consideration will be given to the OHG function.

We now proceed, chapter five, to apply these scattering functions in a multiple scattering regime, with varying geometries, applied to the case of NGC7023.

CHAPTER FIVE

Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin.

John von Neumann

MONTE CARLO SIMULATIONS OF NGC7023

5.1 INTRODUCTION TO NEBULAR SIMULATION

Attempts to understand the propagation of light in reflection nebulae began with analytical considerations of the transfer of radiation in conditions of high symmetry through homogeneous media (Henyey and Greenstein, 1938; van Houten, 1961; Minin 1962). The most simple and therefore, the most investigated conditions are single scattering in uniform spherical nebulae (Vanysek and Svatos, 1964). Early theoretical and numerical work concentrated on the examination of surface brightness variations (Greenberg and Roark, 1967; Roark et al., 1974), because colour or polarisation calculations make analytical modelling prohibitively complex and primitive numerical studies excessively demanding computationally.

The history of the simulation of radiation transfer in nebulae is very much the history of the evolution of the numerical techniques involved. The fundamental concern is the same: to model some property of the light emitted as a function of a range of physical parameters. The first models were of surface brightness (Dahn, 1967), then colour (Greenberg and Hanner, 1970) and finally polarisation (Hanner 1971). At the same time the scope of the problem was extended in other

directions. Gehrels in 1967 was thinking only in terms of single scattering. Zellner (1970) realised that surface brightness measurements proved that there was an important contribution from multiple scattering, which he could not model. However, he was puzzled by the comparatively good results provided by single scattering calculations. Greenberg and Hanner (1970a), modelling both single and double scattering numerically, resolved this difficulty in terms of grains which scatter most of their radiation at 0^{o} incidence. With the increasing use of numerical methods to integrate the Mie scattering functions along the various lines of sight, different geometries became available for study and with them different techniques of integration. In the early seventies spherical structures were abandoned in favour of parallel plane models. Stars were embedded within the nebulae or placed outside them and the planes could also be inclined (Hanner, 1971). The change of the main structure under investigation, and the obvious need to study higher orders of scattering, encouraged the application of Monte Carlo (Roark et al., 1974) and optical depth doubling techniques (White, 1979a,b). The latter is only valid for geometries in which the light is incident parallel to the scattering surface. Since we shall be considering much more complex structures we shall not examine this method further.

The Monte Carlo techniques started badly (Roark et al., 1974) but subsequent generations of models have advanced enormously. Roark et al. used what is known as the "Crude Monte Carlo technique" (i.e. direct modelling of the physical process) to consider the three dimensional transfer problem

in a cylindrical nebula. The fundamental drawback with this technique is the enormous requirement of computer time. To simulate surface brightness alone they needed between 6 x 10^6 to 3 x 10^7 photon interactions involving, in their best model, as many as 1.7 x 10^6 "observable" photons. That "observable" is crucial. In any simple percolation model light will propagate outwards from the source in all directions. Since the observer subtends only a minute area the ratio of calculated to "observed" photons is going to be astronomical. Two different techniques have been applied to reduce these demands on cpu time. First: pre-calculation - the Mie scattering functions operate on the Stokes parameters of the incident radiation as matrix operators. These are constant for any particular angle of incidence, scattering material, wavelength and particle size. Celnikier and Lefevre (1974) introduced the technique of calculating these matrix elements once, for a number of discrete conditions, and storing them in a look-up table. Examination of a stored table is a much faster operation than multiple calculation of the scattering operations. Secondly Witt (1977a) introduced a statistical weighting technique, common in other Monte Carlo regimes, known as importance sampling (Rubinstein, 1981). This procedure applies a deliberate distortion to the probability functions for scattering and uses these distorted distributions to examine most carefully those parts of the scattering process of most importance in the results. The effect of the distortion is corrected afterwards by a suitable weighing function. However the applications of these weighting functions was still very simplistic. Witt (1977 a,b,c) and

Witt and Oshel (1977) were still only capable of detailed studies of surface brightnesses. Daniel (1978, 1980, 1982) extended the methods of Celnikier and Lefevre (1974) to consider polarisation in circumstellar shells and White (1979) using the optical doubling technique has modelled polarisation in plane parallel nebulae.

Still the major source of inefficienty in the calculation remained the loss of nearly all the photons as they propagated away from the observer. Warren-Smith (1979) tackled this problem directly. He considered radiation within an infinite cloud. After the final scattering the direction of propagation will not in general be in the direction of the observer. He solved this problem by rotating the entire structure until the photons did travel towards the observer after the final scattering. Only then are the nebular boundaries applied and the photon accepted if it has never been scattered outwith the nebula. The obvious limitation of this technique is that it demands a constant or spherically symmetric density distribution within the nebula, but the saving in computer time is enormous. Errors on percentage polarisation of only a few per cent are possible for only a few thousand photons : the precise values depend on the detailed geometry. Warren-Smith considered only a planar model, appropriate to NGC1999, however his computer programs are easily adapted to any convex or flat structure of constant density, with an embedded light source. The simulation is incapable of modelling structures with the star outside the nebula or those with concave surfaces. The new calculations build upon this simulation to consider spheroidal structures with an

embedded star and then extend the scope of the model to allow it to consider cavities within the nebula and also geometries in which the star is outside of the nebula altogether. The improved simulation is applied specifically to the case of a hollow conical cavity in an infinite nebula, but could be used over a much larger range of geometries.

5.2 THE IMPLEMENTATION

5.2.1 The General Problem

The general problem in the simulation of light scattering in an illuminated cloud reduces to the integration of the contributions from every element along the line of sight (see Figure 5.1). Thus the problem can be divided into two components: (1) the calculation of the intensity of each element of length δz on the line of sight column and (2) the summation of these elements to obtain the total flux at the observer per surface area of nebula. An obvious consequence of the second point is that we must know the nebular boundaries: z_1 and z_2 in Figure 5.1, and similarly for the x and y-dimensions. Thus where W_j is the total power scattered by a line of sight column, w_j the energy scattered by a volume element, N the number of scattering particles per unit volume and A the area of the element, we obtain:

$$W_j = \int_{z_1}^{z_2} w_j \, NA \, dz \; ; \qquad j = 1,2 \qquad (5.1)$$

Now where I is the intensity of the scattering element and $F_j(\theta)$ the Mie functions of chapter four, a single grain will scatter a power:

$$w_j = \left(\frac{\lambda}{2\pi}\right)^2 \frac{F_j(\theta)}{2} \, I \qquad (5.2)$$

Fig. 5.1

Scattering along a line of sight column

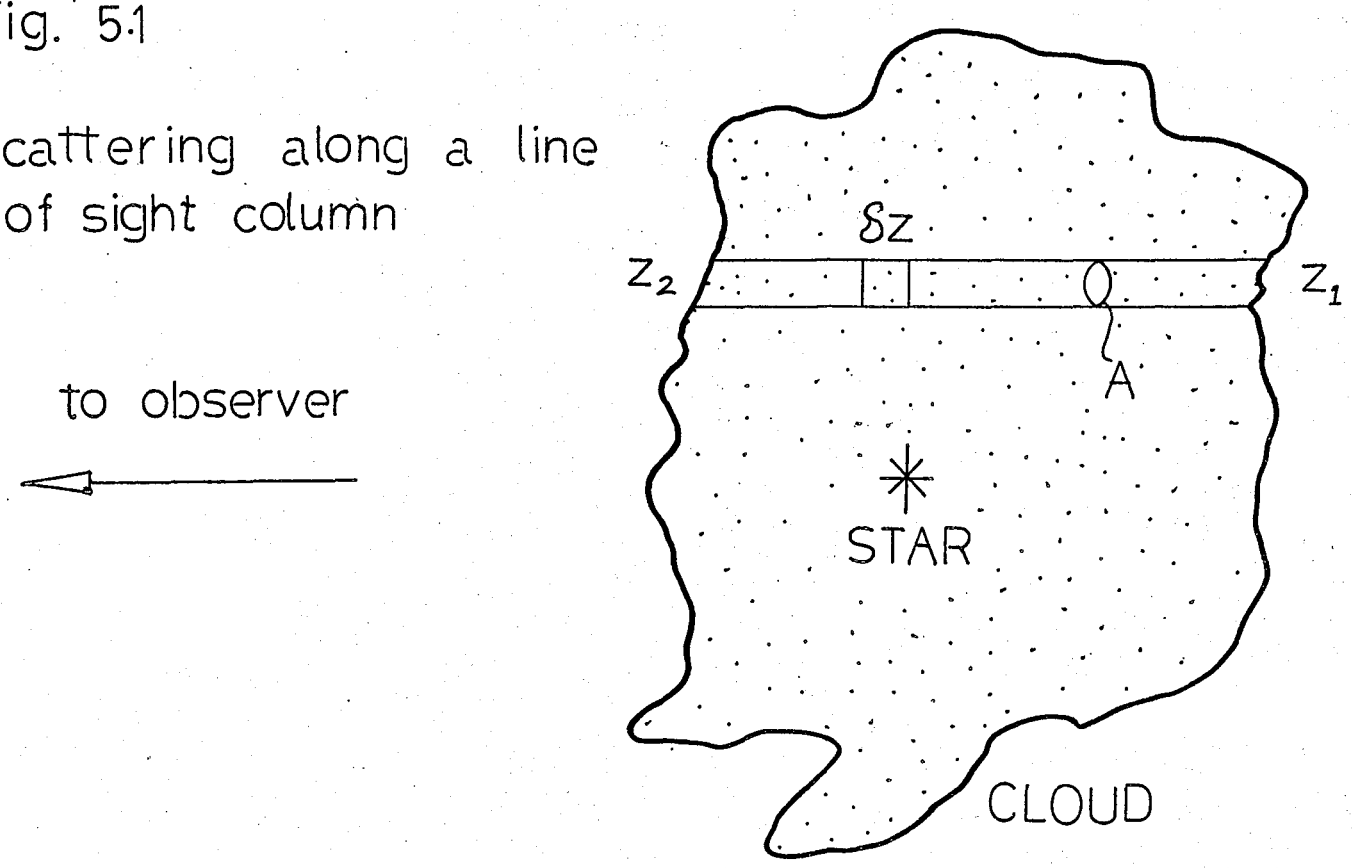
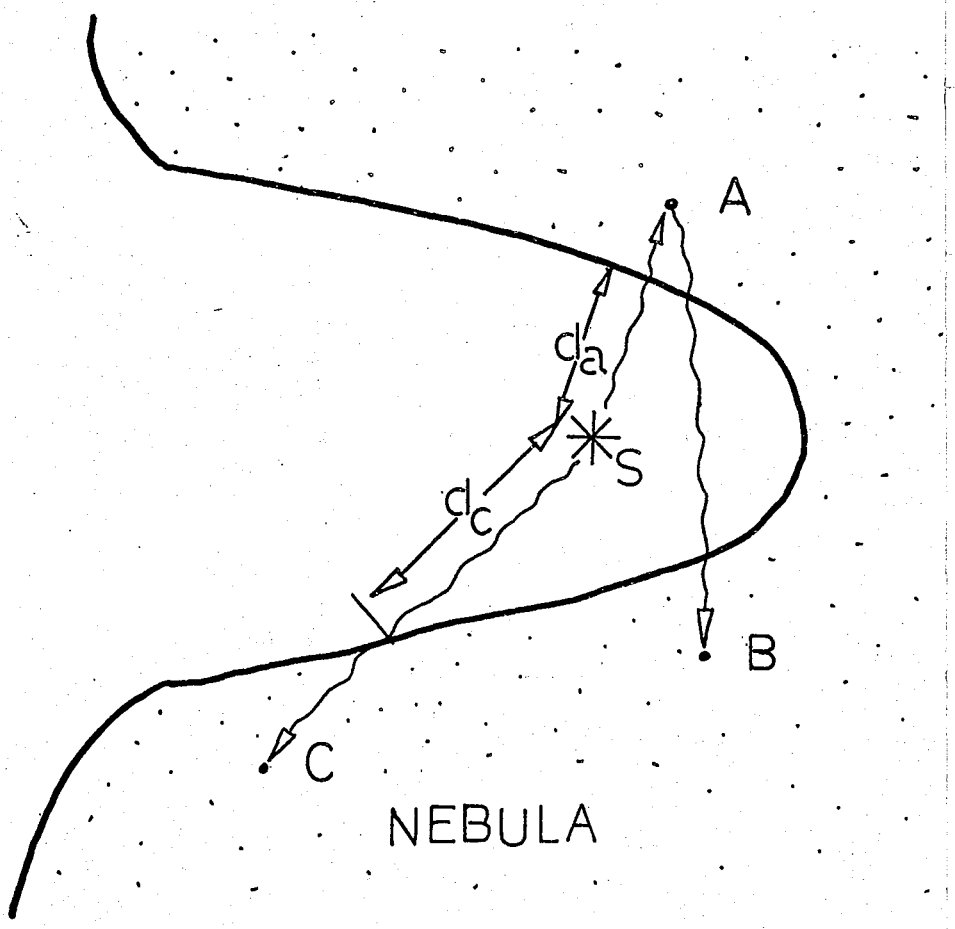


Fig. 5.2

Star in a parabolic cavity



In this case $I = L/4\pi\ell^2$, where L is the luminosity of the star. Thus direct substitution and a change of variable gives:

$$W = \frac{LA\lambda^2}{32\pi^2 r} \int_{\theta_2}^{\theta_1} N(\theta)[F_1(\theta) + F_2(\theta)]d\theta \qquad (5.3)$$

The flux, I_o at the observer, per unit area per square arc second is then simply $W/A\alpha^2$, α being the number of arcseconds in one radian (206, 265). Now the stellar flux, I_* at the observer at distance, D, from the star is:

$$I_* = L/4\pi D^2 \qquad (5.4)$$

Thus direct substitution gives:

$$I_1 = \frac{I_*\lambda^2 D^2}{8\pi^2 r\alpha^2} \int_{\theta_2}^{\theta_1} [F_1(\theta) + F_2(\theta)]N(\theta)d\theta \qquad (5.5)$$

Now consider a star embedded within a nebula of optical depth τ_o, at a distance r from the front face. The light from the star will be extinguished by a factor exp $(-r_o/\tau_o)$. Finally converting to magnitudes and introducing a factor $\gamma(\theta)$ which adjusts for the effects of extinction within the nebula, on the Mie scattering functions, we obtain

$$m_{neb} = m_* - 1.086 \frac{r_o}{\tau_o} - 2.5 \log_{10} \{\frac{\lambda^2 D^2}{8\pi^2 r\alpha^2} \int_{\theta_2}^{\theta_1} [F_1(\theta) + F_2(\theta)]N(\theta)\gamma(\theta)d\theta\} \qquad (5.6)$$

To estimate this integral by the Monte Carlo technique we adopt the procedure of Warren-Smith (1979). The Monte Carlo method is essentially an attempt to evaluate an integral of the form:

$$R = \int_a^b f(x)dx \qquad (5.7)$$

In this case I is a function of $\underline{K}$ the propagation direction, $\underline{S}$ the Stokes vector and z the position on the line of sight. If $M(\underline{K})$ is the Mie scattering matrix the power scattered into a unit solid angle in the direction of the observer in an element of size δz is:

$$\delta z = NA\ \delta z \iint_{\underline{K},\underline{S}} I(\underline{K},\underline{S},z)\ M(\underline{K})\ \underline{S}\ d^3K\ d^4S \qquad (5.8)$$

Now the scattered radiation will be extinguished within the nebula before reaching the observer, according to:

$$\exp\left(\frac{z - z_2}{\tau_o}\right) \qquad (5.9)$$

Thus the integrated power along the line of sight column becomes:

$$R = NA \int_{z_1}^{z_2} \int_{\underline{K},\underline{S}} I(\underline{K},\underline{S},z)\ M(\underline{K})\ d^3K d^4S\ \exp\left(\frac{z-z_2}{\tau o}\right) dz \qquad (5.10)$$

Equation 5.10 can be evaluated by the Monte Carlo method if random samples of the $\underline{K}$, $\underline{S}$ and z with a probability distribution proportional to I can be obtained. This can be done by generating photon paths and following them through the nebula. The random number generator used was selected from the NAG Fortran library. We shall now described Warren-Smith's algorithm to perform this task in the case of a star embedded within a homogeneous convex nebula, as a necessary basis for the extension to more complex geometries. Complete details are given by Warren-Smith (1979).

5.2.2 The Simulation for an Embedded Source

Essentially the simulation performs the following task. It follows a photon propagating into an infinite cloud away from the embedded star. The scattering probability is a function exponentially distributed with distance, so the distance to the first and subsequent scatterings is calculated efficiently by a model which mimics the physical process exactly. When the photon strikes a dust particle it is scattered according to the Mie theory. The form of the resulting Stokes vectors being selected from the look-up table, calculated only once at the start of the simulation. However examination of Figures 4.2 or 4.3 will show that the Mie functions are extremely forward throwing, regardless of refractive index, i.e. material composition. Thus to calculate scattering probabilities for isotropic scattering would be very inefficient. Thus a distorted probability function is applied using the principles of importance sampling. The statistical error of the distortion is removed afterwards by the introduction of a weighting function. Thus with the new propagation direction calculated the photon is allowed to travel on to its next collision, and so on.

After the final encounter the photon will not in general be in the desired line of sight column. To ensure that it is the entire photon path it is rotated to intersect with the line of sight. Usually there will be two such intersections. However, if the column is far removed from the star the photon may never propagate to that distance and there will be no intersection. Thus in general the statistical errors must increase with offset distance from

the star. Given that intersection does occur it is then possible to calculate the contribution of each intersection point to the received flux. Clearly this rotation must distort the probability distribution function for scattering. However multiplication by a geometric weighting factor, $W_g(z)$ is sufficient to adjust for the effects of the distortion.

$$W_g(z) = \frac{1}{|4\pi\rho z|} \qquad (5.11)$$

This rotation of the photon path demands a circularly symmetric density distribution for the nebula, centred on the light source. In fact a constant density is assumed. Finally the spatial boundary conditions are applied and each scattering point tested to check that it is within the nebula. Any outside the nebula are rejected. This process is repeated for every photon and for every point on the nebular surface, for which simulation is required. It is very efficient since nearly every photon is used, most at least twice. The drawbacks are that the calculated values are:

(1) highly correlated

(2) the density distribution must be constant

(3) the star has to be within the nebula, and

(4) no cavities in the surfaces of the nebula are permitted.

Point (1) is partially countered by the use of both the forward and backward scattering points along the lines of sight, since they are anti-correlated. Points (3) and (4) are the subjects of improvements to the model.

5.2.3 The Improvements to the Simulation

The existing model is unable to simulate such geometries as a star illuminating a plane parallel slab situated behind the star, or a star placed in a cavity on the edge of some surrounding nebula. As some observations (c.f. Chapter 2) have indicated that such a possibility might apply in the present case it was thought necessary to construct such a simulation. To do so consider the parabolic cavity in the nebula shown in Figure 5.2. In the original scheme two errors will result if routines are constructed to define nebular boundaries of this form. The lesser problem is that photons can travel from A to B passing outside of the nebula in doing so.

During transit they will be attenuated as if they were passing through an extinguishing medium. Thus the nebula will be slightly less bright than it should be. In practice this should be a small effect in the structures considered since the Mie functions scatter least light at about 90^{o} and most importantly because the grains are extremely strongly forward throwing, so most light will continue to propagate into the nebula in an almost undeviated manner. For these reasons and to avoid computational inefficiency we shall continue under the approximation that this process is unimportant.

As photons travel to either A or C they will suffer extinction according to $\exp(d/\tau_{o})$ where d would be replaced by d_{A} or d_{C} respectively. Unlike the process above this cannot be ignored as it provides the primary illumination for the nebula. Consider two points on the nebular surface,

one three times the distance of the other from the star. The more distant point would receive a flux reduced by $^1/e^2$ (sevenfold) compared to the nearer point, purely as a result of this false extinction. In the calculations. which follow we consider some points as much as ten times removed. without correction the error on the intensity would be over eight-thousandfold. To solve this problem changes were introduced to the existing routines to calculate, from the propagation vectors, the intersection with the nebular boundary for the photons emitted directly by the source. Having done so it is a comparatively more simple matter to adjust the extinction paramters in order to "unextinguish" photons travelling through free space to the nebula. The improved simulation was tested against the old in the asymptotic condition where they tend to the same result: that is for a star just outside and just within a flat slab. The new model can also be tested directly in the same situations that the old could tackle, if it is given suitable boundary conditions. This can be done quite easily by constructing a subroutine that "informs" the model that the nebular boundary is a sphere about the star 100 metres in radius contained within some larger nebular shape. These tests showed the new simulation to be performing as anticipated. Thus armed with these simulation programs it was possible to begin modelling NGC7023.

5.3 APPLICATION TO NGC7023

5.3.1 Basic Modelling Considerations

We now apply the simulation programs to the business of modelling. Essentially this demands fitting two surfaces

simultaneously: intensity and polarisation. Since to satisfy one of these data sets does not necessarily imply a fulfillment of the demands of the other, we are now imposing very severe constraints upon the simulations. Initially however, we have a wide range of free parameters. We may vary the optical depth (i.e. the density) of the nebula. The composition of the scattering materials can be adjusted as can the size distribution functions of the grains. In the present case the evidence restricts that distribution to a power law, or an approximation to a power law, but we must still decide upon the index of the function. The choice of grain composition remains also and most importantly the geometrical structure of the nebula. In general an arbitrary structure will not fulfill the conditions imposed by the observations; this point seems obvious, for if any geometry could satisfy the data there would be no point in modelling. Unfortunately the geometry is not usually known, a fact which has led some authors (Zellner, 1973) to assume no geometry whatever, but simply to vary the length of a line of sight column. Even for single scattering this is an unsatisfactory procedure, but for multiple scattering it is quite impossible. On the other hand to test every conceivable structure is impossible, therefore some method to select for the most probable structures is required. As we know ab initio that NGC7023 has a high optical depth we shall describe a scheme appropriate to conditions where the nebula has τ equal to unity or greater. In such a circumstance, if we can see the source, we may presume:

(i) that the star is close to the front face of the nebula, and

(ii) that the more distant boundary may be considered to be at infinity.

This is a good approximation if $\tau > 1$ since the majority of the photons propagating away from the observer will be scattered before they leave the cloud. Given these conditions, though we may still be ignorant of even the most crude details of the surface geometry, we may now investigate the variations in polarisation and intensity as functions only of the shape of the nearer surface. This is of course, only a qualitive analysis, and is only true for a given set of the other important parameters: composition, density, etc., but it does place a very useful restriction on the problem.

5.3.2 The Plane Slab: A Starting Point

Before we begin modelling we should recall the important observations that must be satisfied. These are:

(1) the centrosymmetry of the entire polarisation pattern. This is proof that only one light source illuminates the cloud and thus simplifies the problem.

(2) the intensity pattern exhibits a monotonic decline with increasing radial distance, from HD200775, but the gradient differs with direction. We must explain both this asymmetry and reproduce the measured gradient.

(3) the percentage polarisation rises and falls with offset. We need to model the position of the peaks, their magnitude and the general behaviour.

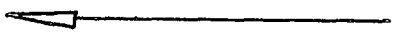
(4) the percentage polarisation is almost constant with wavelength. This observation we can explain already using a power law distribution.

We now consider the properties of a plane parallel slab of infinite depth under the approximations of section 5.3.1. Let us take graphite as a test material. The polarisation properties of graphite are given in Figure 4.6 for a wide range of power laws. In our simple analysis all we need consider are the geometrical constraints (the distance between the star and the front face of the nebula, the nebular tilt, etc.) and the form of this function over the range of scattering angles. Figures 5.3a and 5.3b show such a slab in two different configurations: (a) the most symmetrical case, (b) with a tilted front surface. A crude evaluation of this model for single scattering is quite easy. Imagine that the star, at the point 5, illuminates the nebula in a uniform manner. Light is scattered from dust particles in the cloud towards the observer. To estimate the form of the intensity or polarisation profiles we need only consider line of sight columns through the nebula produced by a distortion of the single scattering functions. The end points of these columns can be considered to be at infinity, i.e. scattering angle of 180^{o} and so from Figure 4.6 the polarisation is 0%. So the resultant integrated polarisation must fall between the value at the front surface, which can be simple read off from Figure 4.6 and zero. Initially we do not know the degree of depolarisation but we may regard it as a smoothly varying function of position on the nebular surface, since the scattering functions are themselves smooth. Now consider the integrated line of sight polarisations at the points A_o, A_1, A_2, etc. At the origin A_o, the angles of scattering are only 0^{o} and 180^{o}. Therefore the polarisation is o%, as

Fig. 5.3a

Plane Parallel Slab

Normal Incidence



to observer

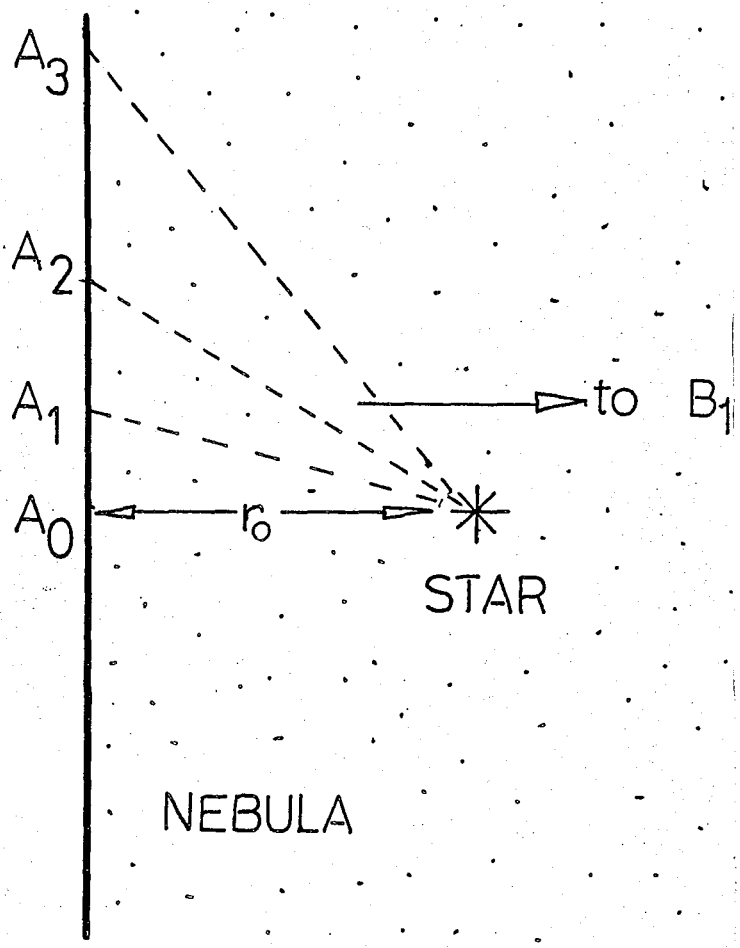
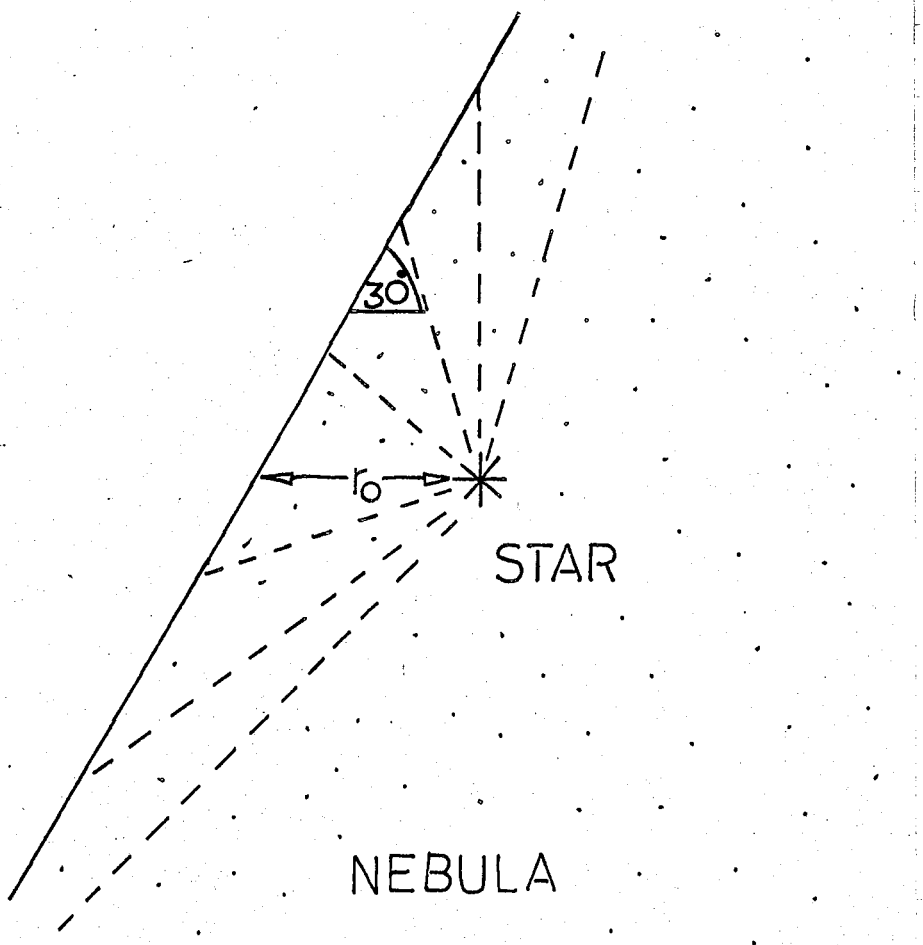


Fig. 5.3b

Tilted Slab

to observer



both these angles have the same value of zero percentage polarisation. Next at points close to, but slightly removed from the origin -A_1 say - we sample almost all possible angles from 0^o to 180^o, with a resultant non-zero percentage polarisation. However, as we move further away we begin to lose angles near to 0^o. At A_2 we have lost all angles less than 30^o. At greater distances still this process continues until asymptotically we shall lose all angles less than 90^o. For the effects of this angular selection we need only consult Figure 4.6. The polarisation rises from a minimum at 0^o to maximum near 90^o, and then falls to 0% at 180^o. Thus as we progressively cut out angles up to 90^o we must increase the integrated polarisation by reducing the less polarised contribution. The exact value of the maximum polarisation depends on the precise form of the scattering functions, but it is now clear that it will occur at an infinite radial distance from the star, perpendicular to the line of sight. For Figure 5.3a the general behaviour of the polarisation is simply a monotonic rise from zero to high values at great angular distances from the star. This pattern will be symmetrical about the source. Figure 5.4a shows the approximate polarisation behaviour expected under this analysis for various power indices for graphite. Now by altering the tilt angle of the front face (Figure 5.3b) we simply vary the range of angles between which the polarisation is to be integrated, and thus the asymptotic values. For a tilt of 30^o we would have external angles of $90^o \pm 30^o$, i.e. 60^o and 120^o. The predicted distribution is shown in Figure 5.4b. The intensity pattern is even simpler; we would expect intuitively

Fig 5.4b

Polarisation versus Radial Distance (normalised to $r_o = 1$)

Geometry as Fig. 5.3b

Graphite Dust

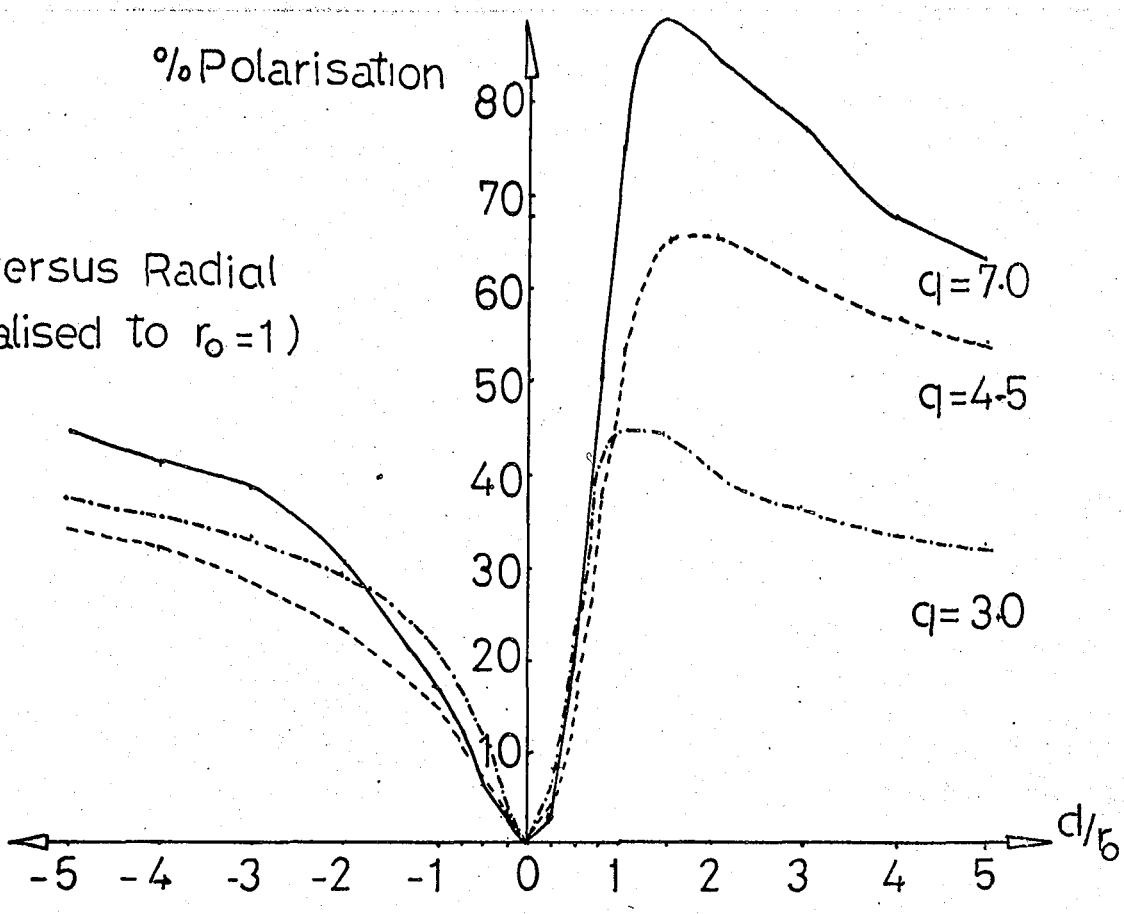
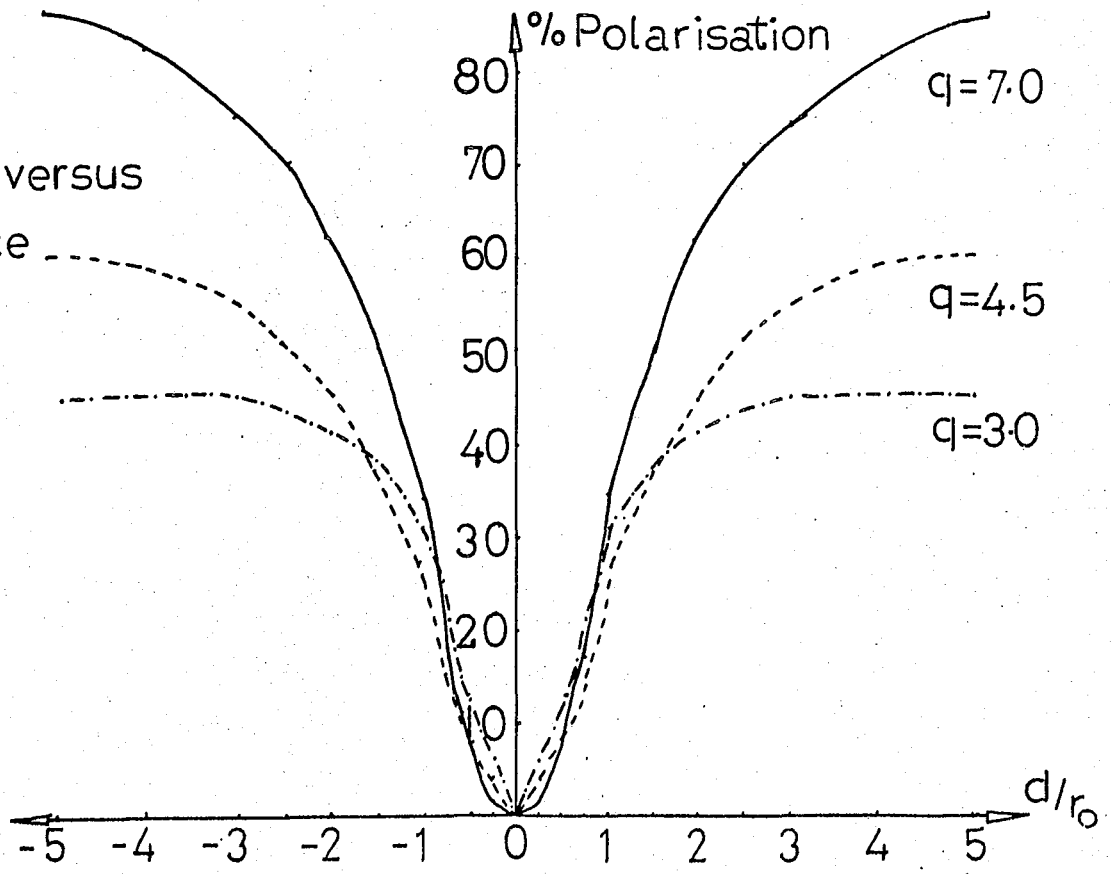


Fig. 5.4a

Polarisation versus Radial Distance (normalised to $r_o = 1$)

Geometry as Fig. 5.3a

Graphite Dust



an approximately linear reduction in intensity due to the $1/r^2$ law. This analysis is very important and it is true that additional materials or rotation of the plane can complicate the results substantially, but the main point is clear, a parallelepiped, irrespective of how it is turned will produce monotonic behaviour of the percentage polarisation with increasing angular displacement from the star, on at least one side of the nebula. In Figure 5.4b it might display a non-asymptotic polarisation peak on the side with an asymptotic scattering angle of 120^{o}. That behaviour would depend on the depth of the star from the nebular surface and also the material of the grains.

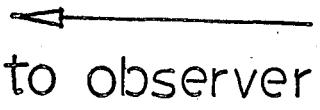
5.3.3 Convex Surface : Embedded Source

The plane surface fails to produce the necessary polarisation behaviour, but it does indicate how a solution may be obtained. The difficulty with the slab geometry is that asymptotically we are summing the polarisation functions only over angles distributed between 90^{o} and 180^{o}. Since most materials have their peaks near 90^{o} we are deliberately selecting for a polarisation pattern showing an increase with distance. So in order to produce a maximum polarisation nearer than at infinity we must restrict the range of angles sampled. Figure 5.5 shows how this can be done. If the front surface of the nebula is pulled backwards progressively with distance it is possible to chose whichever angles are required. At the point, P_1, on the surface the line of sight integration ranges from 0^{o}, to 180^{o}. Thus all angles less than 0^{o} are excluded. At P_2 angles smaller than θ_2^{o} are

Fig. 5.5

Embedded Source

Spherical Surface.



to observer

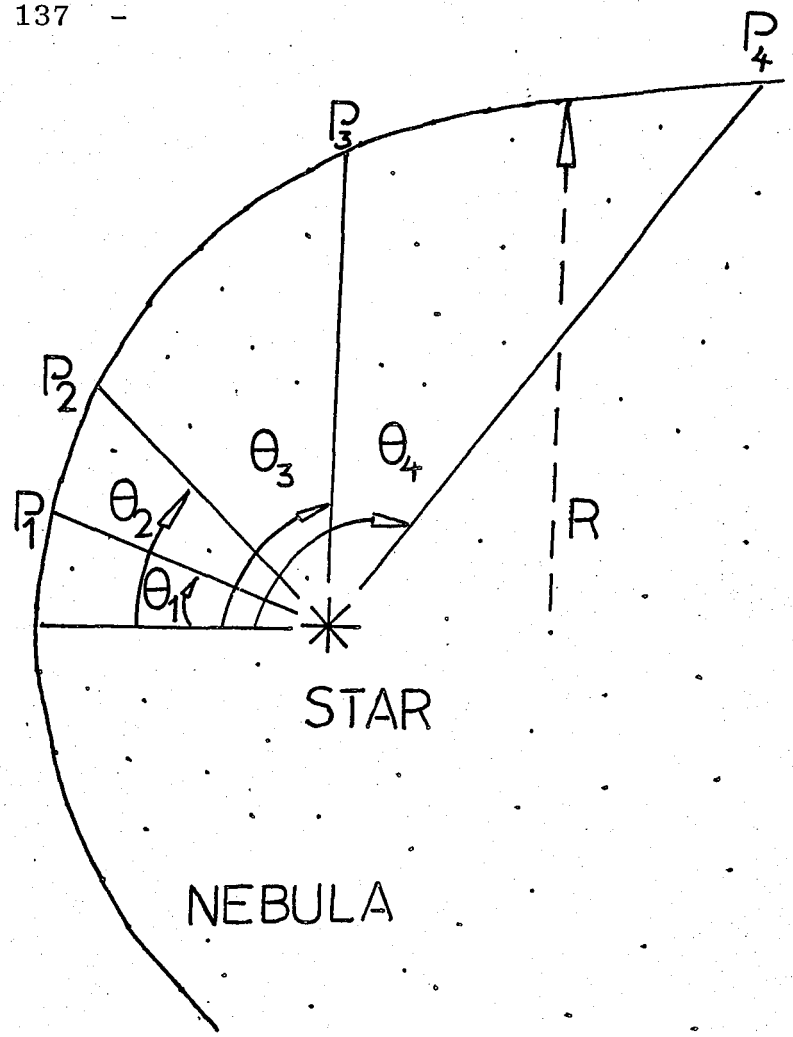
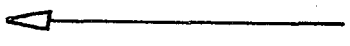
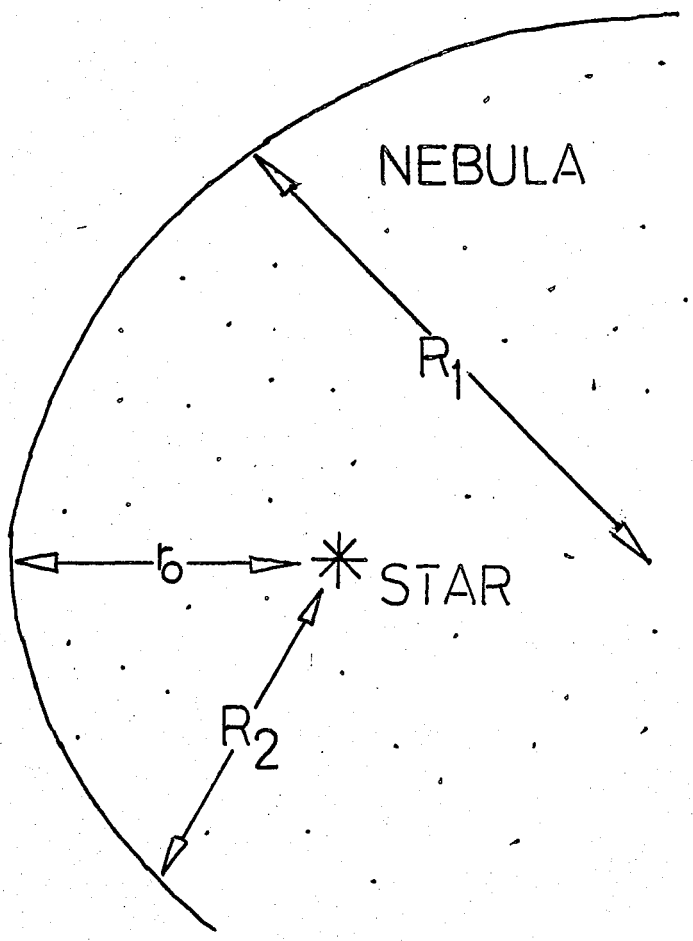


Fig. 5.6

Ovoid Structure



to observer



suppressed. At P_3 we have $\theta_3{}^o = 90^o$ and so we must have approximately peak polarisation for any material (see Figures 4.4 to 4.9).

Now however, beyond P_3 we have P_4. Here for the first time we can cut out angles between 90^o and 180^o. An examination of the polarisation functions for scattering in this region (Figures 4.4 to 4.9) shows that they all decline towards zero at 180^o. As we slice off angles near the peak of the functions we must reduce the integrated polarisation: so the polarisation rises at small angular distances from the star, peaks and then declines to some asymptotic value. The intensity however should still exhibit a monotonic fall off with distance. The slight asymmetry north-south seems to imply that an ellipsoidal structure could explain both polarisation and intensity.

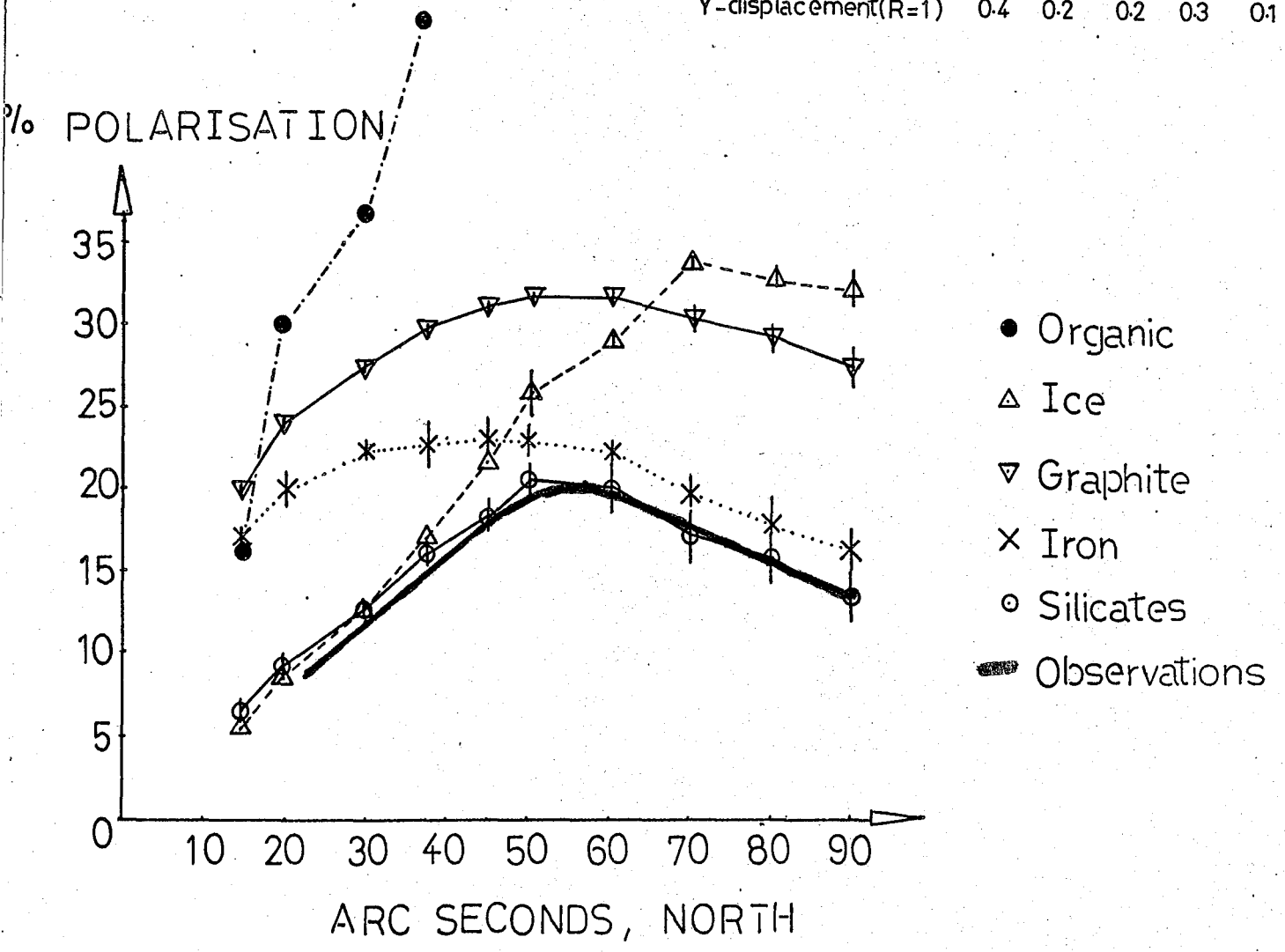
To investigate this hypothesis models have been constructed of ovoid structures composed of circular sections. The variable parameters are those that appear in every model: dust density, material composition, etc. and also the distinct geometrical variables, here: r_o, the distance from the star to the front surface and R, the curvature of the arcs. Integration along a line of sight column and multiple scattering both reduce the resultant maximum polarisation. In all the models considered a maximum number of scatterings of four was found to be sufficient to account for the effects of multiple scattering. Thereafter further scatterings produce only a small charge in the final behaviour of both polarisation and intensity and at a high cost in computer time. This is a result of the stochastic nature of the

process. The main effects of multiple scattering are to brighten the intensity distribution as a whole. Depending on the geometry it might alter the gradient of the fall off and for the polarisation a reduction in the measured values is noted at all positions. Variations in the number density of scatterers can modify the application of these effects, since density and the effective number of scatterings are obviously closely related. For these ovoidal models all the materials in chapter four were examined with several different power laws. For a particular material the various power indices usually preserve the general shape of the scattering functions but raise or lower the peak values of the curves. It is then simply a matter of scaling the effects of single scattering for one index, to those for another, with the results of multiple scattering in the model geometry to estimate the new multiple scattering result for the new power. Since the observations are almost symmetric north-south it is only necessary to attempt to fit one side of the nebula at a time. For if it proves impossible to match one side, then the model fails anyway. If on the other hand a fit is achieved it is then easy to extend it to the other side. This produces a large saving in computer time. With such a range of parameters some calculations produce peaks far removed from the desired position. Figure 5.7 therefore, presents only the best results for each material on the northerly side of the nebula. For most substances it is possible to produce a peak in the correct position and the summation is usually capable of fitting either the rising

Fig. 5.7 Polarisation v. Distance (North): Ovoid Geometry

° Radius of Arc = 9.15×10^{15} m

	Si	Ir	Gr	I	Or
Optical Depth ($\times 2.5 \times 10^{15}$ m)	1	100	100	10	100
Z-displacement (R=1)	0.63	0.58	0.58	056	0.8
Y-displacement (R=1)	0.4	0.2	0.2	0.3	0.1



or falling position of the polarisation profile, but to do all of these simultaneously is a very rigorous requirement which only silicates seem able to fulfil. The solution seems quite simple then; we need only model the other side of the nebula (Figure 5.8) in order to resolve the problem in favour of a silicate composition for the scattering particles. However, this is not so, for although it is possible to fit the polarisation pattern extremely well it is also necessary to produce the correct intensity behaviour. Chapter Three shows the intensity around HD200775. Contrary to the representation of Witt and Cottrell (1980), who characterised it as a power law function of the radial distance, it is a simple logarithmic function. The northerly slope is rather steeper than the southerly. A successful model should reproduce this asymmetry and the measured gradients. Figure 5.9 plots the simulated intensity on a logarithmic scale for comparison with the measured behaviour. It is noted immediately that the model intensities, although in the correct sense are characterised by far too severe a decline with increasing distance from the source. In fact none of the materials tested could produce simultaneously both the correct polarisation and intensity patterns, nor is any other material likely to do so. With hindsight it is possible to understand why from the basic scattering characteristics for single spheres. For intensity these functions are still very much size rather than material dependent. Figures 2.2 and 2.3 show the intensity functions for both silicates and ice. The most striking characteristic is the strong tendency towards forward scattering. Typically

Fig. 5.8 Polarisation v. Distance, Ovoid Geometry, Silicate Grains.

	NORTH	SOUTH
Radius of Arc	9.15×10^{15}m	1.37×10^{15}m
Z-displacement	5.75×10^{15}m	1.0×10^{15}m
Y-displacement	3.65×10^{15}m	4.0×10^{15}m

Optical Depth = 2.5×10^{15}m

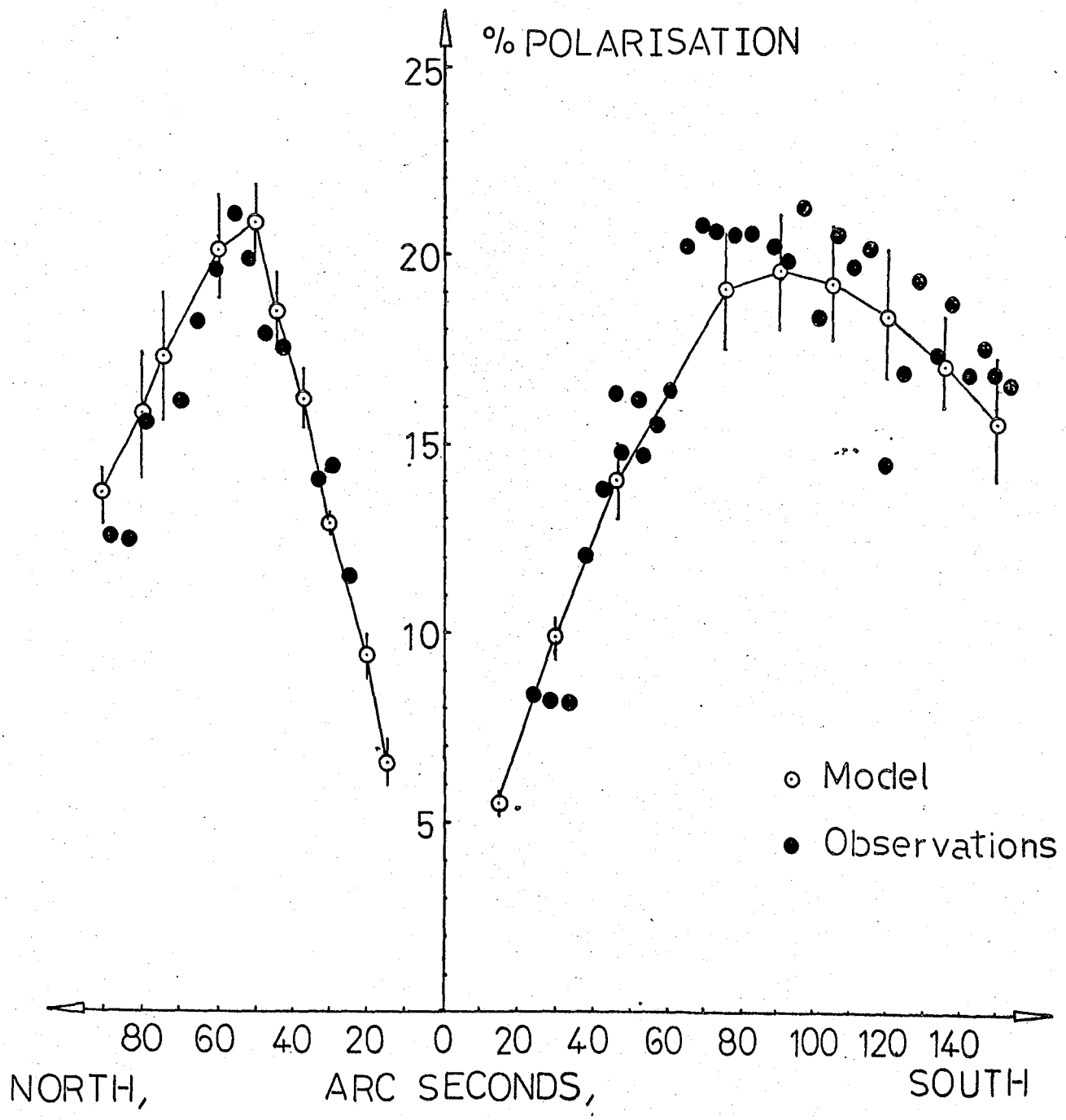
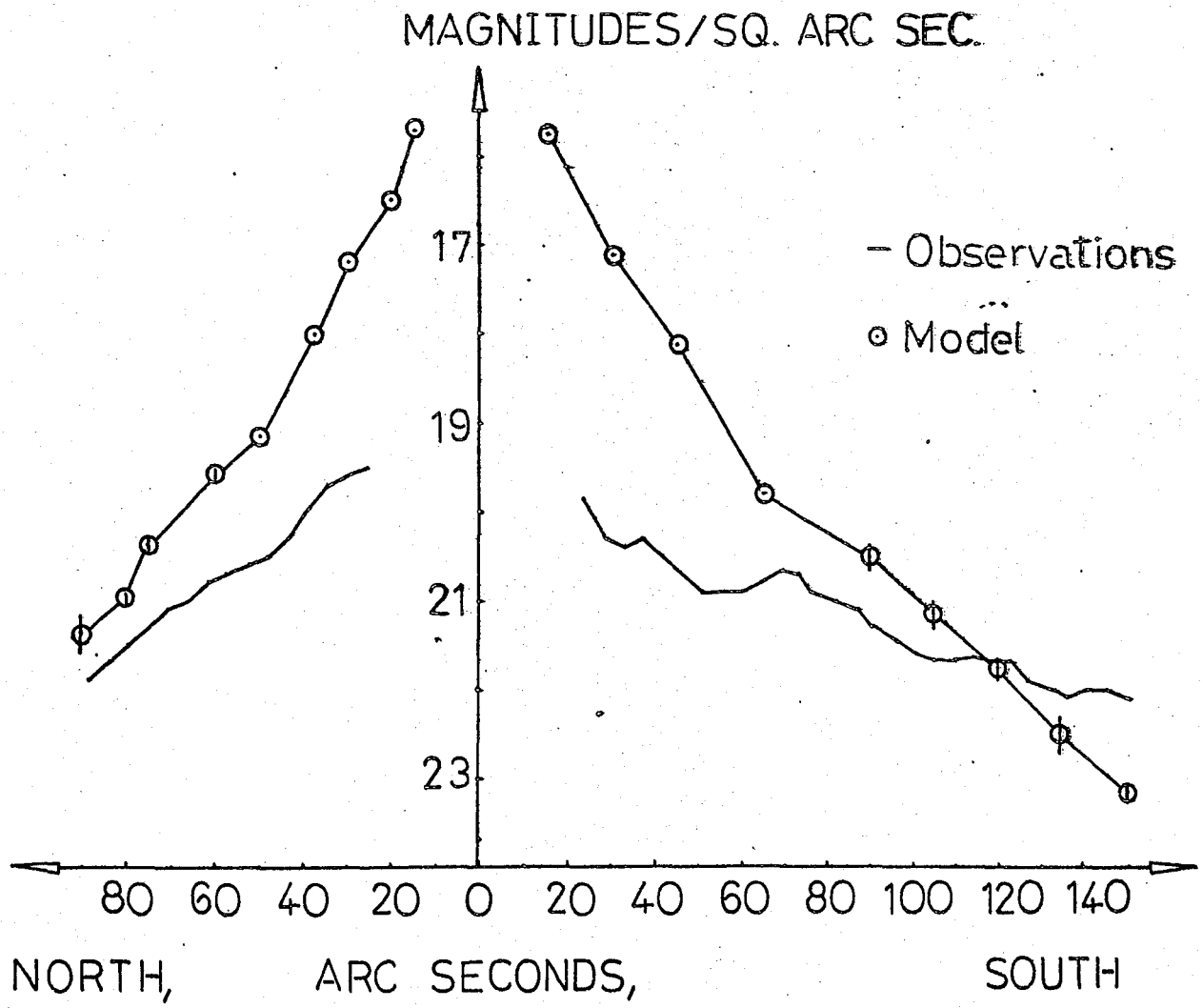


Fig. 5.9 BRIGHTNESS v. DISTANCE, OVOID GEOMETRY SILICATE GRAINS.

Optical Depth = 2.5×10^{15} m

	NORTH	SOUTH
Radius of Arc	9.15×10^{15} m	1.37×10^{15} m
Z-displacement	5.75×10^{15} m	1.0×10^{15} m
Y-displacement	3.65×10^{15} m	4.0×10^{15} m



the flux ratio between scattering at 0^o and 90^o is 150 to 250 fold. Since the maximum polarisation occurs in the vicinity of 90^o this ratio will be approximately that between the intensity near the star and that at the maximum polarisation, neglecting temporarily the effects of the $^1/r^2$ law (see Figure 5.10 for a cloud with a hemispherical front face and the further boundary at infinity). It might be argued that as integration proceeds along the line of sight that the lower intensity of the backscattering will be compensated for by the long integration length. However for an optically thick nebula the effective integration length is on the scale length of the optical depth, so that argument fails. One solution would be to advance the star close to the surface of the nebula. The problem already requires this as this would give the high degree of backscattering needed for the polarisation pattern. However, this tactic fails also since it only increases the relative degree of extinction suffered by the light from the backscattering particles vis a vis the forward scattering grains. In conclusion any convex structure with an embedded star, irrespective of grain composition, will fail to explain the observations of NGC7023.

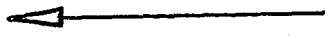
5.3.4 Concave Surface : Exterior Source

The analyses above indicate that no modification of the front surface can reproduce both the polarisation and intensity distributions simultaneously for a star embedded within the nebula. Yet since the object exists there must be a solution. The only remaining possibility is that the star is exterior to the cloud. To understand how this might work we consider Figures 5.11a, b and c. Figure 5.11a examines a plane parallel slab, as before, but now the star is between the observer and

Fig. 5.10

Hemispherical Surface

Infinite Depth



to observer

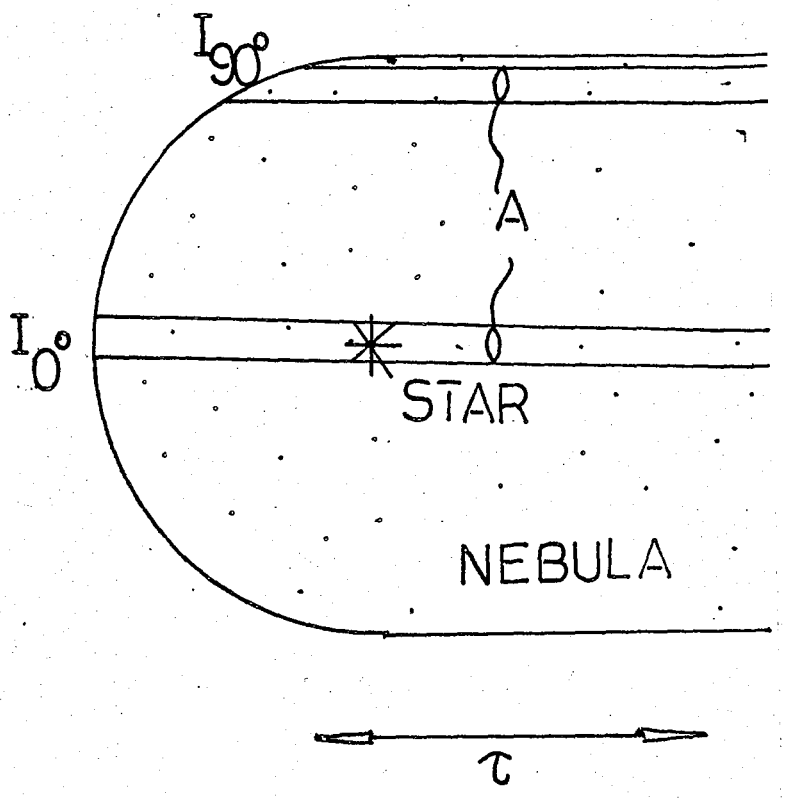


Fig. 5.11a

Semi-infinite Nebula

Plane Surface

External Star

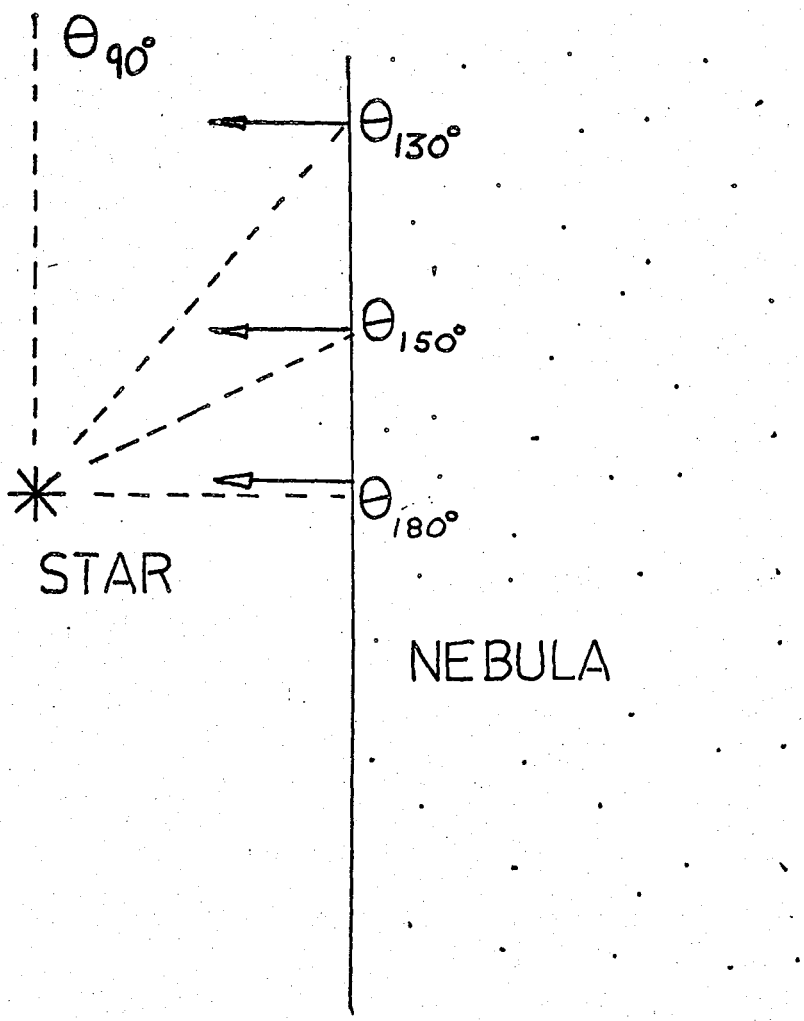
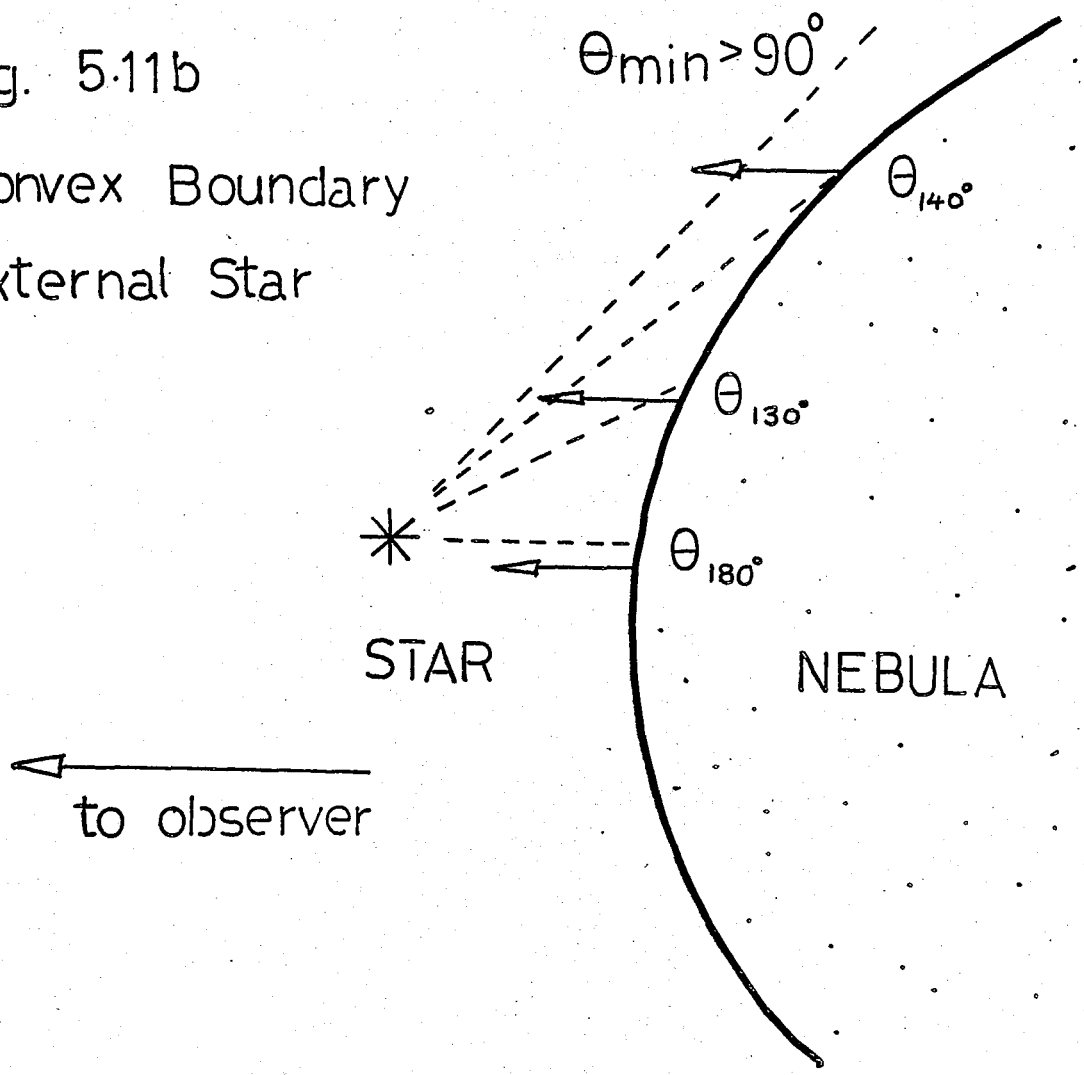


Fig. 5.11b

Convex Boundary

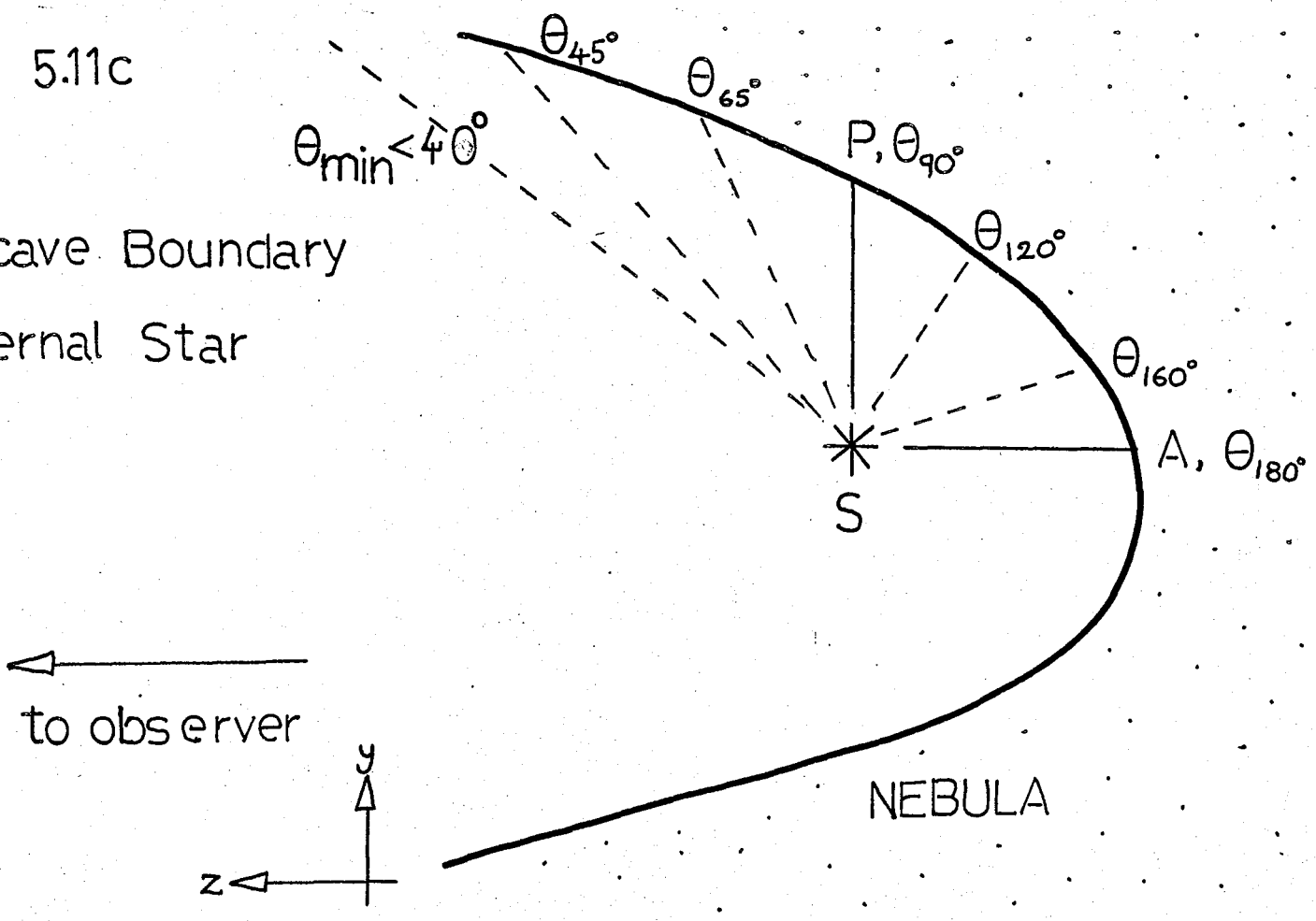
External Star



ig. 5.11c

oncave Boundary

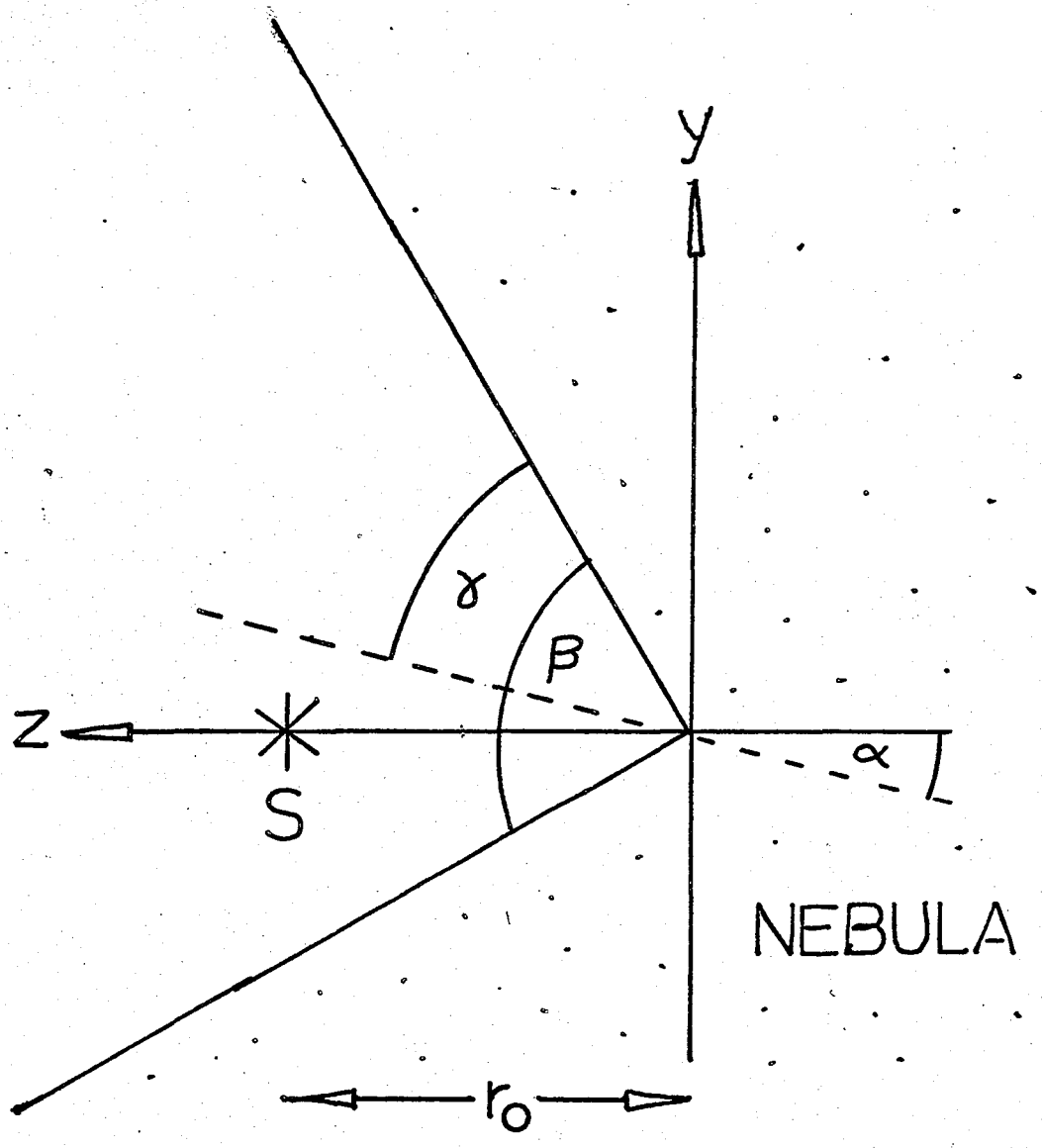
xternal Star



the cloud. Thus all radiation emitted by the nebula is backscattered towards the observer. We remove immediately the disparity in intensity, present in the embedded models, between points close to and distant from the source (in the plane of the sky) which result from the contrast between forward and backward scattering intensity. However for single scattering the maximal polarisation is still only obtained asymptotically, but for the opposite reason: the scattering angle is a minimum (i.e. 90^{o}) at an infinite radial distance. Since the polarisation for scattering from a collection of spheres is maximum near 90^{o} and declines towards 180^{o}, selection between 0^{o}, and 180^{o} will produce a growing integrated polarisation as 0^{o}, tends towards 90^{o}. Therefore some curvature must be applied to the surface. This can be done in two senses, either the boundary bends away from the observer (Figure 5.11b) as for the embedded cloud, or it flexes towards him (Figure 5.11c). The first condition must restrict the scattering angles to the range: $90^{o} < 0^{o} < 180^{o}$, where 0^{o} could have a minimum value substantially greater than 90^{o}. This angular restriction will again produce an asymptotic polarisation maximum. Essentially it is little different from 5.11a; the only differenc is that a lower value of polarisation will be noted at all radial distances. Finally consider Figure 5.11c. Here the range of θ is much extended: $0^{o} < \theta < 180^{o}$. Thus the model is free to sample angles both sides of the polarisation peak, for Mie spheres, ensuring that a non-asymptotic polarisation maximum will occur. The previous discussion re intensity still holds. Just conceivably if θ^{o} were sufficiently close

Fig. 5.12

Conical Cavity Geometry



to observer

to 0^{o} and the nebular surface close to the star the position might even be reversed and the nebula brighten with distance. This is of course, a physically implausible geometry. A more realistic problem is that models will exist where the distance SA is greater than SP. SA and SP are the directions parallel to the Z-axis and perpendicular to it respectively. Thus from the $^{1}/r^{2}$, relationship we could find conditions where the nebular intensity rises as we move away from the star, in the plane of the sky. An intensity peak would occur at the point P, after which there would be a further decline with distance. In these circumstances the polarisation maximum would be roughly coincident with the intensity peak. To avoid this situation, which is not observed, models are limited to SA < SP.

An attempt to produce the optimal structure, in terms of precise modelling of the fine details of the structure of the intensity or polarisation data, would demand adjustment of some very complex surface. This would give little real advantage over a more simple geometry that discovers the underlying geometry. Therefore the cavity in the front face of the nebula has been modelled as a circular cone with a variety of opening angles and tilts. A parabolic or hyperbolic shape would probably have served just as satisfactorily, perhaps even slightly more so very close to the star, where the sharp geometry of the apex of the cone is not physically likely. However, this structure contains the fundamentally important elements and makes a small saving in computer time: it is necessary to calculate the intersection of the photon path with the front surface of the nebula for every

Fig. 5.13 Polarisation v. Distance (South): Conical Geometry.

$SA = 2.0 \times 10^{15}$ m

	Si	Ir	Gr	I	Or
$\gamma°$	35.3	35.3	35.3	26.5	35.3

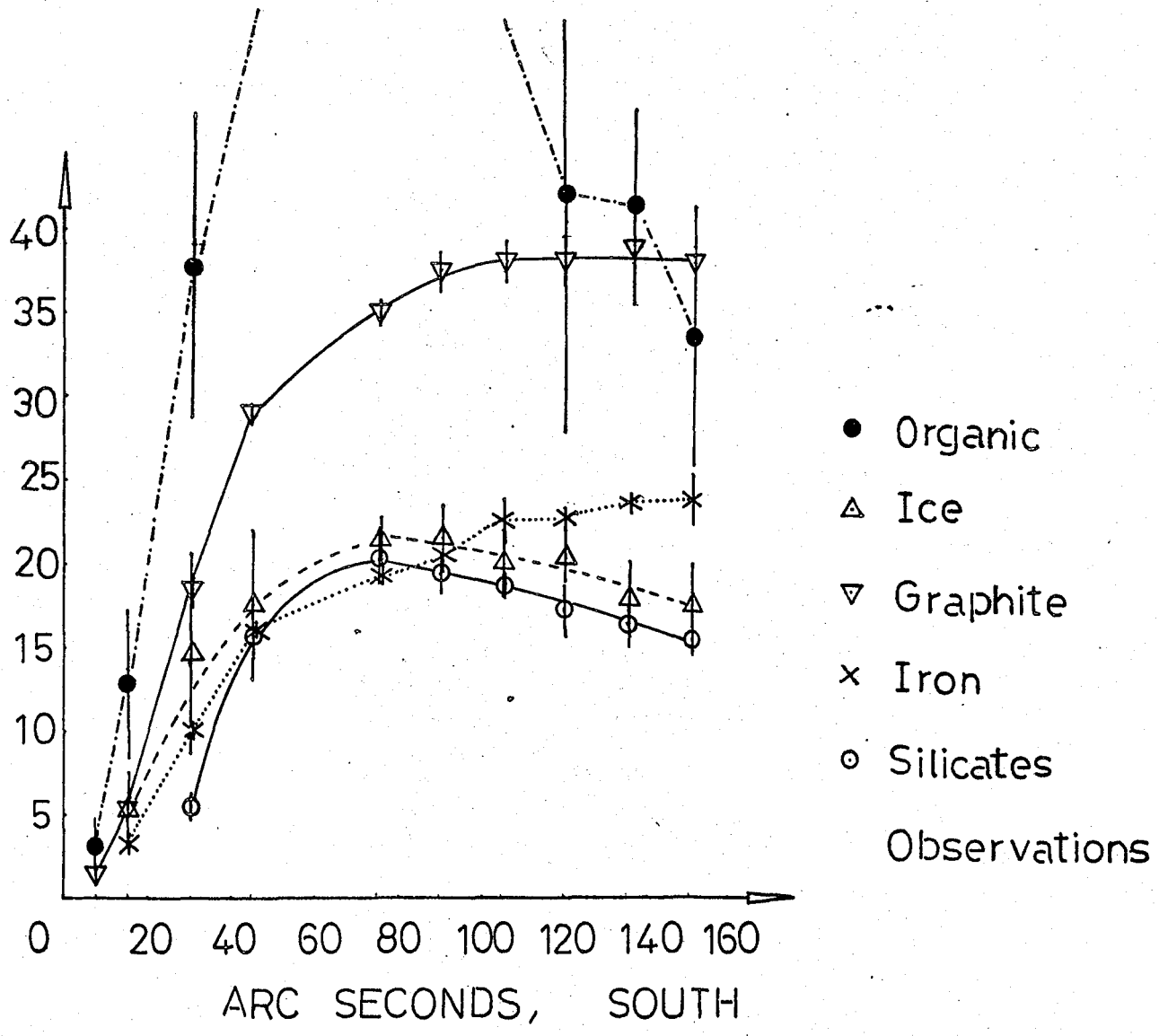


Fig. 5.14 Intensity v. Distance (South):

Conical Geometry.

$SA = 2.0 \times 10^{15}$ m

	Si	Ir	Gr	I	Or
$\gamma°$	35.3	35.3	35.3	26.5	35.3

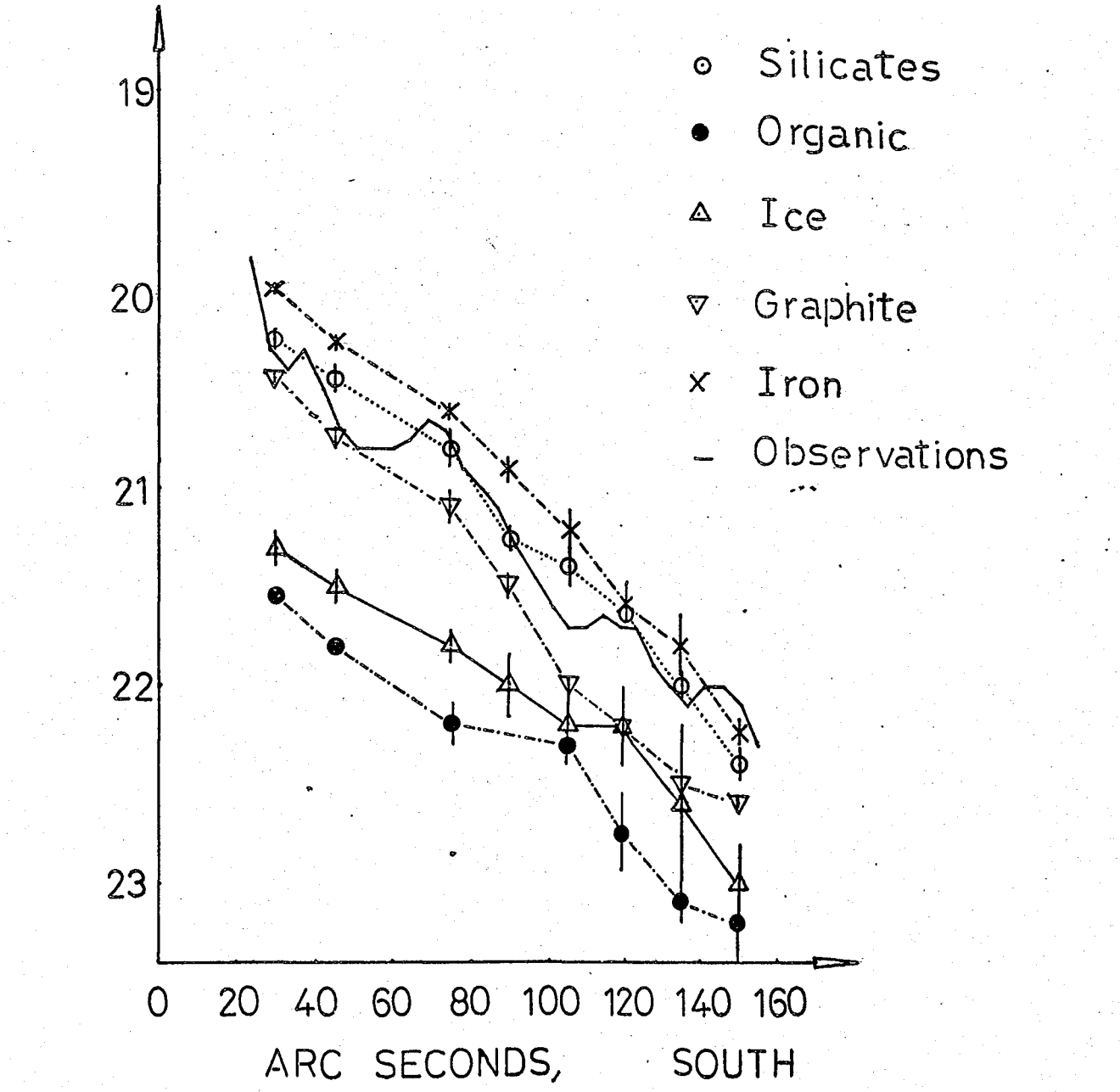


Fig. 5.15 Polarisation v. Distance:

Conical Geometry. $SA = 2 \times 10^{15}$ m

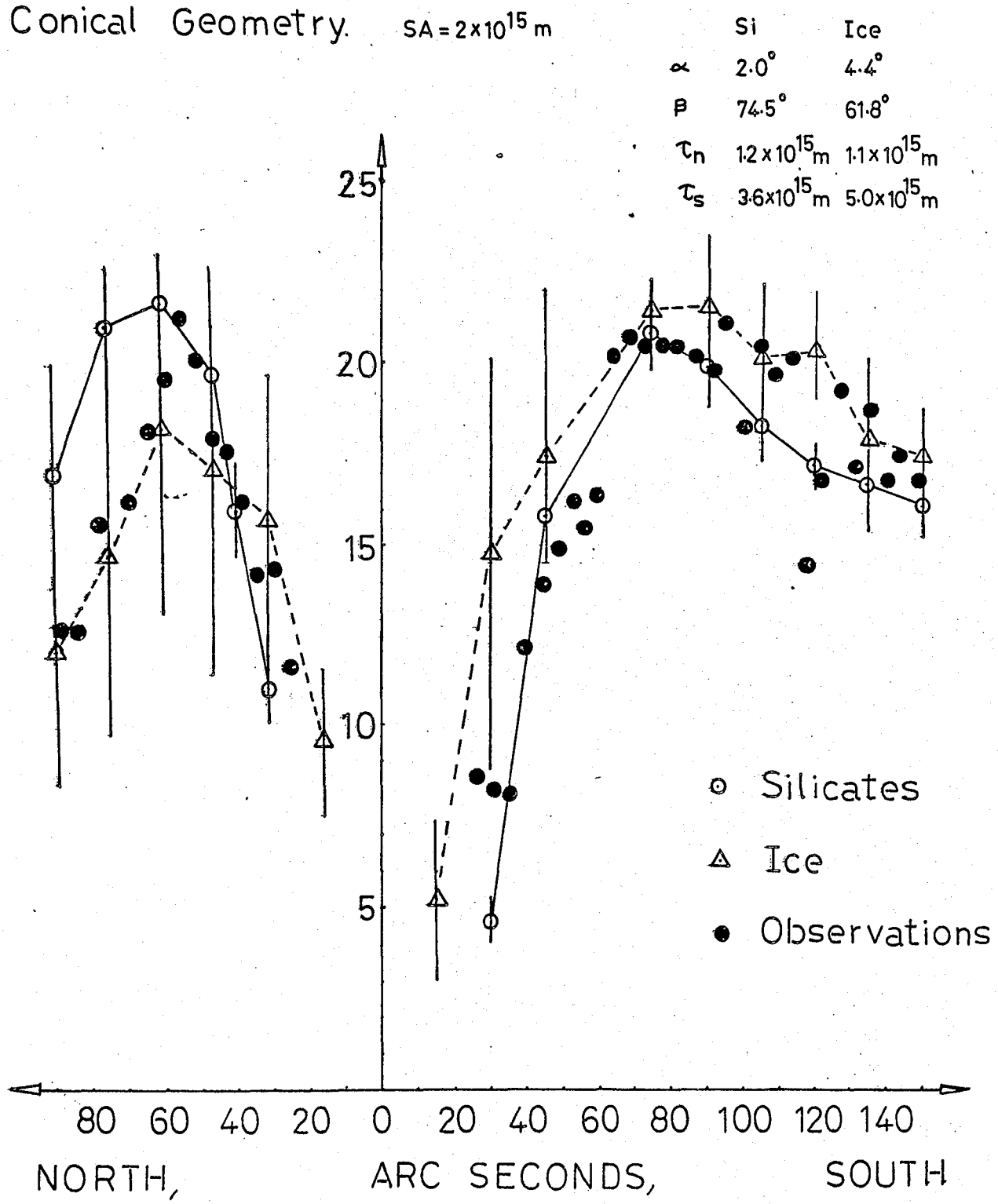
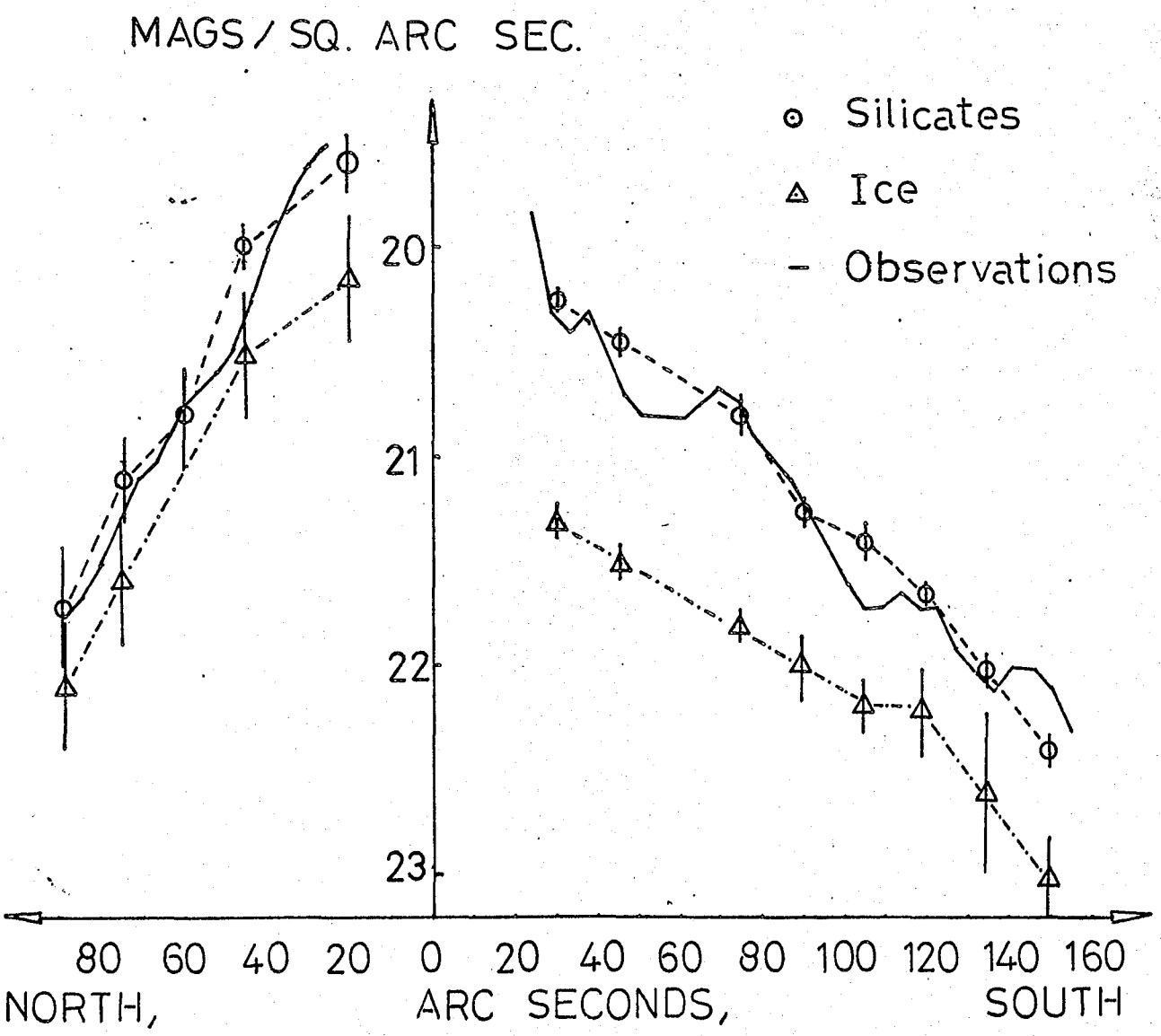


Fig. 5.16 Intensity v. Distance: Conical Geometry.

$SA = 2 \times 10^{15}$ m

	Si	Ice
α	2.0°	4.4°
β	74.5°	61.8°
τ_n	1.2×10^{15}m	1.1×10^{15}m
τ_s	3.6×10^{15}m	5.0×10^{15}m



photon, as they propagate into the cloud. The solution of linear equations are marginally faster than those for a quadratic surface. The "optimal surface" would make absurd demands on cpu time.

Given the hollow conical geometry the range of materials considered previously was again applied for a broad span of the possible orientations and conditions. The density of scatterers in the nebula, the power laws for the materials, the distance of the star to the apex of the cone, r_o, the opening angle of the cone, β, and the tilt angle, α, were all varied and the resulting intensities and polarisations examined. Figures 5.13 and 5.14 show the polarisation and intensity behaviour for the best fits for all these materials. It is very clear that only two materials can produce the observed polarisation and intensity values: silicates and ice. The best fit power law indices for silicates and ice respectively were -4.05 and -3.0, which straddle Mathis et al's. (1977) values (-3.6 to -3.3) obtained by fitting the interstellar extinction for a range of grain compositions. Figures 5.15 and 5.16 give model scans across the whole nebula compared to the measured results. Absolute calibration of the intensity has been calculated from the relationship between the nebular and stellar brightnesses (equation 5.6) using a visual magnitude for the star of 6.8 (van den Bergh, 1966). In the present geometry this becomes:

$$m_{neb} = m_* - 2.5 \log_{10}(R\ 4\pi D^2/\alpha^2) \qquad (5.12)$$

where R is the calculated intensity parameter and there is no extinction by the nebula for the star itself, since the

star is outside the nebula. Using the value for D = 350 pc (the value assumed for the scattering calculations) and $R = 0.73 \times 10^{-34}$ from the result of the simulation, for silicates we obtain: m_{neb} = 20.78 mag/sq. arcsec for the visual at 1' North. This compares with the value given by Vanysek (1963) of 20.77 mag./sq. arcsec. Vanysek's measurement is preferred to the more recent determination by Zellner (1969) since the latter quotes a result averaged over all directions at 0.5'. The intensity variations within the nebula make the use of such a mean unwise. The astonishing similarity of the calculated and measured values is more apparent than real, since Grygar's (1959) values are not consistent with Vanysek, being about 0.7 mag./sq. arcsec down, after accounting for the fall off with distance. Also the brighness of HD200775 varies so that could give a span of about 0.4 mags. to the value of m_* input to (5.12). Ice is consistently 0.5 mag./sq. arcsec dimmer than silicates thus weakening its case, but given the variations in the intensity determinations it cannot be ruled out.

Returning to the discussion of polarisation: biological materials and graphite are obviously excluded, in that their predicted polarisation values even in the most favourable conditions are still far too high. This is a consequence of two factors (1) depolarisation by multiple scattering is not very effective, and (2) the Mie scattering functions for these materials, for the wide range of power laws considered, have high intrinsic polarisations (c.f. chapter four). Iron too can be rejected, for although it can fit the increase of polarisation near the star, it does not peak

at all near the required point. Both silicates and ice produce acceptable results, and both suffer from slight but opposing defects. Silicates tend to fall a little too sharply on nearing the star, while ice has the reverse behaviour, being marginally too polarised near the source. The northerly peak is calculated to be marginally high. The final model has only a slight structural asymmetry. The opening angle β, e.g. for silicates, is 74.5^{o}, while α the tilt angle of the cone is only 2^{o}. Table 5.1 compares the values for both materials of r_{o},α,β and the number density of grains. One difficulty in the fitting was that no single density seemed acceptable for the entire nebula, given the well known density fluctuations in the structure this is hardly surprising (Weston, 1953; Gehrels, 1967; Ney et al., 1980). However the simulation programs are unable to handle this condition directly. It was resolved by modelling first one side of the structure and then the other. Variations in the opening angles or the tilt were unable to produce the same intensity modifications as the density variation. The main complication in the testing of the models was that r_{o}, α, β and τ (the optical depth - for a given material and size distribution inversely proportional to the number density of scatterers) are all closely interlinked in their effects. For example, if the distance between the star and the apex of the cone (Figure 5.12) is shortened this will brighten the nebula near to the star. However, it will also change the angular spread for scattering and thus alter the polarisation pattern. To readjust the polarisation behaviour β of α could be varied, but again this would affect the intensity characteristics.

Table 5.1 Conical Cavity Model Characteristics

Distance of Star from Apex(r) : 2×10^{15}m

	Silicates	Ice
Tilt Angle, α	2^{o}	4.4^{o}
Opening Angle, β	74.5^{o}	61.8^{o}
Optical Depth (North), τ_n	1.2×10^{15}m	1.1×10^{15}m
Optical Depth (South), τ_s	3.6×10^{15}m	5.0×10^{15}m
No. of Scatterers (North), N_n	$0.48\ m^{-3}$	$0.08\ m^{-3}$
No. of Scatterers (South), N_s	$0.16\ m^{-3}$	$0.02\ m^{-3}$

So far while modelling we have considered only the structure dictated by fitting the measured polarisation and intensity distributions. We have not considered how the final geometry might arise or whether it is physically probable. The earlier conditions for stars embedded within clouds hardly need explanation, but this new complex structure should be examined. Already, in chapter two, we have discussed the "bubble" or "blow-hole" models of Castor et al. (1975) and Habing and Israel (1979). Consider a star forming on the edge of a cloud. As it warms it begins to expell matter, in all directions. At the same time it illuminates the surrounding nebula. Being near the surface, the pressure of the expanding gas is sufficient to break through exposing the cavity. We can then develop a hyperbolic structure such as 5.111b. for the nebula in the vicinity of the star. The conical structure is just a rough approximation to this shape. Obviously the dense nebula behind the star cannot be dispelled, so the end result is the star residing in its own cavity, illuminating the nebula to the sides and behind.

One final point: the polarisation and intensity plots given are for visual light. Calculations for red and blue light were also performed to investigate the λ-dependence. The general form of these calculations is so similar as to be almost indistinguisable through the random errors inherent in a stochastic calculation. However observations indicate a general spread of polarisation of about 3% between red and blue. Figure 2.7 & Fig. 2.9 though there is considerable scatter on the measurements. The sense is such that red has the higher values. For the power laws investigated although

the sense was correct the charge was only 0.5% approximately at the peak polarisation - many repetitions of the calculation were performed at this point to reduce the random errors in order to see this difference. Now for an OHG distribution the polarisation charge with wavelength would also be towards increasing polarisation with longer wavelengths, but the effect would be vastly greater: several tens of percent. Thus the evidence hints that although the power law function is a good approximation to the underlying distribution of sizes it is not yet quite the correct representation. Some function which is slightly flatter, but not nearly so flat as the other distribution, may be the solution.

5.4 CONCLUSIONS

New observations of polarisation and intensity at three wavelengths have extended the observational data on NGC7023. Simulations to explain both these and existing measurements have been tested for a wide variety of materials and geometries. All structures which contain an embedded star have been shown to fail. Of those geometries which feature a star outside the nebula, only one broad category can reproduce the polarisation and intensity noted. This is of the general form of a cavity extending into the nebula. The functional dependence of the size distributions of grains has also been investigated. The flatish distributions most prominently the Oort-van de Hulst-Greenberg function are found to fail badly in their wavelength behaviour for the polarisation data. Power law distributions are found to be much more satisfactory, but not quite perfect.

Monte Carlo calculations for an improved simulation have narrowed the range of materials likely to comprise the nebular dust in NGC7023 to two favoured candidates: silicates and ice. Of these silicates are the stronger contender producing better agreement with published nebular brightness determinations. The present tilted cone models have a high degree of symmetry in their structure, as would be anticipated from the observations. The reflection nebula is now to be understood in terms of a burst "blister cavity" around HD200775, on the edges of NGC 7023. The star is sited within the cavity. This model is also consistent with the best infra-red data available (Whitcomb et al., 1981) and seems to resolve the inconsistencies in the interpretations of other optical observers (Witt and Cottrell, 1980a).

REFERENCES

Aannestad, P.A.; 1973. Ap. J. Sup. Ser., 25, 205.

Aannestad, P.A., Purcell, E.M.; 1973. Ann. Rev. Astron. & Ap., 11, 309.

Allen, C.W.: 1963. Astrophysical Quantities, Second Ed., Athlone Press, London.

Aller, L.H.; 1961. The Abundance of the Elements, Interscience Publishers, New York.

Ambartsumyan, V.A., Gondeladze, Sh.G.; 1938. Bull. Abastumansk. Astron. Observ. 2, 37.

Andriesse, C.D., de Vries, J.; 1974. Astron. & Astrophys. 30, 51.

Andriesse, C.D., Piersma, T.R., Witt, A.N.; 1977. Astron. & Astrophys. 54, 841.

Arnett, W.D.; 1969. Ap. J. 157, 1369.

Aveni, A.F., Hunter, J.H.; 1967. Astron. J. 72, 1019.

Axon, D.; 1977. Ph.D. Thesis, University of Durham, England.

Baade, W.; 1944. Ap. J., 100, 137.

Barlow, M.J.; 1978a. Mon. Not. R. astr. Soc., 183, 367.

Barlow, M.J. 1978b. Mon. Not. R. astr. Soc., 183, 397.

Barlow, M.J.; 1978c. Mon. Not. R. astr. Soc., 183, 417.

Bar-Nun, A.; 1975. Ap. J., 197, 341.

Bergh, S. van den; 1966. Astron. J., 71, 990.

Bless, R.C., Savage, B.D.; 1972. Ap. J., 171, 293.

Bodenheimer, G., Tenorio-Tagle, Yorke, H.W., 1979; Ap. J., 233, 85.

Bohlin, R.C., Savage, B.D., Drake, J.F.; 1978. Ap. J., 224, 132.

Burke, J.R., Silk, J.; 1974a, Ap. J., 190, 1.

Burke, J.R., Silk, J.; 1976, Ap. J., 210, 341.

Cameron, A.G.W., Colgate, S.A., Grossman, L., 1973; Nature, 243, 204.

Carrasco, L., Strom, E.M., Strom, K.M.; 1973, Ap. J. 182, 95.

Caroff, L.J., Petrosian, V., Salpeter, E.E., Wagoner, R.V., Werner, M.W.; 1973, Mon. Not. R. astr. Soc., 164, 295.

Castor, J., McCray, R., Weaver, R., 1975; Ap. J., 200, L107.

Celnikier, L.M., Lefevre, J.; 1974, Astron. & Ap., 36, 429.

Cederblad, S.; 1946, Lund. Obs. Medd. Ser.II, No. 119.

Chiao, R.Y., Feldman, M.J., Parrish, P.T.; 1973, I.A.U., Symposium, No. 52, Albany, New York.

Chylek, P., Grams, G.W., Pinnick, R.G.; 1976, Science, 193, 480.

Code, A.D., Davis, J., Bless, R.C., Hanbury Brown, R., 1976; Ap. J., 203, 417.

Cohen, M.; 1974, Pub. Astron. Soc. P., 86, 813.

Cohen, M.; 1982. Pub. Astron. Soc. P., 94, 266.

Cohen, M., Kuhi, L.V., Harlan, E.A., Spinrad, H.; 1981. Ap. J., 245, 920.

Collins, O.C.; 1938. Ap. J., 86, 529.

Coyne, G.V., Gehrels, T., Serkowski, K.; 1974. Ap. J., 79, 581.

Cuderman, J.F., Brady, J.J., 1968; Surface Sci. 10, 410.

Dahn, C., 1967; Ph.D. Thesis, Case Institute of Technology.

Daniel, J.-Y., 1978; Astron. & Ast., 67, 345.

Daniel, J.-Y., 1980; Astron. & Ast., 87, 204.

Daniel, J.-Y., 1982; Astron. & Ast., 111, 58.

Danielson, R.E., Woolf, N.J., Gaustad, J.E., 1965; Ap. J., 141, 116.

Day, K.L., 1974; Ap. J., 192, L15.

Day, K.L., 1976; Ap.J., 210, 341.

Day, K.L., Steyer, T.R., Huffman, D.R., 1974; Ap. J., 191, 415.

Day, K.L., Donn, B., 1978; Ap. J., 222, L45.

Deirmendjian, D., 1969; Electromagnetic Scattering on Spherical Polydispersions, Elsevier, New York.

Debye, P., 1909; Ann. Phys., Lpz. 30, 59.

Dickman, R.L. 1976; Bull. Ann. astr. Soc., 8, 513.

Donn, B., Wickramasinghe, N.C., Hudson, J.P., Stecker, T.P., 1966; Astron. J., 71, 853.

Donn, B., 1968; Ap. J., 152, L129.

Dorschner, J., 1971; Astron. N., 293, 65.

Dorschner, J., Gurtler, J., 1966; Astron. N., 289, 57.

Draine, B.T., 1979; Ap. J., 230, 106.

Duinen, J.R. van., Wu, C.C., Kester, D., 1976; Department of Space Research Groningen Internal Note, ROG NR 76-4.

Duley, W.W., 1970; J.R.A.S. Canada, 64, 33.

Duley, W.W., McCullogh, J.D. 1977, Ap. J. Lett., 211, L145.

Duley, W.W., Miller, T.J., 1978; Ap. J., 220, 124.

Ebert, R., 1957, Z. Astrophys. 42, 263.

Elmegreen, B., 1980; Giant Molecular Clouds in the Galaxy, Pergamon, London.

Elmegreen, D., Elmegreen, B., 1978; Ap.J., 220, 510.

Elvius, A., Hall, J.S., 1965, Astron. J., 70, 138.

Elvius, A., Hall, J.S., 1966; Lowell Obs. Bull, VI, 257.

Falk, S., Lattimer, J.M., Margolis, S.H., 1977; Nature, 270, 700.

Field, G., 1974; Ap. J., 187, 453.

Fitzgerald, M.P., Stephans, T.C., Witt, A.N., 1976; Ap. J., 208, 709.

Friedemann, C., Gurtler, J., Dorshner, J., 1979; Astro. & Space Sci., 60, 297.

Gallagher, J.S., 1977; Astron. J., 82, 209.

Garrison, L.M., 1978; Ap. J., 224, 535.

Gehrels, T., 1967; Astron. J., 72, 631.

Gehrels, T., 1972; Ap. J. Lett., 173, L23.

Gehrels, T., 1974; Astron. J., 79, 590.

Gething, M.R., Warren-Smith, R.F., Scarrott, S.M., Bingham, R.G., 1982; Mon. Not. R. astr. Soc. 198, 881.

Gillet, E.C., Stein, W.A., 1971; Ap. J. 164, 77.

Gillet, F.C., Jones, T.W., Merrill, K.M., Stein, W.A., 1975; Astron. & Astr., 45, 77.

Gilra, D.P.; 1971; Nature, 229 , 238.

Gilra, D.P., 1973; I.A.U. Symposium, No. 52, Albany, New York.

Gilman, R.C., 1969; Ap. J., 155, L185.

Gliese, N., Walter, K., 1951; Z.A., 29, 94.

Grasdalen, G.L., Strom, K.M., Strom, S.E., Ap. J., 184, L53.

Greenberg, J.M., 1960; Ap. J., 132, 672.

Greenberg, J.M., 1968; Stars and Stellar Systems, Chapter 6, University of Chicago Press.

Greenberg, J.M., 1973; I.A.U. Symp. No. 52, Albany New York.

Greenberg, J.M., Hanner, M., 1970; Ap. J., 161, 947.

Greenberg, J.M., Hong, S.S., 1974; Planets, Stars and Nebulae Studied with Photopolarimetry, University Arizona Press, Tucson.

Greenstein, J.L., Aller, L.H., 1947, Pub. Astr. Soc. P., 39, 139.

Grygar, J., 1959; Acta Univ. Carol. Ser. Math. et Phys., 1.

Habing, H.J., Israel, F.P., 1979, Ann. Rev. Astron. & Ast. 17, 345.

Hall, R.C., 1965; Pub. Ast. Soc. P., 77, 158.

Hanner, M.S., 1969; Ph.D. Thesis, Remsselaen Polytechnic Institute, Troy, New York.

Hanner, M.S., 1971; Ap. J., 164, 425.

Hartmann, W.K., 1969, Icarus, 10, 201.

Haslam, C.G.T., Hills, D.L., Mathews, H.E., Salter, C.J., 1978; Astron. & Ast., 70, 575.

Hayes, D.S., Mavko, G.E., Radick, P.R., Rex, K.H., Greenberg, J.M., 1973; I.A.U. Symp. No. 52, Albany, New York.

Hellyer, B., 1970; Mon. Not. R. astr. Soc. 148, 383.

Henyey, I.G., 1936; Ap. J., 84, 609.

Henyey, I.G., Greenstein, J.L., 1938; Ap. J. 87, 580.

Herbig, G.H., 1960; Ap. J. Suppl., 4, 337.

Hiltner, W.A., Hall, J.S., 1949; Science, 109, 165.

Hong, S.S., Greenberg, J.M.; 1978; Astron. & Ast. 69, 341.

Hong, S.S., Greenberg, J.M., 1980; Astron. & Ast., 88, 194.

Houten, C.J. van., 1961; B.A.N., 16, 509.

Hoyle, F., Wickramasinghe, N.C., 1962; Mon. Not. R. ast. Soc. 124, 417.

Hoyle, F., Wickramasinghe, N.C., 1970a, Nature, 226, 62.

Hoyle, F., Wickramasinghe, N.C., 1970b, Nature, 227, 473.

Hoyle, F., Wickramasinghe, N.C., 1979, Preprint.

Hubble, E., 1922; Ap. J., 56, 162.

Hubble, E., 1922; Ap. J., 56, 400.

Hubble, E., 1922; Ap. J., 56, 416.

Hulst, H.C. van de, 1957, Light Scattering by Small Particles, New York, Wiley.

Icke, V., 1979; Astron. & Ast., 78, 352.

Icke, V., Gatley, I., Israel, F., 1980; Ap. J., 236, 808.

Innanen, K.A., 1969; J.R.A.S. Canada, 63, 260.

Israel, F.P., 1978; Astron. & Ast., 70, 769.

Johnson, H.M., 1968, Nebulae and Interstellar Matter, Ed. Middlehurst & Allen, University of Chicago Press, Chicago.

Junge, C.E., 1963; Air Chemistry and Radioactivity, New York, Academic Press.

Kahn, F.D., 1974; Astron. & Ast., 37, 149.

Kandel, R.S., Sibille, F., 1978, Astron. & Ast., 68, 217.

Kaplan, S.A., Pikelner, S.B., 1970; The Interstellar Medium, Harvard University Press, Cambridge, Mass.

Keenan, P.C., 1936; Ap. J., 84, 600.

Kemp, J.C., Wolstencroft, R.D., 1972; Ap. J. Lett., 176, L111.

Kenknight, C.E., Wehner, G.K., 1964; J. Appl. Phys., 35, 322.

Khan, F.D., 1974; Astron. & Ast., 37, 149.

Kleinman, S.G., Joyce, R.R., Sargent, D.G., Gillet, F.C., Telesco, C.U., 1979; Ap. J., 227, 126.

Knacke, R.F., Cudaback, D.D., Gaustad, J.E., 1969a, Ap. J., 158, 151.

Knacke, R.F., Gaustad, J.E., Gillet, F.C., Stein, W.A., 1969b, Ap. J. Lett., 155, L189.

Knacke, R.F., 1980; Protostars & Planets, Ed. T. Gehrels, University of Arizona Press, Tucson.

Knapp, G.R., Brown, R.L., Kuiper, T.B.H., 1975; Ap. J., 196, 167.

Kratschmer, W., Huffman, D.R., 1979; Astrophys. & Space Sci., 61, 195.

Krishna Swamy, K.S., O'Dell, C.R., 1967; Ap. J., 147, 937.

Krishna Swamy, K.S.; Wickramasinghe, N.C., 1968a, Nature, 220, 896.

Krishna Swamy, K.S.; Wickramasinghe, N.C., 1968b, Nature, 217, 1236.

Larson, R., 1969a; Mon. Not. R. astr. Soc., 145, 271.

Larson, R., 1969b; Mon. Not. R. astr. Soc., 145, 297.

Larson, R., 1972; Mon. Not. R. astr. Soc., 156, 437.

Lee, T. 1968; Ap. J., 152, 913.

Lepine, J.R.D., Nguyen-Quang-Rieu, 1974; Astron. & Ast., 36, 469.

Lindblad, B., 1935; Nature, 135, 133.

Loren, R.B., van den Bout, P.A., Davis, J.H., 1973; Ap. J., 185, 167.

Mathis, J.S., 1972; Ap. J., 176, 651.

Mathis, J.S., 1979; Ap. J., 232, 741.

Mathis, J.S., Rumpl, W., Nordsieck, K.H., 1977, Ap. J., 217, 425.

Martel, M.T., 1958; Suppl. Ann. d'Ap., No. 7.

Martin, P.G., 1974; Ap. J., 187, 461.

Martin, P.G., 1978; Cosmic Dust, Its Impact on Astronomy, Clarendon Press, Oxford.

McDonnell, J.A.M.; 1978; Cosmic Dust, Wiley, Chichester.

Mendoza, E.E., 1958; Ap. J., 128, 207.

Merrill, K.M., Russell, R.M., Soifer, B.T., 1976; Ap. J., 207, 763.

Meyer, W.F., 1920; Lick Obs. Bull, 10, 68.

Mie, G., 1908; Ann. d. Phys. 25, 377.

Millar, J., Duley, W., 1978; Mon. Not. R. astr. Soc., 183, 177.

Milkey, R.W., Dyck, H.M., 1973; Ap. J., 181, 833.

Minin, I.N., 1962, Soviet Astron. A.J., 5, 487.

Morton, D.C., 1974; Ap. J., 193, L35.

Nandy, K., Thompson, G.I., Jarman, C., Monfils, A., Wilson, R., 1975; Astron. & Ast., 44, 195.

Ney, E.P., Hatfield, B., Gerhz, R.D., 1980; Proc. Nat. Acad. Sci. U.S.A., 77, No. 1, 14.

O'Dell, C.R., 1965; Ap. J., 142, 604.

Oort, J.H., Van de Hulst, H.C., 1946; B.A.N., 10, 187.

Osterbrock, D.E., 1974; Astrophysics. of Gaseous Nebulae, W.H. Freeman, San Francisco.

Pallister, W.S., 1976; Ph.D. Thesis, University of Durham, England.

Penman, J.M., 1976, Mon. Not. R. astr. Soc., 175, 149.

Perkins, H.G., King, D.L., Scarrott, S.M., 1981; Mon. Not. R. astr. Soc., (Short Communication) 196, 7.

Price, S.D., Walker, R.D., 1976; The AFGL Four Colour Infrared Sky Survey Air Force Geophys. Lab. Tech. Rep. AFGL TR-76-0208.

Purcell, E.M., 1969; Ap. J., 158, 433.

Racine, R., 1966; Ap. J., 73, 233.

Rhijn, P.J. van, 1929; Groningen Pub., 43, 1.

Roark, T., 1966, Ph.D. Thesis, Rennslaer Polytechnic Institute Troy, New York.

Roark, T., Roark, B., Collins, G.W., 1974; Ap. J., 190, 67.

Roberts, M.S., 1957; Pub. Astr. Soc. P. 69, 59.

Romano, G., 1975; Ap. & Sp. Sci., 33, No. 2, 487.

Rosino, L., Romano, G., 1962; Contr. Oss. Astr. Asiago, No. 127.

Rubinstein, R.Y., 1981, Simulation and the Monte Carlo Method, Wiley, New York.

Rush, W.F., 1975, Astron. J., 80, 37.

Rush, W.F., Witt, A.N., 1975, 80, 31.

Savage, B., 1973; I.A.U. Symp. No. 52, Albany, New York.

Savage, B., 1975; Ap. J., 205, 122.

Savage, B., Mathis, J.S., 1979; Ann. Rev. Astr. & Ap., 17, 73.

Salpeter, E.E., 1974a; Ap. J., 193, 579.

Salpeter, E.E., 1974b; Ap. J., 193, 585.

Schalen, C., 1945; Upps. Obs. Ann., 1, No. 9.

Schalen, C., 1948; Upps. Obs. Ann., 2, No. 5.

Schultz, G.V., Wiemar, W., 1975; Astron. & Ast., 43, 133.

Serkowski, K., 1971, Proc. I.A.U. Collq. No. 15., New Directions and Frontiers in Variable Star Research, Veroff. Remeis-Steinwarte Bamberg 9, No. 100, p.11.

Serkowski, K., 1973; Ap. J. Lett., 179, L101.

Serkowksi, K., Mathewson, D.S., Ford, V.L., 1975; Ap. J. 196, 261.

Shapiro, P.R., 1975, Ap. J., 201, 151.

Slipher, V.M., 1912; Lowell Obs. Bull., 2, 26.

Stecher, T.P., 1969; Ap. J., 157, L125.

Stone, M.E., 1970; Ap. J., 159, 277.

Strom, K.M., Strom, S.E., Carrasco, L., Vrba, F.J., 1975; Ap.J., 196, 489.

Strom, K.M., Strom, S.E., Grasdalen, G.L., 1974a; Ap. J. Lett., 187, L83.

Strom, K.M., Strom, S.E., Yost, J., 1971; Ap.J., 165, 479.

Strom, S.E., 1972; Pub. Ast. Soc. P., 84, 745.

Strom, S.E., Grasdalen, G.L., Strom, K.M., 1974c, Ap.J., 191, 111

Strom, S.E., Strom, K.M., Yost, J., Carrasco, L., Grasdalen, G.L., 1975, Ap. J., 197, 77.

Stuart, R.V., Wehner, G.K., 1962; J. Appl. Phys., 33, 2345.

Tabak, R.G., Hirth, J.P., Meyrick, G., Roark, T.P., 1975; Ap. J., 196, 457.

Tenorio-Tagle, G., 1979; Astron. & Ast., 71, 59.

Vanysek, V., 1969; Vistas in Astronomy, 11, 189.

Vanysek, V., Solc, M., 1973; I.A.U. Symp. No. 52, Albany, New York.

Vanysek, V., Svatos, J., 1964; Acta Univ. Carol. Ser. Math et Phys., 1, 1.

Viotti, R., 1969; Mem. Soc. Astron. Italiana, Nuova Ser. 40, 75.

Viotti, R., 1976; Astron. & Ast., 51, 375.

Walborn, N.R., Gull, T.R., 1982; Sky & Telescope, 63, No. 7, 16.

Walker, G.A.H., Yang, T ., Fahlman, G., 1980, Pub. Ast. Soc. P. 92, 411.

Warren-Smith, R.F., Scarrott, S.M., Murdin, P., Bingham, R.G., 1979, Mon. Not. R. astr. Soc., 187, 761.

Warren-Smith, R.F., 1979; Ph.D. Thesis, University of Durham, England.

Weston, E.B., 1951; Astron. J., 1197, 28.

Weston, G.B., 1953; Astron. J., 1206, 48.

Whitcomb, S.G., Gatley, I., Hildebrand, R.H., Keene, J., Sellgren, K., Werner, M.W., 1981; 246, 416.

White, R.L., 1979; Ap. J., 229, 954.

White, R.L., 1979b, Ap. J., 230, 116.

Whitworth, A., 1979; Mon. Not. R. astr. Soc., 186, 59.

Wickramasinghe, N.C., 1963; Mon. Not. R. astr. Soc., 126, 99.

Wickramasinghe, N.C., 1965; Mon. Not. R. astr. Soc., 131, 177.

Wickramasinghe, N.C., 1973; Light Scattering Functions for Small Particles, Pub. Hilger, London.

Wickramasinghe, N.C., Nandy, K., 1968; Nature, 219, 1347.

Wickramasinghe, N.C., Nandy, K., 1971; Mon. Not. R. astr. Soc., 153, 205.

Wickramasinghe, N.C., Nandy, K., 1972; Rep. Prof. Phys. 35, 157.

Wickramasinghe, N.C., Hoyle, F., Nandy, K., 1977; Astrophys. & Space Sci., 47, L9.

Witt, A.N., 1977a; Ap. J. Suppl., 35, 1.

Witt, A.N., 1977b; Ap. J. Supple., 35, 7.

Witt, A.N., 1977c; Ap. J. Suppl., 35, 21.

Witt, A.N., Cottrell, M.J., 1980a; Astron. J., 85, 22.

Witt, A.N., Cottrell, M.J., 1980b; Ap. J. 235, 899.

Witt, A.N., Lillie, C.F., 1972; Scientific Results from the Orbiting Astronomical Observatory OAO-2, Ed. Code. NASA SP-310.

Witt, A.N., Lille, C.F., 1978; Ap. J., 222, 909.

Witt, A.N., Oshel, E.R., 1977d; Ap. J. Suppl, 35, 31.

Witt, A.N., Stephens, T.C., 1974; Astron. J., 79, 948.

Yorke, D.G., Drake, J.F., Jenkins, E.B., Morton, D.C., Rogerson, J.B., Spitzer, L., 1973; Ap. J., 182, Ll.

Zaikowski, A., Knacke, R.F., 1975; Astrophys. Space Sci., 37, 3.

Zellner, B., 1970; Ph.D. Thesis, University of Arizona.

Zellner, B., 1973; I.A.U. Symp. No. 52, Albany, New York.

Zerull, R., 1976; Interplanetary Dust and Zodiacal Light, Ed. Elsasser & Fechtig, Springer-Verlag, Heidelberg.

Zerull, R., Giese, R., 1974; Planets, Stars and Nebulae Studied with Photopolarimetry, Ed. Gehrels, University of Arizona Press, Tucson.

ACKNOWLEDGEMENTS

The author would like to thank his academic supervisor, Dr. S.M. Scarrott, for his advice and tact during the course of this research, and also for smoothing the way on so many foreign trips. Other members of the polarimetry group, both past and present, have been influential in the development of this work, most prominently Dr. R.F. Warren-Smith whose easy approachability and knowledge of simulation techniques were invaluable.

The Physics Department and the Computer Service at Durham have always been most helpful, as have the staff of a number of other institutions : the Royal Greenwich Observatory, Herstmonceux; the Anglo-Australian Telescope, Siding Springs; the United Kingdom Infra-Red Telescope, Hawaii; the University of Hawaii Institute for Astronomy and the Wise Observatory of the University of Tel Aviv. Particular thanks must go to Colonel I. Gillam, the director of the Wise Observatory, for his ever generous hospitality.

The University of Durham and the Science and Engineering Research Council are acknowledged for their financial support both in Durham and elsewhere.

Finally, the excellent typing of Mrs. M. Chipchase is noted with regard.