Astrophysical Studies of the Dynamics of Diffuse Matter in Space

Peter W. J. L. Brand

Presented for the degree of Doctor of Science

University of Edinburgh

1991

The work in this thesis is my own, save those parts whose attribution is given in pages *iii* and *iv* of the thesis.

(signed)

Peter W. J. L. Brand

January 1991

Abstract

The thesis deals with various astrophysical aspects of diffuse matter influenced by violent events. It divides into two parts, the first part dealing with studies of relativistic material (with particles moving at speeds close to that of light), and the second part with non-relativistic material.

The first part investigates the properties of the radiation from *blazars*. These extreme forms of active galactic nucleus radiate powerfully from jets of relativistic electrons interacting with entrained magnetic fields. Our polarimetric studies have determined that the most violent events must be beamed towards us, and require energy released from the vicinity of a black hole; that the emission is most simply explained as due to electron acceleration by a magnetodynamic shock in a relatively quiescent jet of magnetized material; and that the very short timescales and high degree of polarization argue for extremely small—and therefore intense—emitting regions.

The second part of the thesis details work on the diffuse material near star-forming regions, influenced by violent winds or intense UV fluxes. The work investigates in detail the properties of shocks in star-forming regions, establishing the strongly predictive hydrodynamic shock model; pointing out the difficulty of such models in the face of line profile data obtained by ourselves; and investigating analytically the possiblity that such profiles may be the signature of magnetohydrodynamic shocks. Also described is the initial work using fluorescent emission from molecular hydrogen (and other data) to investigate the physics of photodissociation regions. These regions may be the main emission interface in external galaxies undergoing bursts of star-formation.

Table of Contents

Acknowledgements

This work would not have been produced over the past years without the loving encouragement of my wife, Eileen. To her; to my parents who set my feet on the path and who supported with great toleration my first (and later) steps; and to each of the other long-suffering members of my family, go my thanks and my love.

I have been fortunate over the years to have benefitted from the labour and friendship of many research students, whose work is in part presented here, and some of whom have continued their reseach as my colleagues and friends. I have been likewise fortunate in my collaborators. Especially important in the present context are Ian Gatley who taught me much of what I know of observational infrared astronomy, Tom Geballe who is continuing the process, and Michael Smith who took an idea and ran with it.

Finally I thank my colleagues in the Department of Astronomy for making my working environment such a pleasant place, full of friendship.

Attributions

The collegiate nature of my research makes detailed attribution difficult.

Furthermore, most of the projects have involved a student, and it is of the essence of such a relationship that initially the supervisor creates the project and shows the student how to progress, later advises and collaborates, and finally the student leads that part of it that is his (or her) own.

Here I give an assessment of my contribution to the work that I present in this thesis.

Part I

The blazar observational programme started in 1979 with research student Chris Impey (now on the staff of the Steward Observatory, University of Arizona) who was a major contributor to its inception and who played a larger part than any other student in its development. It continued with Paul Holmes, Andrew Mead and Keith Ballard, who in their turn have prepared observing projects, reduced data, done statisical analyses, had good ideas, and been reponsible for the drafting of scientific papers.

My contribution has been in instructing them in how to do these things, encouraging and criticising their execution, suggesting ways forward when progress stalled, inserting new strategies, and in giving overall direction to the programme (and occasionally having the luxury of performing the basic tasks myself). I have also acted as coordinator of the project with respect to our collaborators from other institutions, most notably Prof. J. Hough of Hatfield Polytechnic who built the polarimeter used for most of this research.

Part II

The programme to determine the excitation and dynamics of molecular hydrogen arose out of an early collaboration with Adrian Webster and Tom Geballe. Over the years it also involved as students Mark Toner, Peter Meadows, Mike Burton and Alan Moorhouse, the latter two of whom continue to collaborate in this work. Mike Burton now plays a large role in the evolution of the research. Again I claim a major part of the direction and execution of this programme.

The investigation—as an offshoot from the hydrogen programme—of magnetohydrodynamically moderated C-shocks is a development of early work by myself, and evolved along lines that I determined. On the other hand, it owes a great deal also to the dedicated and inspired work of Mike Smith.

In the extragalactic research programme, I claim approximately 'half share' with Phil Puxley, previously my research student, for the direction of the programme. Phil has been extremely innovative and determined in this work, which forms the basis of his current research, and in which I still collaborate.

In the remaining areas I claim the major part of the work, particularly in the dynamical researches.

In several subsidiary areas arising from the mainstream of my reseach there are offshoots that I became involved in, but did not lead.

Introduction

Overview

This thesis falls naturally into two parts. Part I contains work on relativistic material (*i.e.* material in which velocities are close to that of light), while Part II deals with relatively low energy material. In each part the material is *diffuse*, not in a gravitationally tightly bound configuration such as a star or a black hole. In each part the material is energised by its proximity to such a configuration. In each part the processes of fluid dynamic collimation, of shock excitation, and of magnetohydrodynamics, are important.

The astrophysical problem which we attempt to solve is: how do we use Earth-based observation and universal physical laws to determine these distant physical processes?

Part I
Relativistic Matter : blazars

The first part of this thesis, containing fifteen papers, concerns itself with the astrophysics of 'blazars'. The most powerful sources of radiation in the Universe are quasars, and of these the variety with the most extreme behaviour—the 'blazar'—radiates most of its energy to us by the synchrotron mechanism: relativistic electrons spiral in a magnetic field, and lose energy to radiation at a relativistically enhanced rate.

The dynamical origin of this process is still mysterious. It seems likely that the blazar emitter is a shock-excited region in a collimated relativistic jet powered by a massive black hole at the centre of a galaxy. Jets appear to be the power conduits for all active galaxies which comprise several percent of the total population of galaxies (recent work

has uncovered a jet at the centre of our own galaxy). In the case of some radio galaxies, these jets have transported so much energy into giant lobes of relativistic material around the galaxy that several million solar masses of material must have been annihilated. Only gravitational energy release, via the surroundings of a massive (10^7 – 10^9 solar masses) black hole, can explain this amount of energy. The jet connects the central black hole to the lobe and, like a blowtorch, squirts the power into a particular direction in space.

In observational terms, a blazar is a point-like optical source with a smooth spectrum (but in many cases with very weak absorbtion or emission lines) which varies irregularly by a large factor on a timescale between days and years, a large (>3%) degree of polarization and, nearly invariably, powerful radio emission.

Comprehension of the nature of these objects is one of the goals of astrophysics.

I.1 Surveys

Papers 1-5

This section details an extended survey of the photopolarimetric properties of blazars. Initially (papers 1-3) only the infrared wavebands at 1.25μm (J), 1.65μm (H) and 2.2μm (K), were used, but these were later extended to eight optical and infrared bands between 0.36μm and 2.2μm. Polarimetry and photometry of a sample of identified blazars was obtained. The aim was to have as complete time coverage as the telescope allocation panels would allow, to accommodate time variability.

The reason for the polarimetry is that synchrotron radiation is highly polarized perpendicular to the magnetic field of the source as projected on the sky. Such measurements constrain the magnetic field geometry in the source.

Two major strategies have been followed. On the one hand, detailed studies of single objects are occasionally possible when dramatic changes are happening during an observing period. Such observations are described in the next section. On the other hand, regular monitoring of a sample of objects provides a statistical sample of blazar behaviour. The papers (1-5) of this section describe such a programme.

The resulting database is the most complete sample of blazar behaviour in the optical/infrared region. The conclusions so far are that no clear pattern of behaviour has

yet emerged in the variability so far observed, but that:

- Some blazars' energy budgets can only be explained by the relativistic beaming of radiation toward us, and/or the non-steady non-isotropic release of energy from the vicinity of a black hole.
- Variations of flux and polarization with wavelength and with time are consistent with the non-steady emission of radiation by relativistic electrons accelerated in a magnetohydrodynamic shock, superposed on a steady background emission from an equilibrium distribution of relativistic electrons—possibly the quiescent jet.
- Polarizations can reach such huge levels that the magnetic field within the emitting region must be highly organised, and that this degree of organisation argues for a small emitting region.
- Relatively short variability timescales, typically of days, also argue for small emitting regions, although they may be far from the (assumed) black hole source of energy.
- The power in such events is so great that either extreme relativistic beaming occurs, or the event must involve the total energy output; this constrains beam coollimation, and indicates an intrinsic source size of light days.
- To progress further, monitoring photopolarimetry on an even wider wavelength baseline is required; fortunately this will shortly be possible as a polarimeter becomes available on the James Clerk Maxwell submillimetre telescope.

I.2 Individual Objects

Papers 6-8

A few blazars have been captured during the monitoring programme in the middle of violent change. This section describes the observations and conclusions that can be drawn, for the blazars AO 0235+164, OJ287 and 3C345.

I.3 Multifrequency Campaigns

Papers 8-12

Part of the spinoff from the monitoring programme has been the involvement in occasional campaigns—which require a huge degree of organization, particularly in the acquisition of simultaneous telescope time, and are therefore rare—to observe a blazar at many different wavelengths simultaneously, using many different telescopes to do so.

This simultaneity eliminates apparent spectral peculiarities created by variability between epochs of observation at different wavelengths.

Such campaigns have covered all wavelengths from X-ray to radio, and demonstrate the smooth featureless spectrum of synchrotron radiation, with a spectral slope varying smoothly from a small value at radio wavelengths to a moderately steep (or occasionally very steep) slope at short wavelength.

Such a shape is characteristic of relativistic electron generation and loss typical of shock acceleration, and suggests that this is the universal mechanism.

I.4 Related Topics

Papers 13-15

These papers cover work related to the main theme: an investigation of the nature of the steepest spectrum sources; a test of the assertion that all blazars are radio loud (as predicted by the general model)— no radio-quiet blazars were found, nor have there been any since; and an improved algorithm for searching for absorbtion lines in quasars.

Part II
Non-relativistic matter : star-forming regions

The latter group of thirty-five papers describes work on the diffuse material close to the star-forming regions in our own or neighbouring galaxies. The phenomena under investigation here are caused by the the violent energy outbursts that accompany star formation, or by the irradiation of their surroundings by intense UV radiation from recently formed massive stars.

II.1 Molecular hydrogen

Papers 16-34

The hydrogen molecule exists in interstellar space only in dense dust-shielded clouds impenetrable by visible radiation. Infrared radiation can penetrate these clouds, and H_2 is most easily detected by IR radiation from transitions between the vibrational and rotational levels of the electronic ground state. The lowest such level has an excitation temperature of 500K, and therefore such transitions are generated in hot molecular gas, found in the vicinity of very young or forming stars.

The intensities in these lines have a characteristic distribution if the radiating material is heated by the passage of a shock, and is cooling as a consequence. Papers in the first section (Shock excitation, 16-26) describe an observational campaign to investigate the shocked gas in young stellar outflows. We have demonstrated that the data are consistent with shock excited gas if heating is sudden, but are not consistent with currently fashionable plane parallel magnetohydrodynamically moderated shock models. if we are correct, these shocks are a powerful diagnostic of shock pressure, of dominant cooling processes, and of the extinction by dust in those regions; these are all important parameters in understanding the phenomenon of outflows in star forming regions.

Our work on line profiles has confirmed and extended the result that the line widths cannot easily be explained by *any* of the published models, including our own.

Papers 27-30 concern themselves with C shocks. This work, with M. D. Smith, was designed to explore the one remaining loophole in the process. A complex bow shock type structure was assumed for the shock interaction, and the intensities and profiles due to the entire flow were calculated. We had already demonstrated that non-magnetic shocks even in such configurations could not explain the linewidths. This part of the programme investigated—using the same technique developed for the non-magnetic shocks, namely analytic approximation—the properties of magnetically moderated bow shocks. Currently, the status is that such configurations *could* explain the observed intensities and linewidths, but only for a special shape of bow shock and only in the case of abnormally large magnetic fields. Work continues.

Papers 31-34 deal with fluorescent emission from molecular hydrogen. This programme

started with the discovery of IR fluorescence of H_2 in the reflection nebula NGC2023 by a team (including the author) led by Ian Gatley. The programme has continued on an observational basis by examining the fluorescence in a number of other objects, and by mapping and spectroscopy of ionized and molecular gas and of dust emission. This has enabled a systematic investigation of the physics of the so-called photodissociation front, the boundary between dense molecular material, and hot gas and intense UV radiation from a very young star. Theoretical work by collaborators, and now continuing with students, is matching these results to models, again exploiting the technique of analytical approximation.

II.2 External Galaxies

Papers 35-39

These papers describe attempts to understand the causes and effects of star formation on the scale of entire galaxies. There are strong indications that the physics of the photodissociation region, discussed in the last section, can be applied to the bursts of star formation observed in external galaxies. Both observational and theoretical studies of these processes are under way. At present careful analysis of the extinction in these regions is essential, since this has a profound effect on the mass distribution of newly formed stars, one of the primary parameters of any star formation theory.

The re-radiated starlight appears in the submillimetre region of the spectrum, and maps of galaxies at these wavelengths are only now beginning to appear. In conjunction with diagnostic maps of the molecular material (primarily CO emission, but many other species are now measurable), these surveys are essential in the pursuit of the correct theory of starbirth.

II.3 3.28μm dust emission

Papers 40-43

One of the discoveries of infrared spectroscopy is the pervasiveness of the dust emission features, particularly one at 3.28μm, which has been attributed to a C-H bending mode in a carbon ring structure (a polyaromatic hydrocarbon). The association of these features with fluorescent molecular emission appears to be a powerful diagnostic of photodissociation region physics, and provides a fascinating research area. These papers

provide some information on the physics of this emission process.

II.4-6 Cometary Globules, Star formation in galaxies, Physical processes

Papers 44-54

These groups of papers detail: the discovery and investigation of Cometary Globules, small star forming regions triggered by previous star formation; attempts to synthesize a larger scale scenario for star formation, an ongoing obsession; and investigations of particular processes important in the diffuse matter that forms the essence of this thesis.

TABLE OF PAPERS IN THE THESIS

Part I

Relativistic matter : blazars

1 Surveys

1. Infrared Photometry and Polarimetry of BL Lac Objects
with C.D. Impey, P.M. Williams & R.D. Wolstencroft, Mon. Not. R. astr. Soc. 200, 19-40 (1982).

2. Infrared Polarimetry and Photometry of BL Lac Objects - II
with C.D. Impey, R.D. Wolstencroft and P.M. Williams, Mon. Not. R. astr. Soc. 209, 245-269 (1984).

3. Infrared Polarimetry and Photometry of BL Lac Objects - III
with P.A. Holmes, C.D. Impey and P.M. Williams, Mon. Not. R. astr. Soc. 210, 961 (1984).

4. The optical and infrared emission of blazars
with K. Ballard, A.R.G. Mead, J. Hough. Mon. Not. R. astr. Soc. 243, 640 (1990).

5. Optical and infrared polarimetry and photometry of blazars (IV)
with A.Mead, K. Ballard, J. Hough, C. Brindle, J. Bailey. Astron. and Astrophys. Suppl. 83, 183 (1990).

2 Individual Objects

6. A Polarisation Burst in the BL Lac Object AO 0235+164
with C.D. Impey and S. Tapia, Mon. Not. R. astr. Soc. 198, 1-9 (1982).

7. A Polarization Flare in OJ 287
with P.A. Holmes, C.D. Impey, P.M. Williams, P. Smith, R. Elston, T. Balonek, M. Zeilik, I. Burns, P. Heckert, R. Barvainis, J. Kenny, G. Schmidt and J. Puschell, Mon. Not. R. astr. Soc. 211, 497-506 (1984).

8. Polarimetric observations of the quasar 3C345
with A.R.G. Mead, J. Hough, J. Bailey. Mon. Not. R. astr. Soc. 233, 503-510 (1988).

3 Multifrequency Campaigns

9. Simultaneous Observations of the BL Lac Object I Zw-187
with J.N. Bregman et al., Astrophys. J. 253, 19-27 (1982).

10. Variability of BL Lac Objects at X-ray and Other Frequencies
with A.M.T. Pollock, J.L. Bregman and E.I. Robson, Space Sci. Revs. 40, 607-612 (1985).

11. Simultaneous Multifrequency Observations of the BL Lac Object Markarian 421.
with F. Makino et al., Astrophys. J. 313, 662-673 (1987).

12. Multifrequency observations of BL Lacertae
with J. Bregman et al. Astrophys. J. 352, 574 (1990).

4 Related Topics

13. Infrared Photometry of Flat Spectrum Radio Sources
with C.D. Impey, Nature 299, 814-816 (1981).

14. The Calibration of a radio-independent search for BL Lac Objects
with C.D. Impey, Mon. Not. R. astr. Soc. 201, 849-866 (1982).

15. Detection Lines by a Calibrated Algorithm in a Close Pair of Quasars
with A.S. Trew, Mon. Not. R. astr. Soc. 211, 485-495 (1984).

Part II

Non-relativistic matter : star nurseries

1 Molecular hydrogen

1.1 shock excitation

16. Shocked Molecular Hydrogen in the Supernova Remnant IC443
with M.G. Burton, T.R. Geballe and A.S. Webster. Mon. Not. R. astr. Soc. 231, 617-634 (1988).

17. Ratios of Molecular Hydrogen Line Intensities in Shocked Gas: evidence for cooling zones.
with A. Moorhouse, M.G. Burton, T.R. Geballe, M. Bird and R. Wade. Astrophys. J. 334, L103-106, (1988)

18. The Constancy of the Ratio of the Molecular Hydrogen Lines at 3.8μm in Orion.
with M.P. Toner, A.S. Webster, T.R. Geballe and P.M. Williams. Mon. Not. R. ast Soc. 236, 929-934 (1989).

19. Molecular Hydrogen Line Ratios in Four Regions of Shock-excited Gas.
with M.G. Burton, T.R. Geballe and A.S. Webster. Mon. Not. R. astr.Soc. 236, 409-423, (1989)

20. The velocity profile of the 1-0 S(1) line of molecular hydrogen at Peak 1 in Orion.
with M.P. Toner, T.R. Geballe and A.S. Webster. Mon. Not. R. astr. Soc, 237, 1009-1018, (1989).

21. Shocked Molecular Hydrogen in the Bipolar Outflow NGC2071
with M.G. Burton, T.R. Geballe. Mon. Not. R. astr. Soc. 238, 1513-1527, (1989).

22. High spectral and spatial resolution imaging of the shock waves in Orion
with M.G. Burton, J. Bland, D. Axon, R. Garden, T.R. Geballe, D. Hollenbach, I. McLean and A. Moorhouse. ASP Centennial meeting 'Evolution of the interstellar Medium' ed. L. Blitz (1989).

23. The nature of shocks in molecular clouds
IAU Colloquium 120, Springer, (1989).

24. The velocity profiles of high excitation molecular hydrogen lines
with T.G. Geballe, A. Moorhouse, M.Burton. Mon. Not. R. astr. Soc. 242, 88 (1990).

25. Infrared Images of ionized and molecular hydrogen emission in S106
with S.S.Hayashi, T. Hasegawa, M. Tanaka, M. Hayashi, C. Aspin, I.S.McLean, and I.Gatley, Astrophys. J. 354, 242 (1990).

26. Surprisingly high pressure molecular shocks in the SNR IC 443
with A.Moorhouse, T.R.Geballe and M.G.Burton. Accepted by Mon. Not. R. Astr. Soc. (1990).

1.2 C-shocks

27. Cool C-shocks and high velocity flows in molecular clouds
with M.D. Smith. Mon. Not. R. astr. Soc. 242, 495 (1990).

28. Signatures of C-shocks in molecular clouds (II)
with M.D. Smith. Mon. Not. R. Astr. Soc. 243, 498 (1990).

29. H_2 profiles of C-type bow shocks. (III)
with M. D. Smith. Mon. not. R. astr. Soc. 245, 108 (1990).

30. Bow shocks in molecular clouds : H_2 line strengths (IV).
with M.D. Smith, A. Moorhouse. accepted by Mon not. R. astr. Soc. (1990).

1.3 Fluorescence

31. Level Population and Para/Ortho Ratio of Fluorescent H_2 in NGC 2023
with T. Hasegawa, I. Gatley, R.P. Garden, M. Ohishi, J.F. Hayashi and N. Kaifu.Astrophys. J. 318, L77-80 (1987).

32. Fluorescent Molecular Hydrogen Emission from the Reflection Nebula NGC 2023
with I. Gatley, T. Hasegawa, H. Suzuki, R.P. Garden, J.F. Lightfoot, W. Glencross, H. Okuda and T. Nagata. Astrophys. J. 318, L73-77 (1987).

33. Infrared Spectroscopy of Interstellar Molecular Hydrogen: Decomposition of Thermal and Fluorescent Components.
with M. Tanaka, T. Hasegawa, S. Hayashi, I. Gatley. Astrophys. J. 336, 207-211, (1989).

34. High spectral resolution observations of fluorescent molecular hydrogen
with M.G.Burton, T.R.Geballe, and A. Moorhouse. Astrophys. J. 352, 625 (1990).

2 Star formation in galaxies

35. Emission Line Ratios from HII Regions in Two Nearby Galaxies
with I.M. Coulson and W.J. Zealey, Mon. Not. R. astr. Soc. 195, 353-360 (1981).

36. The Origin of the Far IR Flux from Spiral Galaxies
with S.K. Leggett and C.M. Mountain. Mon. Not. R. astr. Soc. 228, 11-16 (1988)

37. Detection of H53α Emission from M82 : a reliable measure of the ionization rate and its implications
with P.J. Puxley, T.J. Moore, C.M. Mountain, N. Nakai. Astrophys. J. 345, 163 (1990).

38. A 450μm continuum map of M82 : Comparison with CO emission.
with P.Smith, P.J. Puxley, C.M.Mountain. Mon. Not. R astr. Soc. 243, 97 (1990).

39. H53α emission from bright galaxies
with P.J. Puxley, T.J.T. Moore, C.M. Mountain and N.Nakai. accepted by Mon. not. R. astr. Soc. (1990).

3 3.28μm dust emission feature

40. Occurence of the 3.3μm Feature in Galaxies
with T.J. Lee, D.H. Beattie, I. Gatley. H. Hyland and T. Jones. Nature 295, 214 (1982).

41. The Behaviour of the 3.28μm dust feature in NGC 2024
with D.A. Beattie, P.J. Gatley, R.D. Meadows and R.D. Wolstencroft, Occ. Report Royal Observatory 12, 105-112 (1984).

42. Spatial Variations of the 3μm Emission Features with UV-excited Nebulae: photochemical evolution of interstellar PAHs.
with T.R. Geballe, L.F. Allamandola, A.G.G.M. Tielens and A. Moorhouse. Astrophys. J. 341, 278-287, (1989).

43. Infrared images of reflection nebulae and Orion's Bar: fluorescent molecular hydrogen and the 3.3μm feature.
with M.G. Burton, A. Moorhouse, P.F.Roche, and T.R.Geballe. IAU Symposium 135, 'Interstellar Dust', D. Reidel, (1989).

4 Cometary globules

44. Cometary Globules and the Structure of the Gum Nebula
with T. Hawarden, Mon. Not. R. astr. Soc. 175, 19-22p (1976).

45. The Young Cluster NGC 5367 and A 1353 -40
with P.M. Williams, A.J. Longmore and T.G. Hawarden, Mon. Not. R. astr. Soc. 181 709-715 (1977).

46. Dust Clouds in HII Regions. The Dragon in M8
with W.J. Zealey, Astron. & Astrophys. 63, 345-352 (1978).

47. Cometary Globule 1
with T.G. Hawarden, P.M. Williams, A.J. Longmore, J.C. Caldwell and I.M. Coulson Mon. Not. R. astr. Soc. 203, 215-222 (1982).

5 Global theories of diffuse matter

48. Cloud Structure in the Galactic Plane: A Cosmic Bubble Bath?
with W.J. Zealey, Astron. & Astrophys. 38, 363-371 (1975).

49. Spiral Arms and Supernova-dominated Interstellar Medium
with S.R. Heathcote, Mon. Not. R. astr. Soc. 198, 545-562 (1982).

50. The State of Clouds in a Violent Interstellar Medium
with S.R. Heathcote, Mon. Not. R. astr. Soc. 203, 67-86 (1982).

6 Physical processes in star-forming regions

51. Forced Diffusion: a mechanism to create dust clouds
Astrophys. and Space Science 33, 231-234 (1975).

52. The Properties of Weak Shocks in HII Regions
with J.S. Mathis, Astrophys. J. 223, 161-167 (1978).

53. Scattering from a Cylindrical Dust Cloud in an Isotropic Radiation Field
Astron. & Astrophys. 71, 47-50 (1978).

54. Ionisation Front Interactions and the Formation of Globules
Mon. Not. R. astr. Soc. 197, 217-233 (1981).

THE PUBLISHED PAPERS

Mon. Not. R. astr. Soc. (1982) **200**, 19–40

Infrared polarimetry and photometry of BL Lac objects

C. D. Impey*, P. W. J. L. Brand *University of Edinburgh, Edinburgh EH9 2HX*

R. D. Wolstencroft and P. M. Williams *Royal Observatory, Edinburgh EH9 3HJ*

Received 1981 September 14; in original form 1981 March 4

Summary. Infrared polarimetry and photometry have been obtained for a sample of 18 BL Lac objects. The data covers a period of one year and is part of a continuing monitoring programme; all observations were in the J, H and K wavebands. Large and variable degrees of polarization are a common property of the sample. Two BL Lac objects show wavelength-dependent polarization, with the polarization increasing towards shorter wavelengths, and two objects show evidence for position angle rotations over a five-day period. The relationship between changes in polarized and total flux is also discussed. The BL Lac objects cover an enormous range of infrared luminosity; the three most luminous having $L_{IR} > 10^{46}\,\mathrm{erg\,s^{-1}}$ and the other end of the range having infrared luminosities similar to normal elliptical galaxies. These are the first published infrared polarimetric observations for eight of the sample.

1 Introduction

Current models try to explain all active nuclei in terms of the same compact power source in different environments. For instance, the absence of strong emission lines in BL Lac objects can be explained by a lack of gas rather than by insufficient soft X-ray or UV ionizing flux. OVV quasars seem to be transitional between BL Lac objects and quasars; although they have strong emission lines they share the rapid variability and high polarization of the BL Lac compact source. Spectral features are only indirect indicators of the nature of the compact source, once the redshift has been determined. BL Lac objects are important because they allow the opportunity to probe the 'engine' of active nuclei directly.

This paper reports the results of the first year of a monitoring programme of polarized BL Lac-type objects. The aim of the programme, is to establish a database of infrared polarimetry for as many objects as possible obtaining measurements in more than one colour. Recent observations of polarized extragalactic objects have been comprehensively

* Present address: Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA.

reviewed by Angel & Stockman (1980). Most previous polarimetry has been in the optical region (e.g. Angel *et al.* 1978), with only a few objects studied in the near infrared (Puschell & Stein 1980). Infrared polarimetry extends the observational wavelength baseline into a region where BL Lac objects emit most of their energy.

2 Observations

The observations were all carried out at the 3.8-m United Kingdom Infrared telescope (UKIRT) on Mauna Kea using a photovoltaic InSb detector and rotating HR Polaroid. All the data were collected in five separate runs during the period 1979 December to 1980 November. Due to the developing state of the telescope during that period, two runs employed focal-plane chopping at the $f/9$ focus while the three later runs used the $f/35$ chopping secondary. All polarimetric observations used star/sky chopping at a rate of 5–7 Hz. Linear polarization data were collected at 45° increments of the analyser. The instrumental parameters were redetermined for each run and the instrumental sensitivity and polarization were calculated an average of twice a night. The broad band filters and their effective wavelengths were: J (1.25 μm), H (1.65 μm) and K (2.2 μm).

The polarimetry was calibrated with a series of instrumental tests. Nearby unpolarized stars (Serkowski 1974) were observed to determine the instrumental polarization and the results have been corrected accordingly. The (negligible) polarization due to the telescope optics was measured by rotating the photometer cage through 90°. The data have also been corrected for the polarizing efficiency of the HR Polaroid, which varies across the bandwidth of the filters. The correction is only greater than 1 per cent at longer wavelengths, and it can be shown that residual birefringence at the long wavelength side of the K filter does not influence the results. The analyser efficiency was measured by inserting an identical piece of HR Polaroid into the optical path and rotating the analyser through 360°. By performing this test for a complete rotation of the second analyser, the response of the polarimeter to all incident linear polarization vectors was mapped. In addition, the protostellar BN object was measured as a polarized standard on three occasions. In each case the measured polarization at K was within 1 per cent of the published value (Dyck & Beichman 1974; Johnson *et al.* 1981). To check for magnitude-dependent effects, we observed an unpolarized star of similar brightness to most of the BL Lac objects. The resulting instrumental polarization agreed with the errors with the values derived from bright standard stars. Linear polarization measurements were made assuming that there was no elliptical component in the incoming radiation (i.e. Stokes parameter $V = 0$). The basis for this assumption is the negative searches for optical circular polarization ($\lesssim 0.1$ per cent) in extragalactic objects with high linear polarization (Landstreet & Angel 1972; Nordsieck 1972; Kemp, Wolstencroft & Swedlund 1972; Maza 1979). Finally, the principal transmission axis of the rotating analyser was related to the position angle on the sky with an offset accurate to 1°.

Normalized Stokes parameters and their errors were calculated from a least-squares analysis of the intensities sampled at 45° intervals. The fractional degree of polarization ($p = (Q^2 + U^2)^{1/2}/I$) and position angle ($\theta = \frac{1}{2}\tan^{-1}(Q/U)$) were calculated from these quantities. The errors in p and θ were calculated in two ways: (a) in terms of the χ^2 goodness-of-fit of each observation to the reduction algorithm and (b) in terms of a characteristic error incorporating systematic effects which might modulate the signal at frequencies other than the analyser rotation frequency. These systematic effects include blemishes on the analyser, movement of the image on the detector, scattered light in the photometer and imperfect beam profiles. The larger of the two errors was chosen for each observation. At low signal-to-noise, the probability distribution of the true values of p is not Gaussian and

values for θ may be undefined (Vinokur 1965; Wardle & Kronberg 1974). In particular, when the signal-to-noise ratio is $\lesssim 2$, there is a bias introduced to the polarization. No observation with a signal-to-noise ratio of < 5 has been included in this paper.

The photometry was calibrated in absolute units by observing standard stars. The spectral flux calibration is based on the calibration of α Lyr by Oke & Schild (1970) and the model atmosphere calculation by Schild, Petersen & Oke (1971), and the flux for a 0 mag star is taken to be 1520 mJy at J(1.25 μm), 980 mJy at H(1.65 μm) and 620 mJy at K(2.2 μm). Broad-band filter measurements do not uniquely define an energy distribution, so the data were corrected for the difference between the spectral flux distribution of the BL Lac objects and the calibrators. The incident flux distribution was convolved with the filter and Polaroid transmission functions, using power laws of various spectral indices to represent the objects and the airborne spectrophotometry of Strecker, Erickson & Witteborn (1979) to represent the standards. These corrections have been applied to all the flux measurements discussed here. The inhomogeneous nature of the interstellar medium makes the correction due to interstellar extinction difficult to determine, but the problem is eased in this study because the objects are generally at high galactic latitude and the extinction is considerably reduced in the infrared (Heiles 1976; Burstein & Heiles 1978). Using the method of Sandage (1972) and the extension of the extinction law by Johnson (1968), corrections were needed for only two objects, 0735+178 and 0851+202. The error in each flux measurement was chosen from the greater of (a) the error due to the observed signal-to-noise, (b) the error associated with the repeatability of successive standard observations separated by only a minute, and (c) the constancy of the instrumental sensitivity on successive nights. Sensitivity changes of > 10 per cent were observed between observing runs. Since three standard stars were common to two or more runs, it was possible to confirm that there had been real changes in sensitivity rather than an incorrect determination of the zero point.

All the objects in this study were more than 3 mag above the detection limit of the InSb system at the time of measurement. The BL Lac objects were often too faint for the first generation TV acquisition system on the UKIRT, and were centred in the infrared beam using bright offset stars with accurate positions. No uncertainty was introduced by this procedure because the objects were bright enough to peak up on the infrared signal and define half-power boundaries for the two beam positions. The beam positions were checked periodically throughout the night, especially when observing through the meridian. Most of the observations were made through apertures of 10 and 15 arcsec, with a few at 19 and 12. Only two of the BL Lac objects have images sufficiently non-stellar to require an aperture correction. There is no correction between aperture for stellar images because sky compensation is carried out continuously. All observations were made with little or no Moon in the sky.

3 Results

The individual objects are discussed below, and the results are presented in Table 1. In addition to those listed, the following sources were not detected at 2.2 μm, implying a magnitude level of $K \gtrsim 15$: 0048−097 (10.7.80), 0109+224 (20.12.79), 0219+428 (20.12.79), 0818−128 (21.12.79). The dates of the null observations are in brackets. Column (1) gives the Parkes designation and any other source name, column (2) gives the date of the observation and columns (3) and (4) contain the aperture size and effective wavelength of the broad band measurement. The degree of polarization, position angle and flux in millijansky (1 mJy = 10^{-29} W m^{-2} Hz^{-1}) are contained in columns (5), (6) and (7) along with their associated rms errors.

Table 1. Details of polarimetry and photometry for all the objects.

(1) Object	(2) Date	(3) Aperture	(4) Waveband	(5) $p \pm \sigma(p)$	(6) $\theta \pm \sigma(\theta)$	(7) $S \pm \sigma(s)$
0235 + 164	18.12.79	10″	*K*	17.5 ± 2.9	137 ± 4.8	14.6 ± 0.3
(AO)			*H*	22.8 ± 1.8	129 ± 2.3	6.7 ± 0.2
	19.12.79	10″	*K*	18.7 ± 3.1	120 ± 4.8	12.2 ± 0.3
			H	22.3 ± 1.6	114 ± 2.1	8.6 ± 0.3
	20.12.79	10″	*K*	28.7 ± 3.1	125 ± 3.1	13.4 ± 0.4
			H	34.6 ± 2.9	128 ± 2.4	7.6 ± 0.3
			J	36.2 ± 8.6	137 ± 6.8	4.0 ± 0.2
	21.12.79	10″	*K*	26.7 ± 3.6	110 ± 3.8	13.3 ± 0.2
	22.12.79	10″	*K*	14.9 ± 1.9	108 ± 3.0	14.3 ± 0.2
			H	19.4 ± 1.2	109 ± 1.8	10.2 ± 0.4
			J	22.1 ± 2.0	119 ± 2.6	5.6 ± 0.3
	11.07.80	15″	*K*	17.1 ± 5.0	15 ± 8.4	6.5 ± 0.1
			H	10.0 ± 2.1	16 ± 6.1	4.0 ± 0.3
	11.08.80	19″	*K*	15.1 ± 3.4	156 ± 6.4	12.6 ± 0.4
	08.11.80	12″	*K*	6.2 ± 1.9	6 ± 9	17.0 ± 0.2
			H	2.0 ± 4.6	~1	8.9 ± 0.4
			J	9.6 ± 3.9	35 ± 11	5.7 ± 0.2
	09.11.80	12″	*K*	3.9 ± 1.9	157 ± 14	17.0 ± 0.2
0306 + 103	22.12.79	10″	*K*	2.2 ± 1.9	~102	10.0 ± 0.1
(OE 110)	09.11.80	12″	*K*	3.6 ± 4.4	~171	11.6 ± 0.3
0735 + 178	18.12.79	10″	*K*	19.7 ± 1.6	2.6 ± 2.3	16.4 ± 0.2
(PKS)			*H*	23.2 ± 1.3	5.7 ± 1.6	12.0 ± 0.4
	18.12.79	10″	*K*	19.8 ± 2.5	8.6 ± 3.6	14.9 ± 0.4
			H	24.4 ± 2.0	8.5 ± 2.4	11.0 ± 0.4
			J	27.5 ± 5.0	11.3 ± 5.2	7.7 ± 0.4
	20.12.79	10″	*K*	21.5 ± 2.2	9.7 ± 2.9	17.1 ± 0.2
			H	21.4 ± 2.4	6.4 ± 3.2	12.5 ± 0.4
	21.12.79	10″	*K*	21.0 ± 1.3	0.7 ± 1.8	13.6 ± 0.2
			H	22.8 ± 1.7	178 ± 2.1	9.0 ± 0.3
	22.12.79	10″	*K*	19.7 ± 1.8	11.2 ± 2.6	16.7 ± 0.2
			H	23.6 ± 1.5	8.2 ± 1.8	12.4 ± 0.4
			J	22.7 ± 2.0	9.5 ± 2.5	9.2 ± 0.5
	20.04.80	15″	*K*	27.9 ± 2.9	160 ± 3.0	15.7 ± 0.1
			H	29.9 ± 2.5	153 ± 2.4	10.0 ± 0.2
	21.04.80	15″	*K*	22.5 ± 2.9	149 ± 3.7	19.2 ± 0.4
			H	24.5 ± 3.7	156 ± 4.3	14.2 ± 0.4
			J	27.4 ± 4.2	158 ± 4.4	11.0 ± 0.3
	08.11.80	12″	*K*	11.4 ± 0.5	148 ± 1.2	29.3 ± 0.3
			H	12.6 ± 1.1	147 ± 3.0	20.2 ± 0.2
			J	12.4 ± 1.4	145 ± 3.1	16.0 ± 0.2
	09.11.80	12″	*K*	12.6 ± 1.8	154 ± 3.9	33.0 ± 0.3
			J	11.6 ± 2.3	168 ± 5.4	18.2 ± 0.4
0754 + 101	18.12.79	10″	*K*	2.3 ± 0.7	49 ± 8.6	34.8 ± 0.6
(OI 090.4)	19.12.79	10″	*K*	5.1 ± 0.5	31 ± 2.8	35.1 ± 0.9
			H	5.2 ± 1.3	39 ± 7.2	24.1 ± 0.9
	20.12.79	10″	*K*	8.0 ± 0.7	36 ± 2.5	32.0 ± 0.4
	21.12.79	10″	*K*	8.8 ± 2.6	41 ± 8.4	25.5 ± 0.5
			H	11.7 ± 1.2	39 ± 3.0	20.6 ± 0.7
0851 + 202	18.12.79	10″	*K*	5.6 ± 2.9	89 ± 15	25.9 ± 0.3
(OJ 287)			*H*	10.5 ± 1.6	93 ± 4.4	18.5 ± 0.7
	19.12.79	10″	*K*	8.4 ± 0.8	98 ± 2.6	27.2 ± 0.7
			H	10.3 ± 1.4	98 ± 3.9	19.1 ± 0.7
			J	10.1 ± 1.1	91 ± 3.1	15.5 ± 0.9
	20.12.79	10″	*K*	8.9 ± 2.1	98 ± 6.7	29.2 ± 0.4
			H	8.3 ± 1.7	99 ± 5.9	19.0 ± 0.7
	21.12.79	10″	*K*	9.1 ± 2.5	97 ± 7.8	23.0 ± 0.3
			H	8.8 ± 1.0	85 ± 3.3	17.8 ± 0.6

Table 1. *(Continued)*

(1) Object	(2) Date	(3) Aperture	(4) Waveband	(5) $p \pm \sigma(p)$	(6) $\theta \pm \sigma(\theta)$	(7) $S \pm \sigma(s)$
	22.12.79	10″	*K*	10.3 ± 2.6	62 ± 7.3	28.1 ± 0.4
			H	8.6 ± 0.8	64 ± 2.6	20.4 ± 0.7
			J	9.7 ± 1.9	60 ± 5.6	13.8 ± 0.8
	20.04.80	15″	*K*	9.2 ± 1.4	82 ± 4.4	16.4 ± 02
			H	10.4 ± 1.3	78 ± 3.6	10.9 ± 0.2
			J	11.1 ± 6.7	77 ± 17	9.6 ± 0.5
	21.04.80	10″	*K*	10.5 ± 1.2	79 ± 3.3	25.0 ± 0.5
			H	14.2 ± 5.2	82 ± 10	19.9 ± 0.6
			J	14.3 ± 4.8	71 ± 9.7	14.0 ± 0.7
1147 + 245 (OM 280)	18.04.80	15″	*K*	10.1 ± 3.3	115 ± 9.4	8.0 ± 0.1
	20.04.80		*K*	14.8 ± 6.3	109 ± 12	8.8 ± 0.2
1308 + 326 (B2)	19.12.79	10″	*K*	19.6 ± 1.1	150 ± 1.6	11.8 ± 0.3
	21.12.79		*K*	7.6 ± 3.7	145 ± 14	7.7 ± 0.1
	22.12.79		*K*	13.7 ± 4.9	132 ± 10	8.2 ± 0.1
	10.07.80	15″	*K*	8.3 ± 0.5	50 ± 1.7	6.8 ± 0.1
			H	7.2 ± 3.2	76 ± 13	4.4 ± 0.4
1418 + 546 (OQ 530)	20.04.80	15″	*K*	9.2 ± 3.4	89 ± 18	11.1 ± 0.2
			H	9.7 ± 6.1	76 ± 18	8.6 ± 0.2
	10.07.80	15″	*K*	<0.9	~113	9.2 ± 0.2
1514 – 241 (AP Lib)	18.04.80	15″	*K*	7.0 ± 1.4	177 ± 5.7	26.4 ± 0.5
			H	7.4 ± 2.4	175 ± 9.4	20.6 ± 0.4
1641+ 399 (3C 345)	10.07.80	10″	*K*	12.8 ± 3.9	100 ± 8.9	6.7 ± 0.2
1921 – 293 (OV-236)	20.04.80	15″	*K*	13.9 ± 4.0	165 ± 8.3	6.2 ± 0.1
2155 – 304 (PKS)	09.07.80	15″	*K*	3.0 ± 1.4	124 ± 13	36.8 ± 0.3
			H	1.6 ± 0.4	135 ± 7.2	37.5 ± 3.1
			J	<1.3	~119	33.3 ± 0.6
	10.07.80	15″	*K*	1.7 ± 1.0	113 ± 17	37.1 ± 0.3
	11.07.80	15″	*K*	2.9 ± 0.3	130 ± 2.8	40.3 ± 0.4
2223 – 052 (3C 446)	21.12.79	10″	*K*	16.3 ± 3.8	137 ± 6.6	9.4 ± 0.2
	09.07.80	15″	*K*	10.9 ± 2.4	27 ± 6.4	11.0 ± 0.2
	10.07.80	15″	*K*	8.5 ± 0.8	42 ± 2.8	4.2 ± 0.2
			H	8.4 ± 5.0	41 ± 17	6.2 ± 0.5
	11.07.80	15″	*K*	11.1 ± 3.8	34 ± 9.7	9.1 ± 0.2
			H	5.5 ± 2.7	29 ± 14	5.5 ± 0.5
2254 + 074 (OY 091)	09.11.80	12″	*K*	17.4 ± 6.9	21 ± 11	7.1 ± 0.4

(1) Object	(2) Date	(3) Aperture	(7a) $\{S \pm \sigma(S)\}_K$	(7b) $\{S \pm \sigma(S)\}_H$	(7c) $\{S \pm \sigma(S)\}_J$
0912 + 297 (OK 222)	21.04.80	15″	5.4 ± 0.1	2.9 ± 0.1	–
1400 + 162 (OQ 100)	20.04.80	15″	1.5 ± 0.1	–	–
1652 + 398 (Mk 501)	18.04.80	10″	32.6 ± 0.6	29.8 ± 0.6	24.3 ± 1.2
	20.04.80		41.8 ± 0.8	45.5 ± 1.4	37.2 ± 1.8
1727 + 503 (I Zw 186)	20.04.80	15″	4.6 ± 0.3	4.5 ± 0.4	–
	21.04.80	15″	3.7 ± 0.4	–	–
	09.07.80	15″	4.2 ± 0.4	–	–
	10.07.80	15″	3.4 ± 0.2	–	–
	09.08.80	19″	3.3 ± 0.2	3.8 ± 0.2	–
	10.08.80	19″	3.3 ± 0.2	3.6 ± 0.2	–

3.1 0235+164

The extraordinary BL Lac object AO 0235+164 has absorption redshift systems at $z=0.524$ and $z=0.852$ (Rieke *et al.* 1976; Burbidge *et al.* 1976) and is a variable from radio to optical frequencies (Spinrad & Smith 1975). Variable 21cm absorption has been detected in the lower redshift system (Roberts *et al.* 1976; Wolfe 1980), and during two bursts there was evidence for correlated radio/optical activity (Ledden, Aller & Dent 1976; Balonek & Dent 1980). The optical polarization has a range of 5–26 per cent (Angel & Stockman 1980) and in 1975 the radio polarization position angle showed a remarkable linear rotation of 130° in an epoch of otherwise stochastic variations (Ledden & Aller 1978). No infrared polarimetry of this object has been published.

AO 0235+164 is the best-studied object in our sample, with data from four of the five observing sessions. The source has a large and variable degree of polarization at 2.2 μm and underwent a polarization burst in 1979 December, which is discussed in detail by Impey, Brand & Tapia (1982). By 1980 July, the 2.2 μm flux had fallen by more than a factor of 2 from its peak in December and the source was still highly polarized. Only in 1980 November, was the 2.2 μm polarization low, and through the observations the position angle spanned a range of $\sim 150°$. The data show wavelength dependence of polarization with the degree of polarization increasing towards shorter wavelengths, a trend verified by simultaneous optical polarimetry during the December burst. There is evidence both for position angle rotation and a change in the spectral shape over the five consecutive nights monitoring in December.

3.2 0306+103

0306+103 was found to be an optical variable with a featureless spectrum by Leacock *et al.* (1976). It is a highly variable BL Lac candidate with an archival range of variability of $\Delta m_B > 2$ mag (Leacock *et al.* 1976). On two separate observing periods this source was found to be less than 3 per cent polarized at 2.2 μm and the fluxes were identical. However, polarization in BL Lac objects is episodic and given the general correlation between polarization and variability (Stein, O'Dell & Strittmatter 1976) it would be surprising if high polarization was not eventually detected in this object.

3.3 0735+178

This object was shown to have an absorption redshift of $z=0.424$ and high and variable optical polarization by Carswell *et al.* (1974). Further optical polarimetry was presented by Nordsieck (1976) and Puschell & Stein (1980) and the monitoring of Angel *et al.* (1978) shows moderate stability around $p \sim 22$ per cent compared with the wild fluctuations of previous data. Rieke *et al.* (1977) obtained a single measurement of $p=18.5$ per cent at 2.2 μm. The infrared data presented here show that on each occasion the polarization at 2.2 μm was greater than 10 per cent. Over five nights in December, 0735+178 was stable in both polarization and position angle which is unusual behaviour in a BL Lac object (Angel & Stockman 1980). In 1980 April, the 2.2 μm polarization changed by 5 per cent in successive nights and over the entire period of observation the flux varied by nearly a magnitude. In common with AO 0235+164, 0735+178 shows polarization which increases towards shorter wavelengths.

3.4 0754+101

Optical polarimetry of this source has a range of 3–26 per cent, including a change from 25 to 12 per cent on successive nights (Tapia *et al.* 1977; Craine, Duerr & Tapia 1978). The

rapid, variable polarization was confirmed by Angel *et al.* (1978), who observed a change from 5 to 10 per cent in only 4 hr. This is the most rapid change of polarization observed in any BL Lac object. The infrared polarization is also variable (4–14 per cent), and there is evidence for a position angle rotation between the visible and the infrared regions (Puschell & Stein 1980; Rieke *et al.* 1977). Our data over four consecutive nights show a smooth increase in polarization from 2 to 9 per cent, while the total flux decreases by ~ 30 per cent. These changes may indicate dilution between a polarized and an unpolarized component in the source. The two nights' observation at 1.65 μm also follow the same trend.

3.5 0851+202

OJ 287 is one of the most rapidly variable BL Lac objects known. It has a tentative emission line redshift of $z = 0.306$ (Miller, French & Hawlet 1978) and has shown large amplitude ($\Delta m_B > 4.0$ mag) and rapid ($\Delta t < 1$ hr) variability at radio, infrared and optical wavelengths (Visvanathan & Elliot 1973; Veron & Veron 1975; Wolstencroft, Gilmore & Williams submitted). There is a good evidence in the photometry of Epstein *et al.* (1972) and the polarimetry of Rudnick *et al.* (1978) that a common emission process is acting from radio to optical wavelengths. The optical polarization shows a range of 1–29 per cent (Angel & Stockman 1980) and the 8 GHz polarization of 15 per cent measured by Aller & Ledden (1978) is the highest observed in the integrated radio flux of an extragalactic object. Two single infrared measurements of low precision have been published by Rudnick *et al.* (1978) and Puschell & Stein (1980).

As found by Angel *et al.* (1978) at optical wavelengths, OJ 287 is characterized by sudden changes superimposed on smoothly varying flux and polarization. Kikuchi *et al.* (1976) found evidence for periods of wavelength-dependent polarization, but there is no evidence for it here over a baseline of 1–2 μm.

In five nights of monitoring in 1979 December, OJ 287 showed a smooth increase of 2.2 μm polarization from 5 to 10 per cent. The position angle was stable until there was a sudden rotation of ~ 35° on the fifth night, with similar changes being seen in all three infrared wavebands. Four months later, the infrared flux changed by ~ 50 per cent between successive nights with no appreciable change in the degree of polarization. Both p and θ were near the December levels.

3.6 0912+297

This source was identified by Wills & Wills (1976), and was found to have a range of optical polarization of 0–13 per cent by Kinman (1976). The maximum radio polarization of 14.7 per cent is one of the highest seen in any BL Lac object (Wardle 1978). 0912+297 was not highly polarized when observed in 1980 April, but the flux measurement at 2.2 μm combined with the variability limits in Craine (1977) indicate that $V-K \gtrsim 3.5$, which is typical for BL Lac objects.

3.7 1147+245

1147+245 was also first identified by Wills & Wills (1976) and has no published redshift. Optical polarimetry by Kinman (1976), Serkowski & Tapia (1975) and Angel *et al.* (1979) shows it to have a range of 1.5–13 per cent with wide variations in position angle. Two measurements presented here, the first in the infrared, give a polarization of 10–15 per cent at a position angle of ~ 110°. The 2.2 μm flux increased by 15 per cent over a three-night period.

3.8 1308+326

B2 1308+326 has absorption-line redshifts at $z=0.996$ and $z=0.879$ (Miller *et al.* 1978), and at its brightest is one of the most luminous objects in the Universe (Liller 1976). Puschell *et al.* (1979) reported optical polarimetry with variations on a time-scale of 15 min, and high and variable infrared polarization. Moore *et al.* (1980) give polarimetry from the same outburst and find the optical and infrared variations to be correlated. 1308+326 has a history of strong bursts at radio, infrared and optical wavelengths. Our infrared polarimetry confirms the rapid variability found by other investigators. In 1979 December, the 2.2 μm polarization fell by 12 per cent in two nights, then rose by 6 per cent over the next 24 hr. By 1980 April, the flux had fallen by nearly 2 mag from its December peak and the source had dipped below the polarization detection limit of the instrument. However, another rapid optical flare was reported in July (Wisniewski, Mufson & Pollock 1980) with a peak near July 4 and changes of $\Delta m_B \sim 0.9$ in 24 hr. On July 10, the 2.2 μm polarization was ~ 8 per cent and the position angle had rotated by fully 90° since the December measures. This is obviously an extremely interesting object to monitor on all time-scales.

3.9 1400+162

1400+162 has a weak emission line redshift of $z=0.245$ (Miller *et al.* 1978) and is one of only two BL Lac objects with double-lobed radio structure. The position angle of the optical polarization is ~ 15° away from the axis of the extended radio structure (Baldwin *et al.* 1977). In 1980 April, the source was too faint for infrared polarimetry, with a K magnitude 0.6 mag fainter than the flux given in the Baldwin paper.

3.10 1418+546

OQ 530 was confirmed as a BL Lac object by Craine *et al.* (1978), who found high polarization in a photographic survey, and subsequently obtained photoeletric measures in the range 5–19 per cent. The object has an archival variability of $\Delta m_B \sim 4.8$ and is variable on time-scales from days to years (Miller 1978). In 1980 April, the polarizations at K and H were both ~ 9 per cent. Three months later, the source was 20 per cent fainter and the K polarization was undetectable, with an upper limit of ~ 1 per cent.

3.11 1514−241

AP Librae is located in the centre of an elliptical galaxy with an emission line redshift of $z=0.049$ (Disney, Peterson & Rodgers 1974; Miller *et al.* 1978). The range of optical polarization is 2–7 per cent, and the position angle is relatively stable at $\theta = 6 \pm 16°$ (Angel *et al.* 1978; Strittmatter *et al.* 1972). Capps & Knacke (1978) found 2.2 μm polarization in the range 3–6 per cent, but with a position angle ~ 90° different from the optical data. We report measuring K and H polarizations of ~ 7 per cent at a position angle (~ 175°) which confirms the long-term stability of the optical measurements. The data was obtained through a 15 arcsec aperture and contributions from the distributed starlight may lead to a higher than usual error on this observation.

3.12 1641+399

3C 345 is a violently variable, strong-lined source with an emission redshift of $z=0.595$ (Burbidge, Crown & Smith 1977). *VBLI* measurements show it to be expanding super-luminally at a rate of ~ 5c (Kellerman 1978). Kinman (1977) has studied eight years of

polarization data and finds a preferred position angle (80–110°) when the source is bright which is roughly aligned with the *VLBI* expansion axis (~ 105°). Knacke, Capps & Johns (1979) found strong wavelength dependence of polarization in this source. The increase from < 6 per cent in the optical region to 32 per cent at 2.2 μm is in the opposite sense to the two cases presented in this work. The single 2.2 μm measurement of $p = 13$ per cent in this paper has a position angle (~ 100°) well aligned with the super-luminal expansion axis.

3.13 1652+398

Mk 501 is an extended object, with a redshift of $z = 0.034$ based on emission lines in the galaxy component (Ulrich *et al.* 1975). The optical polarization is low (2–4 per cent) and its wavelength dependence has been used to separate the galaxy and non-thermal components of this source (Maza, Martin & Angel 1978; Ulrich *et al.* 1975). Although it was bright in the infrared when our observations were made, the polarization was too low to be detected. However, over a three-day period the flux rose by ~ 35 per cent in the *J*, *H* and *K* wavebands, which represents a considerable brightening of the non-thermal component.

3.14 1727+503

The compact galaxy I Zw 186 has an emission-line redshift of $z = 0.055$ (Oke 1978). The optical polarization is low (4–6 per cent), but the radio (3.7 cm) polarization is surprisingly high for an extended BL Lac object, at 11.3 per cent (Kinman 1976; Wardle 1978). The source was not highly polarized at 2.2 μm, and photometry over three observing sessions shows moderate (~ 25 per cent) variability of the compact component.

3.15 1921−293

The radio source OV-236 was found to be a variable in a 90 Ghz survey of flat spectrum source (Landau, Epstein & Rather 1980). During its outbursts OV-236 is the strongest quasar in the sky at 30 GHz and there was evidence for correlated radio/optical activity during the period 1977–79 (Gilmore 1980). In 1980 April, this source was confirmed as a BL Lac object by a 2.2 μm polarization measurement of 14 ± 4 per cent. Subsequently, the 2.2 μm flux decreased by ~ 35 per cent, so the source displays another BL Lac-type property of infrared variability. Wills & Wills (1981) have recently published a redshift of $z = 0.353$ for OV-236.

3.16 2155−304

H2155−304 is a *HEAO 1* X-ray source, with an optical indentification that has a tentative redshift of $z = 0.17$ (Schwartz *et al.* 1979; Charles, Thorstensen & Bowyer 1979). It is a rapid X-ray and optical variable, with an optical polarization of 5 per cent (Snyder *et al.* 1980; Griffiths *et al.* 1980). In 1980 July, the infrared polarization at three wavelengths and on three nights was < 3 per cent at a position angle of 120–130°. There were no examples of large or abrupt flux changes between nights or over a baseline of one month. The source has a $J-K$ colour of ~ 1, which is less than usual for a violently variable quasar or BL Lac object.

3.17 2223−052

3C 446 has strong emission lines and a redshift of $z = 1.404$ (Burbidge *et al.* 1977). Optical variability was observed by Visvanathan (1973) and Miller (1978) noted that at maximum

light the spectrum had much weaker lines characteristic of a BL Lac object. The optical polarization ranges from 4 to 17 per cent, and Stockman & Angel (1978) noticed a position angle rotation of $\sim 10^\circ$ in a few hours. The polarization position angle covers the complete possible range in 3C 446. In 1979 December, 3C 446 was found to be highly polarized at 2.2 μm ($p = 16$ per cent). By 1980 July, the polarization was down to 10 per cent and there had been a large swing in position angle, but the flux was nearly identical. 3C 446 has a very strong infrared excess with a $K-H$ colour of ~ 0.9, and is evidently a transitional case between quasars and BL Lac objects.

3.18 2254+074

OY+091 has no published redshift and is surrounded by a large number of very faint galaxies (Craine, Johnson & Tapia 1975). Optical variability was noted by Pollock (1975), and a range of optical polarization from 14 to 21 per cent has been measured by Kinman (1976). It was found to be highly polarized (~ 15 per cent) at 2.2 μm in a single measurement made in 1980 November.

4 Discussion

Of the 18 BL Lac objects in the sample, 11 have published redshifts. The original definition of BL Lacs as having featureless spectra has been relaxed since high resolution, good signal-to-noise spectra have revealed weak emission and absorption lines in many objects. The known BL Lac objects form a very inhomogeneous sample, so this study makes no attempt at completeness. For example, the selection techniques for BL Lac objects heavily favour radio-emitters and the identification procedures select against strong infrared-emitters (Rieke & Lebofsky 1979). A more practical limitation on this sample was that the sources had to be bright enough in the infrared for polarimetry. There is also the ubiquitous effect that the most well-studied objects are those which displayed the most unusual behaviour. Nevertheless, the sample includes one-third of the objects in the compilation by Angel & Stockman (1980), and is large enough to be considered representative of the infrared properties of BL Lac objects.

Twelve of the sample were observed on more than one occasion and 16 were observed at more than one wavelength. The principal wavelength of the study was 2.2 μm. The properties of the sample can be summarized as follows. High and/or variable linear polarization was observed in 14 out of 18 objects. Nine BL Lacs had $P_{MAX} > 12$ per cent and two had $P_{MAX} > 25$ per cent. The polarized and total flux measurements for 46 observations at 2.2 μm are plotted in Fig. 1, showing that the distribution of polarizations is weighted well away from the detection limit of the instrument. Most objects were variable in flux, polarization and position angle on all time-scales investigated. The minimum time-scale probed was fixed observationally at one day, but we can define the following maximum rates of change from the database: flux – 50 per cent/day, degree of polarization – 12 per cent/day and position angle – 30°/day.

4.1 RANGE OF INFRARED LUMINOSITY

The near-infrared luminosity of 11 BL Lac objects can be calculated assuming (a) the cosmological nature of the redshift and (b) isotropic emission. The luminosity in the rest frame of the BL Lac object is given by $L(\nu) = 4\pi\,(cz/H_0)^2(1+z)^{-1}F[\nu/(1+z)]$ assuming a Friedmann model with cosmological constant $\Lambda = 0$, and taking $H_0 = 75\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $q_0 = 1$.

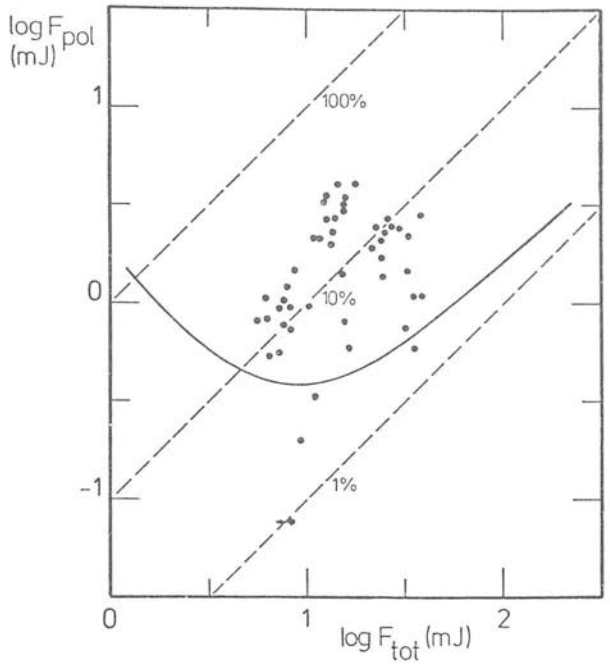


Figure 1. Polarized flux plotted against total flux for 46 observations in the K band. The solid line corresponds to the detection limit as a function of the magnitude and polarization of the source.

Choosing $q_0=0$ would increase the luminosity by $\sim(1+z/2)^2$ (Mattig 1958). The observed 2.2 μm fluxes are converted into 2.2 μm fluxes in the BL Lac rest frame by $F[\nu/(1+z)] = F(\nu)(1+z)^{\alpha}$ where α is the near-infrared spectral index from 1 to 2 μm. The flux distribution-dependent part of the K correction will be small. Table 2 contains the redshift (1), distance (2), the maximum 1–2 μm luminosity observed (3), the range of luminosity (4), and the number of separate obervations of each object (5). The most striking feature is the enormous range of infrared luminosities: a factor of 3×10^3. It is a range which extends from 1727+503 ($8\times10^9 L_{\odot}$) with a normal infrared flux for an elliptical galaxy to 3C 446 with a peak luminosity of 8.3×10^{46} erg s^{-1}, making it one of the most luminous objects in the

Table 2. Luminosities and their variations for 11 BL Lac objects.

Object	(1) z	(2) d (Mpc)	(3) L_{MAX} (erg s^{-1})	(4) L_{MAX}/L_{MIN}	(5) No. Obsns
0235+164	0.582	3410	6.6×10^{46}	2.2	12
0735+178	0.424	1700	1.7×10^{46}	1.4	9
0851+202	0.306	1220	7.2×10^{45}	1.9	7
1308+326	0.996	3980	3.9×10^{46}	5.8	6
1400+162	0.244	970	2.7×10^{44}	–	1
1514–241	0.049	200	1.7×10^{44}	–	1
1641+399	0.595	2380	8.3×10^{45}	–	1
1652+398	0.034	140	1.3×10^{44}	1.3	2
1727+503	0.055	220	3.2×10^{43}	1.4	6
1921–293	0.353	1400	2.4×10^{45}	1.4	3
2155–304	0.17	680	3.0×10^{45}	1.2	6
2223–052	1.404	5620	8.3×10^{46}	1.2	4

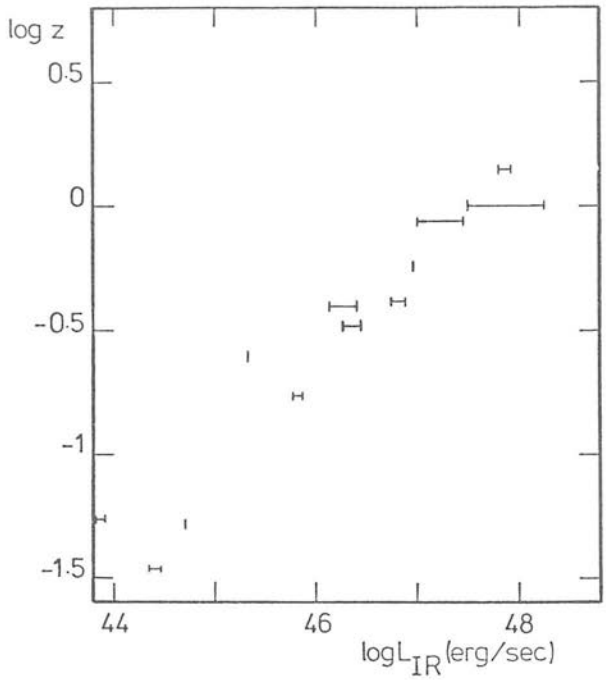


Figure 2. Redshift versus infrared luminosity for 11 BL Lac objects. The range of luminosity from the monitoring is shown as a horizontal error bar. A bandwidth of $1-10\,\mu$m has been used for the infrared flux.

Universe. Fig. 2 shows the distribution of redshift against luminosity, with the selection effect in favour of distant, luminous sources apparent. The luminosity spread is greater than the amplitude of variability in a given object; the most extreme archival variability is ~ 6.7 mag for 3C 279 which is a factor of 500 (Eachus & Liller 1975). It is therefore reasonable to question whether objects like Mk 501, AP Lib and 1727 + 503 are powered by the same source as the three ultraluminous objects 0235 + 164, 1308 + 326 and 3C 446. The 11 sources *all* have the common properties of variability, polarization infrared excess and power law spectra; these are the features which indicate a non-thermal emission process, and lead to a BL Lac classification.

The estimates of continuum peak luminosity are conservative since only a limited infrared bandwidth has been used. For any spectral index $\alpha > 1$ (where $\alpha = -d \ln F/d \ln \nu$), the power per unit bandwidth will be greater at $2.2\,\mu$m than in the blue part of the optical region. If, as shown by Elias *et al.* (1978), the strong radio and millimetre fluxes of BL Lac-type objects form a smooth extrapolation of the near infrared data, then most of the energy is emitted in the range $1-100\,\mu$m (in the absence of a strong, hard X-ray component). Even these underestimated energy densities pose problems for theories of the BL Lac power source. The peak luminosities of three of the sources require mass inflow rates of $\sim 100 M_\odot$ per year, if the mass–energy conversion efficiency is 10 per cent. Although these high rates do not have to be maintained continuously for long periods, the accumulation of large masses of gas near the nucleus is a problem (Gunn 1978). Fig. 3 shows a plot of minimum variability time-scale against infrared luminosity (adapted from Elliot & Shapiro 1974) where the diagonal line represents the maximum energy density for accretion on to a supermassive, rotating object. The three ultraluminous sources cross the line, providing good evidence that non-spherical accretion and relativistic accretion and relativistic source motion apply to at least some

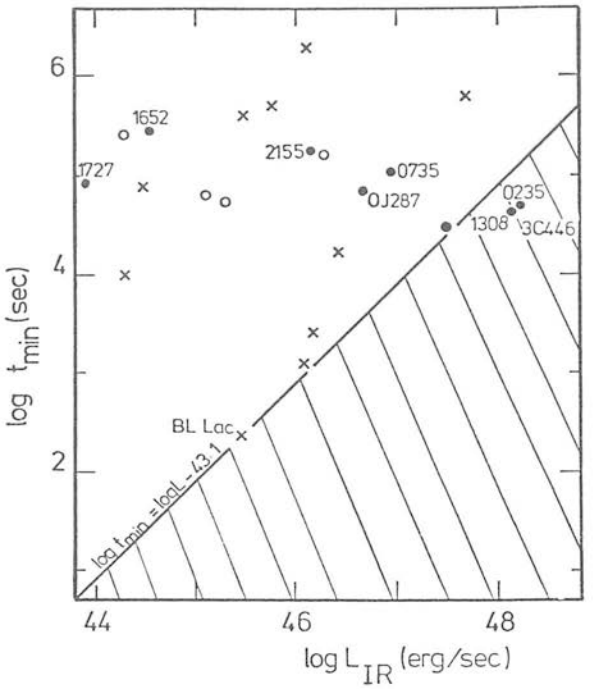


Figure 3. Minimum time-scale of variability versus infrared luminosity (Elliot & Shapiro 1974). A band-width of 1–10 μm has been used for the infrared flux. ×, Optical quasar data from Elliot & Shapiro (1974). ○, Optical quasar data from Gilmore (1978). ●, Infrared observations from this work.

BL Lac objects. Ways of reducing the energy budget in BL Lac objects will be discussed in a future paper (Brand & Impey, in preparation).

4.2 RELATIONSHIP BETWEEN POLARIZATION, VARIABILITY AND IR-EXCESS

It is clear that there is a strong link between the degree of polarization, the amplitude of variability and the near-infrared spectral slope of this sample. Table 3 shows the maximum 2.2 μm polarization (1), the degree of optical variability from the literature (2) and the near-infrared spectral index measured between J(1.25 μm) and K(2.2 μm) for a total of 12 BL Lac objects. Fig. 4(a) shows the archival range of (photographic) variability plotted against maximum 2.2 μm polarization from this study. There is no strict correlation between ΔB and P_{IR}^{MAX} but every object which is highly polarized (> 15 per cent) is also an extreme variable. The diagram also shows the average variability and polarization of a complete radio sample, and it is clear that BL Lac objects form the extreme tail of the properties of radio sources. Undoubtedly selection effects contribute to this diagram, since variables have been a natural target for polarimetry. However, Angel & Stockman (1980) have pointed out that many polarized objects were identified as radio sources before their variability was studied.

Fig. 4(b) plots near-infrared spectral index against maximum 2.2 μm polarization and the correlation is striking, especially considering the inhomogeneity of the data. It is recognized that broad-band colours over a baseline of 1–2 μm are not ideal for defining a continuum slope and a power law spectral index. For example, Frogel *et al.* (1978) have shown how the infrared colours of low-z elliptical galaxies are susceptible to the strong and complex absorption features of their constituent stars. Also, the near-infrared colours of quasars are

Table 3. Maximum polarization, variability and spectral index.

Object	(1) P_{MAX}^{IR}	(2) P_{MAX}^{OPT}	(3) ΔB	(4) α_{IR}
0235 + 164	34.6	44	5.2	2.1
0306 + 102	3.2	–	2.2	0.54
0735 + 178	29.9	31	2.5	1.1
0754 + 101	11.7	26	1.0	0.95
0851 + 202	14.3	32	4.0	0.89
0912 + 297	–	13	1.9	1.8
1147 + 245	14.8	13	1.4	1.3
1308 + 326	19.6	25	5.6	1.3
1418 + 546	9.7	19	4.8	0.98
1514 – 241	7.4	7	2.5	0.85
1641 + 399	12.8	16	2.0	–
1652 + 398	–	4	0.5	0.45
1727 + 503	–	6	1.9	–0.12
1921 – 293	13.9	8	2.5	1.4
2155 – 304	3.0	7	–	0.63
2223 – 052	16.3	17	3.4	1.4
2254 + 074	17.2	21	1.6	1.1

strongly affected by the appearance of redshifted emission lines (especially Hα) through the filters. The case for a power law in BL Lac objects rests on the absence of strong emission lines and the dominance of the non-thermal component. Soifer & Neugebauer (1980) have summarized the evidence that the optical/infrared continua of BL Lacs are well approximated by power laws. Fig. 4(b) indicates that sources with a larger infrared excess tend to

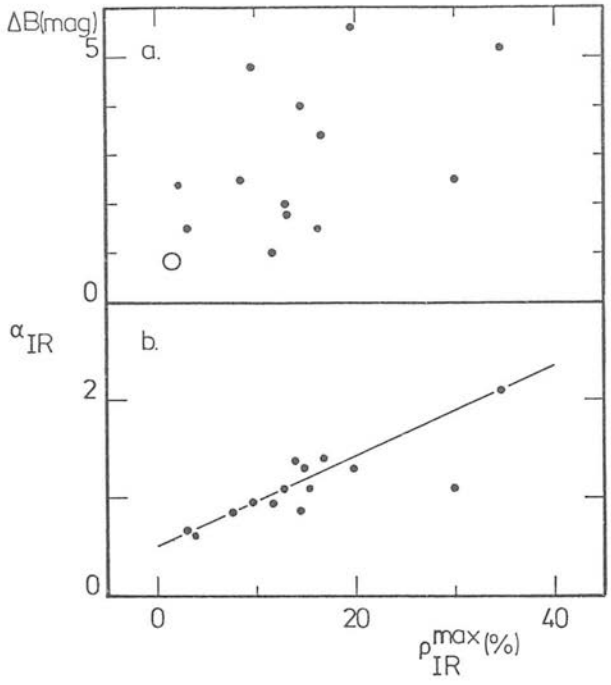


Figure 4. (a) Optical variability against maximum polarization. Open circle represents average of complete radio sample (Grandi & Tifft 1974) for ΔB and using Stockman & Angel (1978) for P_{MAX}. Other values for ΔB from Angel & Stockman (1980). (b) IR spectral index against maximum polarization. All data from this work. Correlation coefficient for 14 points is $r^2 = 0.81$ with $\sigma(p) = 9$ per cent and $\sigma(\alpha) = 0.38$.

have higher maximum polarizations. It is well known that polarized and variable BL Lac objects tend to have steeper optical/infrared spectra than (generally unpolarized) QSOs. However, this is the first time that a correlation between spectral shape and polarization has been noted *within* a sample of BL Lac objects. It is noteworthy that both the relationships discussed here are independent of the interpretation of the distances of BL Lac objects.

4.3 CHANGES IN POLARIZATION AND FLUX

One advantage of polarimetry is that it is a differential technique: the degree of polarization is a dimensionless quantity and does not require calibration of the total intensity. Changes in ratios of the Stokes parameters Q/I and U/I can be measured to a limit determined only by photon statistics. For this reason, some of the comprehensive monitoring programmes have not used continual intensity calibrations (e.g. Angel *et al.* 1978). However, one of the aims of this programme was to investigate the relationship between total and polarized flux in BL Lac objects. So the photometry has been calibrated and corrected for extinction, enabling a study of the three Stokes parameters Q, U and I. Inevitably, the calibration procedure means that sensitivity to small changes (2–3 per cent) in the total and polarized flux is lost. Only by accumulating data it is possible to look for bulk trends in transitions shown by BL Lac objects.

Fig. 5(a) shows internight changes in polarized and total flux plotted in absolute units. The data include differences between consecutive nights' observation for all objects monitored. Measurements at J, H and K are represented. The diagram can be used to understand the mechanism for variability in BL Lac objects. For example, it has been suggested that the source might consist of highly polarized subunits, with the flux and polarization variability caused by the dilution of the varying contributions of the subunits (Blandford & Rees 1978). It is also important to determine whether the polarization is imprinted in the emitting region or generated along the line-of-sight.

The collected data shows that there is a broad relationship between changes in total and polarized flux which is highly significant. However, one of the reasons that Fig. 5(b) is smeared out is because of the intrinsic spread in the properties of the sample. The two extremes of correlation and no correlation between polarization and flux are both interesting. Unfortunately, only a few objects have sufficient data points to populate the plot of $\Delta I_{\rm pol}$ versus $\Delta I_{\rm TOT}$ meaningfully. Fig. 5 shows the data for AO 0235 + 164 and the data for the combined pair 0735 +178/0851 +202, and the extremes of behaviour are apparent.

Changes in polarization and flux on 0235 +164 appear to be totally uncorrelated. For example, during the 1979 December burst, the 2.2 μm polarization rose by 10 per cent between consecutive nights while the flux increased by 10 per cent. However, on the tail of the event the polarization *fell* by ~ 12 per cent while the flux increased by 8 per cent. In 1308 +326 there was a substantial rise in polarization by 6 per cent in 24 hr while the total flux was constant within the errors. These variations are difficult to understand in terms of the canonical synchrotron picture. At infrared frequencies, propagation and optical depth effects are presumably small so the polarization changes reflect the magnetic geometry in the synchrotron source. A source structure must be found where changes in emitted flux are not tied to the strength and orientation of the magnetic field.

On the other hand, 0735 +178 and 0851 +202 present a monotonic behaviour where the polarization and flux changes are strongly correlated. Twenty-three out of 26 of the points for these objects lie in two quadrants of the diagram. The histogram of azimuth angle $\theta = \tan^{-1} \Delta I_{\rm pol}/\Delta I_{\rm TOT}$ shows a sharp peak around $\theta \simeq 8^{\circ}$ corresponding to $\Delta I_{\rm pol}/\Delta I_{\rm TOT} \simeq 0.10$. The width of the distribution is consistent with a single value of θ convolved with a

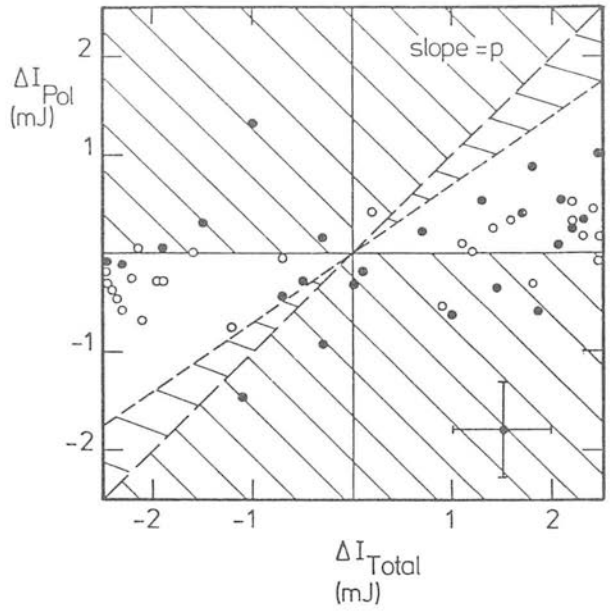


Figure 5. (a) Internight changes in polarized and total flux with the data divided into: •, 0235 +164 + the rest; ○, 0735 +178/0851+202. Shaded area corresponds to physically realizable area in a 'building block' model of the polarization changes (e.g. Blandford & Rees 1978).

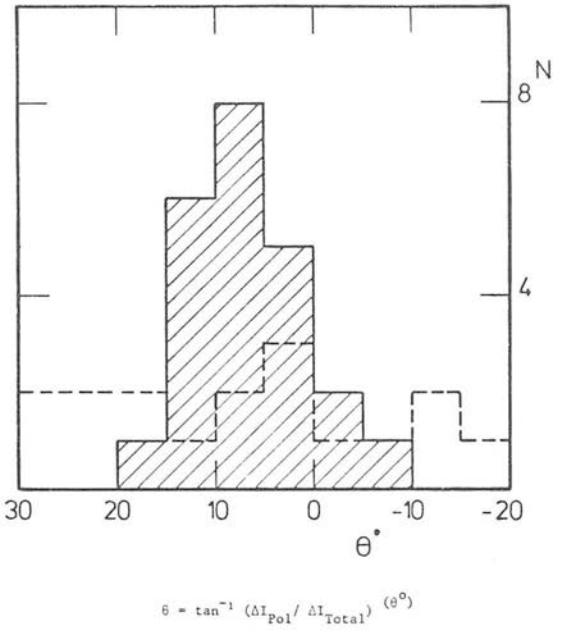


Figure 5. (b) Histogram of internight changes in polarized and total flux shown in terms of the ratio $\Delta I_{\rm pol}/\Delta I_{\rm TOT}$. Shaded area is 0735 +178 and 0851+202; unshaded area is 0235 +164 and the other monitored BL Lacs.

point-spread-function of half-width equal to the error bar on a single data point. In other words, the data for both objects can be interpreted in terms of a fixed ratio of $\Delta I_{\rm pol}/\Delta I_{\rm TOT}$. A natural framework for this interpretation is a two-component model where one component has a steady flux and arbitrary polarization and the other has a variable component of fixed polarization p. The observed changes are then caused by the fluctuations of the second component, and the polarization of the variable component is given very simply by $\Delta I_{\rm pol}/\Delta I_{\rm TOT} = \tan\theta$. From Fig. 5(b), $p \simeq 15$ per cent. It is important that static component of low (or zero) polarization does not represent an underlying galaxy. For dilution to apply the two components must have roughly equal intensity, and the luminosities of 0735+178 and 0851+202 preclude any contribution from a galaxy. Therefore, there are *two* non-thermal components in these objects and it is tempting to equate them to a jet and an accretion disc around the central power source.

4.4 THE $Q-U$ PLANE

A visually appealing way of presenting the polarization changes is a $Q-U$ diagram with the normalized Stokes parameters Q/I and U/I as the y and x axes. The length of the vector from the data point to the origin represents p, and the angle between the vector and the Q axis represents θ. The data for short monitoring runs on four objects are plotted in the $Q-U$ plane in Fig. 6. Contours of constant polarization are circles in the $Q-U$ plane, while contours of constant position angle are radii in the $Q-U$ plane. Even with relatively few data points, there are several interesting features of the variability which are not obvious from the tabulated data. The most striking case is AO 0235+164 where the polarization rose from a baseline of ~16 per cent at one position angle and fell back to the same baseline along a different position angle. The entire burst involved a position angle rotation of ~ 30°. It is not possible to do more than note the potential significance of these trajectories. However, if paths of constant polarization and position angle do dominate the variations of BL Lac

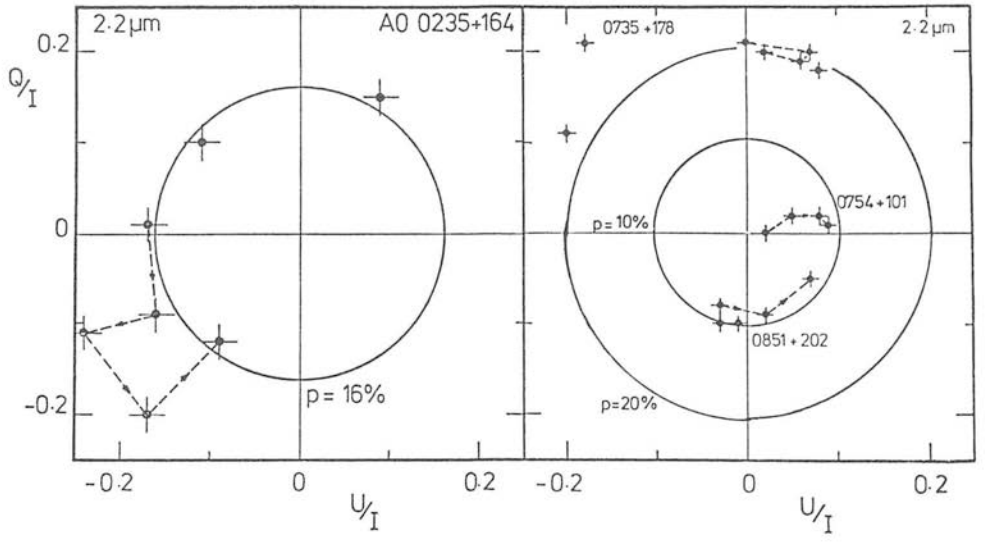


Figure 6. Normalized Stokes parameters for AO 0235+164, 0735+178, 0754+101 and 0851+202. Error bars for Q/I and U/I are approximated by $\sigma(p)/\sqrt{2}$. Trajectories of constant polarization are circles in the $Q-U$ plane; trajectories of constant position angle are radii in the $Q-U$ plane. The dashed lines connect observations made on consecutive nights.

objects, there will be important implications for the magnetic field geometry of the emitting region.

4.5 POSITION ANGLE ROTATIONS

Most objects in the sample have erratic variations in the position angle. There are small but significant changes between successive nights, and changes of much larger amplitude on a time-scale of months. Angel & Stockman (1980) have noted that BL Lac objects divide into those with a preferred position angle and those where the position angle takes all values. The latter species are nearly always the most luminous sources, which is confirmed by our data. The three most luminous sources in the sample (0235+164, 1308+326 and 2223−052) are also those with the widest variation in position angle. The relative stability of 0735+178 and 0851+202 shows that this is not simply a selection effect.

Apart from the five-day rotation of AO 0235+164 (Fig. 6), there was also an abrupt rotation in OJ 287. Fig. 7 shows five nights of monitoring of OJ 287, and after four nights at a steady angle there is a $\sim 30^\circ$ rotation within 24 hr. The rotation is of similar amplitude at all three wavelengths. An accumulation of radio monitoring data by Aller, Hodge & Aller (1981) shows that linear position angle rotations such as the one observed in AO 0235+164 (Ledden & Aller 1978) may be relatively commonplace. In that case the behaviour of AO 0235+164 and OJ 287 from our data is not unique, and continued monitoring will reveal many more changes in θ. At infrared frequencies, propagation effects are expected to be small, and position angle rotations directly reflect a rotation of the magnetic field in the emitting volume.

4.6 WAVELENGTH-DEPENDENT POLARIZATION

Only two of the BL Lac objects show wavelength-dependent polarization, 0235+164 and 0735+178. In both cases the effect is systematic between the three infrared wavelengths and

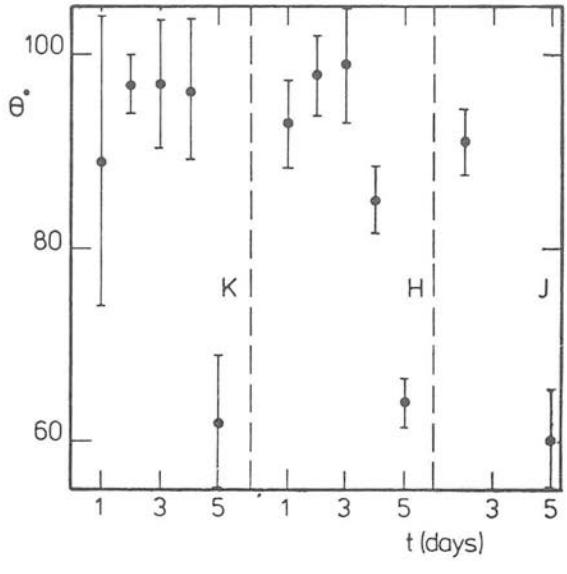


Figure 7. Position angle rotation in OJ 287. Measurements were made on five consecutive nights (see Table 1 for details), and in three wavebands.

persists during the night-to-night polarization variations. The shape of the $p(\lambda)$ curve for AO 0235+164 is constant during the large changes in degree of polarization which occurred during the December burst. The slope of the $p(\lambda)$ curve is variable in the case of 0735+178, being quite steep in 1979 December and nearly flat in 1980 November. The polarization increases with decreasing wavelength in both cases. Fig. 8 shows the $p(\lambda)$ plot for both objects at various times during the monitoring.

Most BL Lac objects show constant polarization over the optical part of the spectrum (Moore *et al.* 1980), and the few observations over the longer infrared baseline generally support this. The exceptions include OJ 287 (Kikuchi *et al.* 1976), 0735+178 (Nordsieck 1976) and BL Lac (Puschell & Stein 1980). The polarization increase towards the blue in BL Lac is too large to be caused by galactic dilution. For the low redshift objects Mk 421 and Mk 501, Maza *et al.* (1978) explained a slight wavelength dependence in the optical region with a two component model. It is straightforward to discover whether the wavelength dependence found here is intrinsic to the BL Lac object or the effect of thermal dilution from a surrounding galaxy. It is assumed that cosmological redshifts apply, and that the compact non-thermal source lies at the centre of a normal galaxy. To determine the visual luminosity, the infrared spectral index is used to extrapolate the spectral flux down to the appropriately redshifted optical window. $H_0 = 75\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $q_0 = 0$ are assumed throughout. The rest frame luminosities per unit frequency are:

$L(0.55\mu) = 4 \times 10^{31}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}$ AO 0235+164

$L(0.55\mu) = 6 \times 10^{29}\,\mathrm{erg\,s^{-1}\,Hz^{-1}}$ PK5 0735+178

First-ranked cluster galaxies (usually giant ellipticals) have a range of $-23.4 < M_v < -21.4$ (Burbidge 1974; Frogel *et al.* 1978), and the corresponding absolute magnitudes for 0235+164 and 0735+178 are $M_v = -28.9$ and -24.3 respectively. The visual luminosity of

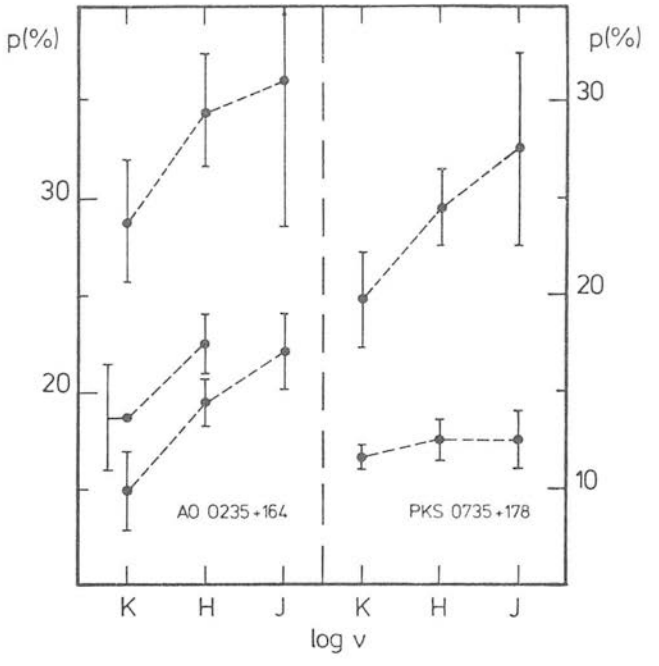


Figure 8. Wavelength dependence of polarization for AO 0235+164 and PKS 0735+178.

0235+164 is more than 100 times that of a first-ranked giant elliptical, therefore dilution from a galactic component has a negligible effect on the polarization measured for the source. 0735+178 is less luminous, but once again the wavelength dependence must be intrinsic to the non-thermal source. In the worst case, assume that the surrounding galaxy is at the top of the range of ellipticals (−23.4) and that the BL Lac object is at its faintest measured flux level. The only correction in the other direction occurs if the spectral index decreases at optical wavelengths, in which case the optical luminosity inferred from extrapolating the IR flux would be an over-estimate. Looking at the energy distribution of an elliptical galaxy in the BL Lac rest frame (Whitford 1971; Grasdalen 1980) between 1 and 2 μm, the relative contribution of the elliptical is decreasing because the BL Lac flux is rising rapidly and the galaxy flux is falling. Therefore the galaxian dilution will *decrease* with increasing wavelength, in the opposite sense to the observed wavelength dependence. There is no indication of extended image structure on Palomar Sky Survey plates of 0735+178, so it is reasonable to attribute $p(\lambda)$ entirely to the compact BL Lac object.

The degree of polarization of incoherent synchrotron radiation in a uniform magnetic field depends only on the spectral index of the flux distribution. For simple power law spectra, the linear polarization will be independent of wavelength. Nordsieck (1976) has shown that for objects with curve spectra, the polarization depends on the local spectral index. For the two objects described here, spectral curvature cannot be responsible for $p(\lambda)$, and there is no simple property of a synchrotron source which will lead to such a rapid $dp/d\lambda$ over such a short baseline. Ways of generating wavelength-dependent polarization will be discussed in a future paper.

5 Conclusions

The first results of a monitoring programme of BL Lac objects have been presented. Simultaneous infrared polarimetry and photometry of 18 objects show that, as in the optical region, high and variable linear polarization is a common characteristic that clearly distinguishes BL Lac objects from other active nuclei. The range of infrared luminosity is enormous; the three most extreme objects requiring non-spherical accretion and/or relativistic source motions to explain their energy production.

The position angles of some BL Lacs were found to vary widely, and systematic rotations on a time-scale of days were seen in two objects. Wavelength-dependent polarization was seen in only two BL Lac objects, but in these cases it is not due to dilution by an underlying galaxy and is difficult to explain in terms of the canonical synchrotron mechanism.

For one object the flux and polarization changes are completely unrelated, while for two others the night-to-night changes are consistent with a mixture of two non-thermal components.

Both the extension of wavelength baseline to infrared wavelengths and the simultaneous observation of flux and polarization are important developments in the monitoring of BL Lac objects.

Acknowledgments

We thank the Panel for the allocation of telescope time, and T. J. Lee and the UKIRT staff for excellent technical support under difficult conditions. It was through their efforts that the polarimeter was successfully interfaced to the UKIRT control software. Much credit is due to R. J. Beetles, W. A. Cormack and the staff of the ROE workshop for the design and construction of the Edinburgh polarimeter. CDI acknowledges the receipt of an SRC studentship.

References

Aller, H. D., Hodge, P. E. & Aller, M. F., 1981. *Astrophys. J.*, **248**, L5.

Aller, H. D. & Ledden, J. E., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 53, ed. Wolfe, A. M., University of Pittsburgh.

Angel, J. R. P. *et al.*, 1978. *Pittsburgh Conference on BL Lac Objects*, p. 117, ed. Wolfe, A. M., University of Pittsburgh.

Angel, J. R. P. & Stockman, H. S., 1980. *A. Rev. Astr. Astrophys.*, **18**, 321.

Baldwin, J. A., Wampler, E. J., Burbidge, E. M., O'Dell, S. L., Smith, H. E., Hazard, C., Nordsieck, K. H., Pooley, G. & Stein, W. A., 1977. *Astrophys. J.*, **215**, 408.

Balonek, T. J. & Dent, W. A., 1980. *Astrophys. J.*, **240**, L3.

Blandford, R. D. & Rees, M. J., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 328, ed. Wolfe, A. M., University of Pittsburgh.

Burbidge, E. M., 1974. *High Energy Astrophysics and Its Relation to Elementary Particle Physics*, p. 109, MIT.

Burbidge, E. M., Caldwell, R. D., Smith, H. E., Liebert, J. & Spinrad, H. 1976. *Astrophys. J.*, **205**, L117.

Burbidge, G. R., Crown, A. H. & Smith, H. E., 1977. *Astrophys. J. Suppl.*, **33**, 113.

Burstein, D. & Heiles, C., 1978. *Astrophys. J.*, **225**, 40.

Capps, R. W. & Knacke, R. F., 1978. *Astrophys. Lett.*, **19**, 113.

Carswell, R. F., Strittmatter, P. A., Williams, R. E., Kinman, T. D. & Serkowski, K. 1974. *Astrophys. J.*, **190**, L101.

Charles, P., Thorstensen, J. & Bowyer, S., 1979. *Nature*. **281**, 285.

Craine, E. R., 1977. *Quasi-stellar and BL Lac Objects*, Pachart.

Craine, E. R., Duerr, R. & Tapia, S., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 99, ed. Wolfe, A. M., University of Pittsburgh.

Craine, E. R., Johnson, K. & Tapia, S., 1975. *Publs astr. Soc. Pacif.*, **87**, 123.

Disney, M. J., Peterson, B. A. & Rodgers, A. W., 1974. *Astrophys. J.*, **194**, 179.

Dyck, H. M., & Beichman, C., 1974. *Astrophys. J.*, **194**, 57.

Eachus, L. J. & Liller, W., 1975. *Astrophys. J.*, **200**, L61.

Elias, J. H., Ennis, D. J., Gezari, D. Y., Hauser, M. G., Houck, J. R., Lo, K. Y., Matthews, K., Nadeau, D., Neugebauer, G., Werner, M. W. & Westbrook, W. E., 1978. *Astrophys. J.*, **220**, 25.

Elliot, J. L. & Shapiro, S. L., 1974. *Astrophys. J.*, **192**, L3.

Epstein, E. E. *et al.*, 1972. *Astrophys. J.*, **178**, L51.

Frogel, J. A., Persson, S. E., Aaronson, M. & Matthews, K., 1978. Astrophys. J., **220**, 25.

Gilmore, G., 1978. *PhD thesis*, University of Canterbury.

Gilmore, G., 1980. *Nature*, **287**, 612.

Grandi, S. A. & Tifft, W. G., 1974. *Publs astr. Soc. Pacif.*, **86**, 873.

Grasdalen, G. L., 1980. *IAU Symposium*, **92**, 269.

Griffiths, R. E., Briel, U., Chaisson, L., Davis, M., Doxsey, R. E., Hjellming, R., Schwartz, D. A. & Tapia, S., 1980. *Harvard Astrophysics preprint.*

Gunn, J. E., 1978. *Active Galactic Nuclei*, eds Hazard. C. & Mitton, S., University of Cambridge.

Heiles, C., 1976. *Astrophys. J.*, **204**, 379.

Impey, C. D., Brand, P. W. J. L. & Tapia, S., 1982. *Mon. Not. R. astr. Soc.*, **198**, 1.

Johnson, H. L., 1968. *Stars and Stellar Systems VII*, p. 167, University of Chicago.

Johnson, P. A., Rieke, G. H., Lebofsky, M. J. & Kemp, J. C., 1981. *Astrophys. J.*, **245**, 871.

Kellerman, K. J., 1978. *Physica Scripta*, **17**, 257.

Kemp, J. C., Wolstencroft, R. D. & Swedlund, J. B., 1972. *Astrophys. J.*, **173**, L113.

Kikuchi, S., Mikami, V., Konno, M. & Inove, M., 1976. *Publs astr. Soc. Jap.*, **28**, 117.

Kinman, T. D., 1976. *Astrophys. J.*, **205**, 1.

Kinman, T. D. 1977. *Nature*, **267**, 798.

Knacke, R. F., Capps, R. W. & Johns, M., 1979. *Nature*, **280**, 215.

Landau, R., Epstein, E. E. & Rather, J. O. G., 1980. *Astr. J.*, **85**, 363.

Landstreet, J. D. & Angel, J. R. P., 1972. *Astrophys. J.*, **174**, L127.

Leacock, R. J., Smith, A. G., Edwards, P. L., Pollock, J. T., Scott, R. L., Gearhart, M. R., Pacht, E. & Kraus, J. D., 1976. *Astrophys. J.*, **206**, L87.

Ledden, J. E. & Aller, H. D., 1978. *Pittsburgh Conference on BL Lac objects*, p. 60. ed. Wolfe, A. M., University of Pittsburgh.

Ledden, J. E., Aller, H. D. & Dent, W. A., 1976. *Nature*, **260**, 752.

Liller, W., 1976. *IAU Circular*, **2939**.

Mattig, W., 1958. *Astr. Nachr.*, **284**, 109.
Maza, J., 1979. *PhD thesis*, University of Toronto.
Maza, J., Martin, P. G. & Angel, J. R. P., 1978. *Astrophys. J.*, **224**, 368.
Miller, J. S., 1978. *Comm. Astrophys. Space Sci.*, **7**, 175.
Miller, J. S., French, H. B. & Hawley, S. A., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 176, ed. Wolfe, A. M., University of Pittsburgh.
Moore, R. L., Angel, J. R. P., Rieke, G. H., Lebofsky, M. J., Wisniewski, W. Z., Mufson, S. L., Urba, F. J., Miller, H. R., McGimsey, B. Q. & Williamson, R. M., 1980. *Astrophys. J.*, **235**, 717.
Nordsieck, K. H., 1972. *Astrophys. Lett.*, **12**, 69.
Nordsieck, K. H., 1976. *Astrophys. J.*, **209**, 653.
Oke, J. B., 1978. *Astrophys. J.*, **219**, L97.
Oke, J. B. & Schild, R. E., 1970. *Astrophys. J.*, **161**, 1015.
Pollock, J. T., 1975. *Astrophys. J.*, **198**, L53.
Puschell, J. J. & Stein, W. A., 1980. *Astrophys. J.*, **237**, 331.
Puschell, J. J., Stein, W. A., Jones, T. W., Warner, J. W., Owen, F., Rudnick, L., Aller, H. & Hodge, P., 1979. *Astrophys. J.*, **229**, L11.
Rieke, G. H., Grasdalen, G. L., Kinman, T. D., Hintzen, P., Wills, B. J. & Wills, D., 1976. *Nature*, **260**, 154.
Rieke, G. H. & Lebofsky, M. J., 1979. *A. Rev. Astr. Astrophys.*, **17**, 477.
Rieke, G. H., Lebofsky, M. J., Kemp, J. C., Coyne, G. V. & Tapia, S., 1977. *Astrophys. J.*, **218**, L37.
Roberts, M. S., Brown, R. L., Brandage, W. D., Rots, A. H., Haynes, M. P. & Wolfe, A. M., 1976. *Astr. J.*, **81**, 293.
Rudnick, L., Owen, F. N., Jones, T. W., Puschell, J. J. & Stein, W. A., 1978. *Astrophys. J.*, **225**, L5.
Sandage, A., 1972. *Astrophys. J.*, **178**, 1.
Schild, R. E., Peterson, B. M. & Oke, J. B., 1971. *Astrophys. J.*, **166**, 95.
Schwartz, D. A., Doxsey, R. E., Griffiths, R. E., Johnson, M. D. & Schwarz, J. 1979. *Astrophys. J.*, **229**, L53.
Serkowski, K., 1974. *Methods of Experimental Physics*, **12A**, 36, Academic Press.
Serkowski, K. & Tapia, S., 1975. *Bull. Am. astr. Soc.*, **7**, 499.
Snyder, W. A. *et al.*, 1980. *Astrophys. J.*, **237**, L11.
Soifer, B. T. & Neugebauer, G., 1980. *IAU Symposium*, **96**, 329.
Spinrad, H. & Smith, H. E., 1975. *Astrophys. J.*, **201**, 275.
Stein, W. A., O'Dell, S. L. & Strittmatter, P. A., 1976. *A. Rev. Astr. Astrophys.*, **14**, 173.
Stockman, H. S. & Angel, J. R. P., 1978. *Astrophys. J.*, **224**, L67.
Strecker, D. W., Erickson, E. F. & Witteborn, F. C., 1979. *Astrophys. J. Suppl.*, **41**, 501.
Strittmatter, P. A., Serkowski, K., Carswell, R. F., Stein, W. A., Merrill, K. M. & Burbidge, F. M., 1972. *Astrophys. J.*, **175**, L7.
Tapia, S., Craine, E. R., Gearhart, M. R., Pacht, E. & Kraus, J. D., 1977. *Astrophys. J.*, **215**, L71.
Ulrich, M. H., Kinman, T. O., Lynds, C. R., Rieke, G. H. & Ekers, R. D., 1975. *Astrophys. J.*, **198**, 261.
Veron, P. & Veron, M. P., 1975. *Astr. Astrophys.*, **39**, 281.
Vinokur, M., 1965. *Ann. D'Astrophys.*, **28**, 412.
Visvanathan, N., 1973. *Astrophys. J.*, **179**, 1.
Visvanathan, N. & Elliot, J. J., 1973. *Astrophys. J.*, **179**, 721.
Wardle, J. F. C., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 39, ed. Wolfe, A. M., University of Pittsburgh.
Wardle, J. F. C. & Kronberg, P. P., 1974. *Astrophys. J.*, **194**, 249.
Whitford, A. E., 1971. *Astrophys. J.*, **169**, 209.
Wills, D. & Wills, B. J., 1976. *Astrophys. J. Suppl.*, **31**, 143.
Wills, D. & Wills, B. J., 1981. *Nature*, **289**, 384.
Wisniewski, W. Z., Mufson, S. L. & Pollock, J. T., 1980. *IAU Circular*, **3500**.
Wolfe, A. M., 1980. 2nd European IAU Meeting on Variability in Stars and Galaxies.
Wolstencroft, R. D., Gilmore, G. & Williams, P. M., *Mon. Not. R. astr. Soc.*, in press.

Mon. Not. R. astr. Soc. (1984) **209**, 245–269

Infrared polarimetry and photometry of BL Lac objects – II

C. D. Impey *Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA*

P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

R. D. Wolstencroft *Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

P. M. Williams *United Kingdom Infrared Telescope Unit, 900 Leilani Street, Hilo, Hawaii 96720, USA*

Received 1983 November 15; in original form 1982 April 14

Summary. Photometry and polarimetry in the *JHK* wavebands have now been obtained for 25 BL Lac objects, more than doubling the amount of data contained in Paper I (Impey *et al.* 1982a). Several new objects have been monitored for periods of up to five days, and accumulated data is sufficient for a statistical analysis of polarization properties. The selection effects operating on this sample are examined first. A power-law spectrum is consistent with the spectra of all but three objects. Important new results are (1) The four most luminous sources show indirect evidence for relativistic bulk motion of the emitting volume; (2) The range of position angles from monitoring correlates well with infrared luminosity; (3) The maximum 2.2 μm polarization and maximum 1–2 μm spectral index are correlated; (4) The sample divides into two distinct parts: one where the change in total and polarized flux are closely correlated, and the other where they are not correlated at all; (5) Strong wavelength-dependent polarization is observed in only two out of thirteen objects; (6) The amplitude of polarization variations between day and month time-scales is constant, while the amplitude of flux variations increases; (7) There is good evidence that transitions in polarization and position angle are 'decoupled'; i.e., changes in position angle take place at a constant degree of polarization.

1 Introduction

This paper presents results from a continuing program to monitor the infrared polarization of BL Lac objects. Data from the first year of monitoring has previously been published

(Impey *et al.* 1982a, Paper I). Infrared polarimetry and photometry have now been obtained for 19 BL Lac objects, with photometry alone for another six. The data contained here doubles the amount of polarimetry in Paper I, and adds a considerable amount of new photometry. Where possible, polarization measurements have been made in more than one colour. The sample now represents a considerable fraction of the number of known BL Lac objects (Angel & Stockman 1980; Weiler & Johnston 1980). As expected, the infrared continua of many BL Lac objects are dominated by strongly variable flux and polarization. However, a sufficient database has now been established to examine the polarization properties of BL Lac objects statistically.

2 Observations

The observations were all carried out at the 3.8-m United Kingdom Infrared Telescope (UKIRT) on Mauna Kea using a photovoltaic InSb detector and rotating HR Polaroid. All the polarization data were collected in two separate runs during the period 1981 April/May. The photometry was collected over the period 1980 April to 1981 May. The two runs in 1981 April/May used the *f*/35 chopping secondary with star/sky chopping at a rate of 7–8 Hz. The instrumental parameters were re-determined for each run and the instrumental sensitivity and polarization were calculated an average of twice a night. The broad band filters used and their effective wavelengths were: J(1.25 μm), H(1.65 μm) and K(2.2 μm).

The calibration and reduction procedures for this project were described in Paper I, the same procedures have been followed for the data presented here. The analyzer efficiency and position angle zero point were re-determined each time the polarimeter was mounted on the telescope. The polarimetry was calibrated with nearby unpolarized stars, and the instrumental polarization was repeatable over a run to an rms of 0.3 per cent. Fluxes in absolute units have been corrected for the effect of the different flux distributions of the objects and the calibrators in the broad passbands. Repeated measurements of standard stars showed that the instrumental magnitudes during the two runs in 1981 April/May had rms deviations of no more than 1.8 per cent and 2.4 per cent. For faint BL Lac objects where polarimetry was not possible, the analyzer was removed from the light path to enable faster photometry to be carried out. Nine observations were made with the analyzer both in and out. Due to uncertainties in the primary calibrators, the zero point of the absolute flux scale is only accurate to about 5 per cent.

The BL Lac objects were acquired either directly on the integrating Quantex TV system or indirectly by offsetting from a nearby bright star with accurate coordinates. The offsetting was in general accurate to $\pm$ 0.5 arcsec, and in all cases was checked by peaking up on the 2.2 μm signal of the object. The beam positions were checked on each new object, and also whenever the telescope passed through the meridian. All the observations in 1981 April/May were made using a 10 arcsec aperture, and since all of the BL Lac objects were stellar, no aperture corrections were needed. Flat-topped beam profiles and accurate autoguiding ensured the integrity of the photometry. All the observations were made in or near dark of Moon.

3 Results

Individual objects are discussed below, with the exception of those whose polarimetric histories were given in Paper I. The results are presented in Table I, where column (1) gives the Parkes designation and any other source name, column (2) gives the date (UT) of the obser-

Table 1. Details of polarimetry and photometry for all the objects.

(1) Object	(2) Date	(3) Aperture	(4) Waveband	(5) $p \pm \sigma(p)$	(6) $\theta \pm \sigma(\theta)$	(7) $S \pm \sigma(S)$
0735+178 (PKS)	4 Apr 81	10"	K	13.1 ± 3.4	28 ± 7.1	11.6 ± 0.3
			H	15.0 ± 7.6	40 ± 7.6	9.8 ± 0.3
	5 Apr 81	10"	K	14.4 ± 4.6	17 ± 8.7	18.3 ± 0.7
			J	12.7 ± 4.5	170 ± 9.7	9.0 ± 0.7
	6 Apr 81	10"	K	14.0 ± 1.1	10 ± 2.1	16.5 ± 0.5
			H	16.6 ± 3.3	12 ± 5.4	9.2 ± 0.3
			J	20.9 ± 5.0	10 ± 6.5	5.5 ± 0.2
	7 Apr 81	10"	K	10.4 ± 2.6	12 ± 4.7	17.9 ± 0.5
			H	16.4 ± 2.0	12 ± 2.3	11.4 ± 0.2
	8 Apr 81	10"	K	15.4 ± 2.0	167 ± 3.5	18.4 ± 0.6
			H	15.3 ± 2.1	170 ± 2.7	13.7 ± 0.3
	30 Apr 81	10"	K	24.7 ± 1.1	155 ± 1.3	23.9 ± 0.7
			H	22.7 ± 2.7	155 ± 3.2	18.3 ± 0.9
	2 May 81	10"	K	26.4 ± 2.2	163 ± 2.3	28.2 ± 0.8
			H	29.6 ± 2.3	161 ± 2.1	20.4 ± 1.0
			J	32.6 ± 3.0	163 ± 2.5	14.7 ± 0.3
	3 May 81	10"	K	20.8 ± 0.8	163 ± 1.1	23.0 ± 1.2
			H	20.7 ± 1.8	167 ± 2.4	17.2 ± 0.9
			J	26.4 ± 3.7	162 ± 3.8	13.3 ± 0.7
	4 May 81	10"	K	22.7 ± 2.4	158 ± 2.9	16.8 ± 0.5
			H	23.5 ± 3.7	152 ± 4.3	7.3 ± 0.4
			J	13.3 ± 6.8	168 ± 14	6.2 ± 0.1
0754+101 (OI 090.4)	4 Apr 81	10"	K	18.9 ± 6.7	112 ± 9.7	10.9 ± 0.3
	5 Apr 81	10"	K	12.9 ± 4.8	109 ± 10	13.6 ± 0.6
	6 Apr 81	10"	K	12.1 ± 6.2	89 ± 13	12.0 ± 0.4
	7 Apr 81	10"	K	4.2 ± 3.4	172 ± 22	16.5 ± 0.5
	8 Apr 81	10"	K	2.7 ± 1.5	102 ± 15	12.0 ± 0.4
0818-128 (OJ-131)	7 Apr 81	10"	K	14.4 ± 2.5	122 ± 4.7	13.9 ± 0.4
			H	17.9 ± 6.1	128 ± 9.3	10.2 ± 0.3
			J	21.4 ± 9.9	126 ± 13	9.4 ± 0.2
	8 Apr 81	10"	K	13.7 ± 3.2	91 ± 6.2	11.9 ± 0.4
	2 May 81	10"	K	8.7 ± 2.1	91 ± 6.6	19.1 ± 0.6
			H	6.0 ± 1.9	104 ± 8.6	15.6 ± 0.8
	3 May 81	10"	K	14.0 ± 5.0	128 ± 4.8	16.7 ± 0.8
0851+202 (OJ 287)	5 Apr 81	10"	K	16.8 ± 3.6	151 ± 5.9	11.7 ± 0.4
	6 Apr 81	10"	K	12.0 ± 2.3	142 ± 5.2	15.2 ± 0.5
			H	15.6 ± 2.5	152 ± 4.4	10.3 ± 0.2
	7 Apr 81	10"	K	12.9 ± 2.4	165 ± 5.1	17.1 ± 0.5
			H	15.9 ± 3.0	150 ± 5.2	10.9 ± 0.2
	8 Apr 81	10"	K	11.9 ± 1.5	133 ± 3.4	14.9 ± 0.4
	2 May 81	10"	K	18.9 ± 4.2	125 ± 3.8	20.6 ± 0.6
			H	19.7 ± 1.9	124 ± 2.6	13.0 ± 0.7
			J	15.0 ± 1.4	113 ± 2.5	8.9 ± 0.2
0912+297 (OK 222)	6 Apr 81	10"	K	12.2 ± 5.0	2.9 ± 11.2	8.0 ± 0.2
	8 Apr 81	10"	K	5.1 ± 4.9	128 ± 26	6.8 ± 0.2

Table 1 – *continued*

(1) Object	(2) Date	(3) Aperture	(4) Waveband	(5) $p \pm \sigma(p)$	(6) $\theta \pm \sigma(\theta)$	(7) $S \pm \sigma(S)$
1156+295 (4C 29.45)	4 Apr 81	10"	K	11.2 ± 0.7	21.8 ± 1.7	54.2 ± 1.1
			H	12.5 ± 2.3	21.5 ± 5.0	34.5 ± 0.7
	5 Apr 81	10"	K	14.3 ± 0.9	31.6 ± 1.7	74.8 ± 1.5
			J	15.3 ± 1.7	22.3 ± 3.0	47.7 ± 1.4
	6 Apr 81	10"	K	10.0 ± 1.8	13.0 ± 4.9	40.7 ± 0.8
			H	10.3 ± 0.9	22.8 ± 2.4	27.2 ± 0.5
			J	10.4 ± 1.4	18.7 ± 3.7	22.0 ± 0.7
	7 Apr 81	10"	K	5.5 ± 1.1	0.5 ± 4.0	49.9 ± 1.0
			H	4.7 ± 2.1	166 ± 12	32.4 ± 0.6
			J	3.6 ± 2.1	161 ± 16	25.5 ± 0.8
	8 Apr 81	10"	K	8.9 ± 1.6	158 ± 4.9	32.9 ± 0.7
			H	11.3 ± 1.1	157 ± 2.7	23.9 ± 0.5
	30 Apr 81	10"	K	8.3 ± 5.4	54 ± 18	8.7 ± 0.3
	2 May 81	10"	K	8.3 ± 2.5	108 ± 8.2	12.0 ± 0.4
	3 May 81	10"	K	5.6 ± 0.9	161 ± 4.4	14.4 ± 0.6
	4 May 81	10"	K	11.6 ± 3.0	128 ± 7.1	9.7 ± 0.2
1253-055 (3C 279)	6 Apr 81	10"	K	16.2 ± 4.1	67.8 ± 6.9	5.1 ± 0.2
	8 Apr 81	10"	K	9.7 ± 6.2	98.4 ± 17	6.2 ± 0.2
1308+326 (B2)	3 May 81	10"	K	11.4 ± 3.1	85.8 ± 7.4	9.0 ± 0.9
	4 May 81	10"	K	20.1 ± 3.3	99.5 ± 4.5	8.1 ± 0.2
1418+546 (OQ 530)	6 Apr 81	10"	K	17.9 ± 7.8	59.3 ± 12	6.5 ± 0.4
	2 May 81	10"	K	10.2 ± 1.1	151 ± 2.9	16.4 ± 0.5
	4 May 81	10"	K	6.6 ± 4.9	23 ± 20	14.2 ± 0.9
1514-241 (AP Lib)	4 May 81	10"	K	2.8 ± 2.3	179 ± 22	25.5 ± 0.5
	4 May 81	10"	H	6.0 ± 2.1	132 ± 10	22.0 ± 0.4
1641+399 (3C 345)	5 Apr 81	10"	K	6.4 ± 1.4	120 ± 6.0	18.3 ± 0.6
	6 Apr 81	10"	K	11.3 ± 1.5	108 ± 3.6	19.3 ± 0.6
	7 Apr 81	10"	K	8.1 ± 0.9	118 ± 3.0	18.6 ± 0.6
	8 Apr 81	10"	K	6.8 ± 1.5	101 ± 6.0	20.8 ± 0.6
			H	11.7 ± 0.9	96 ± 2.1	12.7 ± 0.3
	30 Apr 81	10"	K	16.2 ± 1.2	148 ± 2.0	15.2 ± 0.5
			H	14.7 ± 3.6	152 ± 4.8	9.3 ± 0.5
	1 May 81	10"	K	11.6 ± 2.3	86 ± 3.8	14.9 ± 0.4
			H	11.4 ± 3.4	95 ± 4.8	8.8 ± 0.4
	2 May 81	10"	K	8.8 ± 1.9	99 ± 5.9	13.6 ± 0.4
	3 May 81	10"	K	7.3 ± 3.8	91 ± 10	10.6 ± 1.6
	4 May 81	10"	K	3.3 ± 1.7	84 ± 14	9.1 ± 0.5
2200+420 (BL Lac)	1 May 81	10"	K	10.6 ± 0.8	3.3 ± 2.1	51.3 ± 1.5
			H	15.1 ± 3.5	2.1 ± 6.3	26.4 ± 1.3
	2 May 81	10"	K	14.2 ± 0.8	9.7 ± 1.5	57.2 ± 1.7
			H	14.9 ± 2.2	4.8 ± 2.7	31.2 ± 1.6
			J	11.3 ± 1.6	4.4 ± 2.9	22.8 ± 0.5
	3 May 81	10"	K	4.2 ± 1.1	6.5 ± 7.2	103.2 ± 15.4
	4 May 81	10"	K	9.8 ± 1.1	20.6 ± 3.1	38.9 ± 1.9
			H	8.1 ± 0.9	21.9 ± 3.0	28.5 ± 1.5

Table 1 – *continued*

(1) Object	(2) Date	(3) Aperture	(7a) $\{S \pm \sigma(S)\}_K$	(7b) $\{S \pm \sigma(S)\}_H$	(7c) $\{S \pm \sigma(S)\}_J$
0235+164 (AO)	10 Jul 80	15"	9.2 ± 0.1	–	–
	9 Aug 80	19"	11.8 ± 0.3	6.3 ± 0.5	–
	10 Aug 80	15"	12.6 ± 0.4	7.1 ± 0.6	4.6 ± 0.4
	11 Aug 80	15"	–	6.4 ± 0.4	4.2 ± 0.3
0754+101 (OI 090.4)	7 Apr 81	10"	–	10.5 ± 0.3	5.7 ± 0.2
	8 Apr 81	10"	–	11.0 ± 0.2	8.5 ± 0.2
0818-131 (OJ -131)	8 Apr 81	10"	–	9.3 ± 0.2	7.2 ± 0.1
0851+202 (OJ 287)	12 Mar 81	7.5"	22.9 ± 1.1	15.1 ± 0.8	11.0 ± 0.6
	13 Mar 81	7.5"	33.7 ± 1.7	24.8 ± 1.2	17.6 ± 0.9
	14 Mar 81	7.5"	28.1 ± 1.4	20.5 ± 1.0	14.0 ± 0.7
	8 Apr 81	10"	–	12.3 ± 0.2	9.7 ± 0.3
0912+297 (OK) 222)	7 Apr 81	10"	6.3 ± 0.3	4.3 ± 0.3	3.8 ± 0.2
	8 Apr 81	10"	–	4.9 ± 0.3	3.6 ± 0.1
1147+245 (OM 280)	20 Apr 80	15"	–	6.1 ± 0.1	–
	21 Apr 80	10"	9.1 ± 0.2	5.7 ± 0.2	–
1156+295 (4C 2945)	8 Apr 81	10"	–	–	19.7 ± 0.6
	30 Apr 81	10"	–	5.6 ± 0.2	3.8 ± 0.2
	2 May 81	10"	–	7.9 ± 0.2	5.7 ± 0.2
	3 May 81	10"	–	9.9 ± 0.3	6.6 ± 0.2
	4 May 81	10"	–	6.1 ± 0.2	4.2 ± 0.2
1253-055 (3C 279)	7 Apr 81	10"	–	3.4 ± 0.2	–
	8 Apr 81	10"	–	3.7 ± 0.2	–
	3 May 81	10"	2.0 ± 0.4	1.7 ± 0.3	0.8 ± 0.2
	4 May 81	10"	2.6 ± 0.1	1.9 ± 0.1	1.2 ± 0.2
1308+326 (B2)	20 Apr 80	15"	–	2.6 ± 0.2	2.0 ± 0.2
	6 Apr 81	10"	1.1 ± 0.1	0.7 ± 0.1	–
	8 Apr 81	10"	1.0 ± 0.1	–	–
	3 May 81	10"	–	6.6 ± 0.7	5.3 ± 0.5
	4 May 81	10"	–	5.7 ± 0.1	4.8 ± 0.4
1335-127 (PKS)	3 May 80	8"	17.6 ± 0.5	–	6.7 ± 0.2
	7 May 81	10"	4.3 ± 0.2	2.3 ± 0.2	–
1400+162 (OQ 100)	7 May 81	10"	3.6 ± 0.2	2.5 ± 0.2	–
1418+546 (OQ 530)	20 Apr 80	15"	–	–	6.2 ± 0.3
	21 Apr 80	10"	12.9 ± 0.2	10.1 ± 0.3	–
	2 May 81	10"	–	8.0 ± 0.2	5.0 ± 0.1
	4 May 81	10"	–	10.7 ± 0.2	7.7 ± 0.2
1514-241 (AP Lib)	18 Apr 81	10"	–	–	15.5 ± 0.7
	4 May 81	10"	–	–	16.3 ± 0.3
1538+149 (4C 14.60)	2 May 81	10"	1.7 ± 0.1	1.3 ± 0.1	1.0 ± 0.1

Table 1 – *continued*

(1) Object	(2) Date	(3) Aperture	(7a) $\{S \pm \sigma(S)\}_K$	(7b) $\{S \pm \sigma(S)\}_H$	(7c) $\{S \pm \sigma(S)\}_J$
1641+399 (3C 345)	7 Apr 81	10"	–	11.8 ± 0.2	8.1 ± 0.2
	8 Apr 81	10"	–	–	8.7 ± 0.2
	2 May 81	10"	–	8.0 ± 0.2	5.0 ± 0.1
	4 May 81	10"	–	5.3 ± 0.2	3.6 ± 0.1
1727+503 (I Zw 187)	8 Apr 81	10"	4.6 ± 0.2	3.3 ± 0.2	–
1749+096 (OT 081)	30 Apr 81	10"	4.3 ± 0.1	2.9 ± 0.1	1.4 ± 0.1
1921-293 (OV -236)	20 Apr 80	15"	–	3.8 ± 0.1	–
	21 Apr 80	10"	5.5 ± 0.1	–	–
	9 Jul 80	15"	4.6 ± 0.1	3.9 ± 0.2	–
	30 Apr 81	10"	5.6 ± 0.2	3.8 ± 0.1	2.6 ± 0.2
	1 May 81	10"	6.9 ± 0.1	4.0 ± 0.2	2.6 ± 0.2
	2 May 81	10"	6.5 ± 0.1	4.2 ± 0.2	2.9 ± 0.1
	4 May 81	10"	7.5 ± 0.3	5.0 ± 0.3	3.2 ± 0.2
2155-304 (PKS)	9 Aug 80	15"	34.8 ± 1.0	29.0 ± 1.8	28.5 ± 2.2
	10 Aug 80	15"	37.9 ± 1.1	31.7 ± 2.0	–
	11 Aug 80	15"	40.0 ± 1.0	34.2 ± 2.6	28.5 ± 1.8

vation and columns (3) and (4) contain the aperture size and waveband of the measurement. The degree of polarization, position angle and flux in millijansky (1mJy = 10^{-29} Wm^{-2} Hz^{-1}) are contained in columns (5), (6) and (7) along with their associated rms errors. The final part of Table 1 contains photometry alone, with flux measurements at *K*, *H* and *J* in columns (7a, b and c).

3.1 0818 – 128

OJ-131 is the Ohio Radio source with a flat spectrum which was identified with a blue polarized object by Tapia *et al.* (1977). Tapia *et al.* list radio fluxes from 408 MHz to 10.8 GHz, and they also mention the lack of emission lines visible in low-dispersion spectra. The linear polarization in the optical region was very high and variable (9–36 per cent). OJ-131 has a range of variability of 2.9 mag (Craine, Duerr & Tapia 1978), and at peak flux is one of the optically brightest BL Lac objects known. Optical polarimetry by Angel *et al.* (1978) over a two month period gave a range of 12–25 per cent at polarization position angles of 60–95°. Infrared monitoring shows it to have moderate infrared polarization (9–14 per cent) and a shallower infrared continuum than is normal for a BL Lac object ($\alpha \approx 0.7$, where $S(\nu) \propto \nu^{-\alpha}$). There were flux variations of up to 15 per cent on successive nights.

3.2 1156 + 295

4C 29.45 is a variable compact source with a flat spectrum from 178 MHz–5 GHz (Conway *et al.* 1974). It is identified with a strong-lined quasar at $z = 0.728$ (Schmidt 1974), which has strong optical polarization of up to 10 per cent with no preferred position angle. The optical brightness varies by at least two magnitudes (Moore & Stockman 1981). 4C 29.45 was the subject of widespread interest during early 1981 because it was undergoing a dramatic flare at optical/infrared wavelengths.

The UKIRT monitoring covered a 4 day run in 1981 April and a 5 day run in May 1981. At the beginning of April the source was at the peak of its flare, with an enormous 2.2 μm flux of 75 mJy ($K = 9.87$) and a strong infrared excess ($J-K = 1.61$). When the non-thermal component was at its strongest, there was no evidence for wavelength dependent polarization between J, H, and K bands. Thereafter, the 2.2 μm flux fell by a factor of 2.3 in only three days. The infrared polarization followed the flux, rising to 14 per cent (K) at the time of peak flux and then falling to 5 per cent two days later. During these same four days the position angle rotated by approximately 50°. By the end of the month, the flare had subsided. On April 30, the 2.2 μm flux was down by a factor of 8.6 from its peak of three weeks earlier. The polarization was still highly variable from night to night, and the position angle swung by large amounts (up to 50°) from night to night. Unfortunately in the faint phase there was not sufficient signal-to-noise to study wavelength-dependent polarization. The polarization behaviour of 4C 29.45 was even more erratic in its low-flux state than in its high-flux state. These infrared measurements were part of a multi-wavelength collaboration on this object (Glassgold *et al.* 1983).

3.3 1253 – 055

This source is one of the optically violent variable (OVV) quasars, with strong emission lines and a history of large-amplitude optical variability. 3C 279 is in fact the most variable extra-galactic source known, varying by over 6.7 mag (Eachus & Liller 1975). At peak brightness it is one of the most luminous objects in the universe ($z = 0.536$). 3C 279 is a strong X-ray emitter (Tananbaum *et al.* 1980), and shows a steep power-law continuum of slope $\alpha = 1.3$ in the range 1–10 μm (Rieke *et al.* 1977). The radio properties of 3C 279 are intriguing. It has a compact variable source which has been extensively studied by *VLBI* techniques (Cotton *et al.* 1979; Kellerman *et al.* 1971). The fringe visibilities change with time, which implies a superluminal velocity of $\sim 21c$ for the compact components (Seielstad *et al.* 1979; Cotton *et al.* 1979). Optical polarimetry of 3C 279 has been published by Visvanathan (1973). Kinman (1967) and Elvius (1968), and the range is 4–19 per cent with large variations in position angle. These are the first published infrared measurements, and they show high polarization at 2.2 μm and variability by a factor of 3 over a one-month period. The infrared spectral index ($\alpha = 1.4$–1.7) agrees with previous optical and infrared measurements.

3.4 1335 – 127

This is an object from the Parkes 2.7 GHz flat spectrum sample (Condon, Hicks & Jauncey 1977) which was shown to be a very red source ($\alpha \simeq 3.0$) by Impey & Brand (1981). Aller, Aller & Hodge (1981) have monitored the radio polarization at 14.5, 8.0 and 4.8 GHz and find variability in both the flux and polarization at these frequencies. There is evidence for differential rotation between the position angles at high and low frequencies. Two flux measurements at 2.2 μm, separated by almost a year, show large amplitude variability by over a factor of 4. The source has so far been too faint to attempt a polarization measurement.

3.5 1538 + 149

4C 14.60 has a large variability amplitude of > 2.8 magnitudes (Usher 1975) and a steep and variable spectral index in the optical region (Tapia *et al.* 1977). Kinman (1976) has made

a single optical polarization measurement of 22 per cent and Wills *et al.* (1980) have measured values in the range 17–20 per cent. Infrared photometry at J, K and K shows a spectral index $\alpha \simeq 0.9$, which is considerably flatter than the optical slope. The difference is more than can be accounted for by a break in the spectrum due to synchrotron losses. 4C 14.60 has no published reshift.

3.6 1749 + 096

OT 081 has a range of optical polarization of 3–9 per cent (Kinman 1976) and it has a steep optical continuum of spectral index $\alpha = 2.2$ (Tapia *et al.* 1977). The source has no emission line redshift. Wills *et al.* (1980) have made optical polarization measurements in the range 4–12 per cent. The spectral index from the infrared measurements ($\alpha \simeq 2.0$) is in close agreement with the optical data. If the source has no large amplitude variability, then the inferred $V-K$ value is 4.0, which is typical for BL Lac objects.

3.7 2200 + 420

The prototype object BL Lacertae was thought for many years to be a variable star, until it was identified with the radio source VR0 42.22.01 (Schmitt 1968; MacLeod & Andrew 1968). BL Lac has been extensively studied and its properties have been reviewed by Stein, O'Dell & Strittmatter (1976) and Miller (1978). The non-thermal source has now been demonstrated convincingly to lie at the centre of a giant elliptical galaxy at $z = 0.07$ (Miller & Hawley 1977; Miller *et al.* 1978). The optical polarization of BL Lac varies from 2–23 per cent and over all position angles (see e.g. Angel *et al.* 1978). In addition, simultaneous optical/infrared polarimetry has been published by Knacke, Capps & Johns (1976) and Puschell & Stein (1980), and near-simultaneous polarimetry from centimeter to visual wavelengths has been obtained by Rudnick *et al.* (1978). However, the most impressive results have been achieved by a world-wide collaboration which monitored BL Lac almost continuously for a week in 1978 (Moore *et al.* 1982). This group found the Stokes parameters Q and U to vary quasi-periodically on a time-scale of ~ 1 day, with a superimposed longer time-scale variation involving a rotation of the position angle. Where infrared points were available they tracked the optical polarization and position angle.

In four nights of observation in 1981 May, BL Lac showed erratic polarization behaviour with variations of up to 10 per cent between successive nights. On the other hand, the position angle was stable $\pm 5^\circ$. There was no evidence for wavelength dependent polarization during that period. The flux variations were very large, and the 2.2 μm flux reached a peak of 103 mJy ($K = 9.5$), after which it fell by a factor of 2.6 in *only* 24 hours. This is one of the largest brightness changes ever recorded in a BL Lac object, and is rivalled only by an optical change of a factor of 2.5 in 24 hours reported in BL Lac by Veron (1978). It is clear that these dramatic variations are not sampled from the same sort of light curves that were observed in 1979 September.

3.8 OTHER OBJECTS

(1) 0735 + 178 was well-observed during 1981 April/May, during which time the object went through a broad flux and polarization flare. The 2.2 μm flux and polarization went from ~ 18 mJy and ~ 12 per cent in early April to ~ 28 mJy and ~ 25 per cent in early May. Between the nights of April 6 and 7 there was a 25° position angle rotation, seen at two different wavelengths. (2) Over a five night period, the 2.2 μm polarization of 0754 +

101 fell from 19 per cent to 3 per cent only a 10 per cent change in the total flux. The position angle was roughly constant. (3) 0912 + 297 was shown to have high infrared polarization for the first time, with moderate (~ 15 per cent) variability of its infrared flux. (4) 1308 + 326 again demonstrated rapid polarization changes (~ 10 per cent in 24 hours), but at a completely different position angle from previous measurements. (5) There were also large position angle rotations in 1418 + 546, when the position angle changed by ~ 50° in two days. (6) 16 + 1 + 399 was interesting because over a five day period there was a smooth monotonic decrease in both degree of polarization and infrared flux.

4 Discussion

Together with Paper I, this paper brings the number of BL Lac objects we have measured up to 25, 19 of which have been measured polarimetrically in the infrared, 10 for the first time. The principal wavelength of the study was 2.2 μm, but 13 objects have polarimetry at shorter wavelengths also. 15 of the BL Lac objects in the sample have redshifts, so useful (but model-dependent) physical information can be derived for them. Although BL Lac objects generally have featureless continua, spectra with good signal to noise ratio have reavealed weak emission and absorption lines in many objects. For the purpose of this work OVV quasars are included in the sample since they share the properties of the BL Lac compact source: variability, polarization and a steep infrared continuum.

The sample includes nearly half of the objects in the compilation by Angel & Stockman (1980), and it is probably large enough for the data to be considered representative of the infrared properties of BL Lac objects. However, there are clearly selection effects operating. Since techniques involving slitless spectroscopy and searches for UV excess are not appropriate for BL Lacs, most have been found from accurate positions of flat spectrum radio sources. There is currently *no* known example of a radio-quiet BL Lac object, and it is possible that this absence is not simply a selection effect. For example, BL Lac objects are intrinsically strong radio emitters compared to quasars, and most BL Lacs selected by non-radio properties have turned out to be radio-emitters. The question may be answered by studies currently under way which select objects by their optical polarization variability and X-ray emission (Borra, 1983; Hawkins 1983; Impey & Brand 1982; Stocke *et al.* 1982; Chanan *et al.* 1982). A more serious problem for this study is the fact that radio source identification procedures select against strong infrared-emitters (Rieke & Lebofsky 1979). For example, if a sample of flat spectrum sources is observed in the infrared without regard to optical identification, there is a smooth distribution of near-infrared spectral indices from $\alpha = 0$ to $\alpha = 3$ (where $S(\nu) \propto \nu^{-\alpha}$). Some of the reddest objects are BL Lac objects (Impey & Brand 1981). However, the spectral index distribution of the sample of BL Lac objects peaks around $\alpha = 1.5$–1.7 and drops off for sources with spectra steeper than $\alpha \simeq 2$ (next section), so the very red objects are clearly being missed.

4.1 SPECTRAL ENERGY DISTRIBUTION

The near-infrared continuum in BL Lac objects is presumed to be due to synchrotron radiation from relativistic electrons. The two most compelling reasons for this assumption are the high polarization in BL Lac objects, and the continuity of the infrared flux with the high frequency radio spectrum, where the synchrotron interpretation is secure. Thus, the infrared flux is interpreted as the high energy tail of the compact radio component. However, this latter interpretation is not at all secure (BL Lac: Rieke & Kinman 1974, 1413 + 135; Beichman *et al.* 1981, 0235 + 164; Rieke *et al.* 1976).

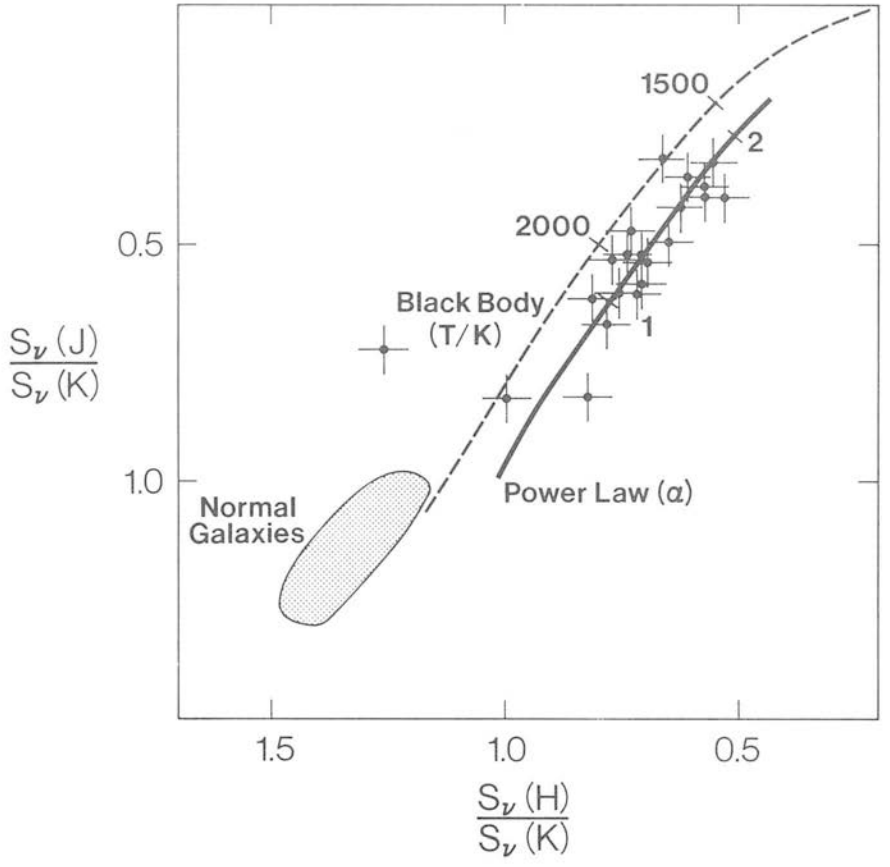


Figure 1. Infrared flux ratios are plotted for 21 BL Lac objects based on photometry at *J*, *H* and *K*. The region of colours for normal (elliptical) galaxies is shown as a shaded area, and the two loci of blackbody and power-law spectra are marked in units of temperature and spectral index respectively.

A major aim of this work is to understand the polarization and flux properties of the central regions of BL Lac objects. Therefore, we must be sure that the infrared measurements are isolating the non-thermal component. A relatively model-free way of doing this is to use a two-colour plot. It is noted that flux ratios and two-colour diagrams are very appropriate for studying composite systems (Becker 1937; Wade 1980).

In Fig. 1, infrared flux ratios are plotted for 21 BL Lac objects. For objects where more than one set of measurements was made, we plotted the data set which is nearest the object's mean spectral index. Two lines are superimposed, one of which is the locus of blackbody colours (with temperature marked), and the other of which is the locus of power-law spectra (with spectral index marked). The area corresponding to the colours of normal elliptical galaxies is defined from the work of Frogel *et al.* (1978). Most of the sample lie close to the power-law line and a power law is a very good representation of the near-infrared spectral shapes. In fact, the scatter about the line is consistent with the photometric errors, and there is no need to invoke any other component. All but three objects fall in the range corresponding to $0.8 < \alpha < 1.8$. The three BL Lacs lying outside this range (Mk 501, 1727 + 503 and 2155 − 304) all have nonstellar images on deep, direct plates. Therefore, their colours are probably affected by the presence of an underlying galaxy component. Even three of the objects without three-colour photometry can be included in this argument, since their two-

colour ratios of $S(H)/S(K) = 0.53, 0.63, 0.68$ put them far into the non-thermal region of Fig. 1.

4.2 LUMINOSITY PARAMETERS

The calculation of near-infrared luminosity assumes (1) the cosmological nature of the redshift and (2) isotropic emission. Fifteen of the sample have published redshifts. The luminosity density in the rest frame of the BL Lac object is calculated from

$$L(\nu) = 4\pi\left(\frac{cz}{H_0}\right)^2 (1+z)^{-1}\, F[\nu/(1+z)]$$

assuming a Friedmann model with cosmological constant $\Lambda = 0$ and taking $H_0 = 75$ km s^{-1} Mpc^{-1} and $q_0 = 1$. As in Paper I, observed 2.2 μm fluxes are converted into BL Lac rest frame fluxes by $F[\nu/(1+z)] = F(\nu)\,(1+z)^{\alpha}$ where α is the near-infrared spectral index from 1–2 μm. Table 2 shows the object name (1), redshift (2), luminosity distance (3), the maximum 1–2 μm luminosity observed (4), the number of observations of each object (5). The infrared luminosities are lower limits in almost all cases since only limited infrared bandwidths have been used. For a bandwidth of 1–10 μm and a typical spectral index of $\alpha > 1.4$, the integrated IR luminosity is increased by an order of magnitude, assuming no synchrotron turnover out to 10 μm. The range of luminosity within the sample is very large: a factor of 1.2×10^3.

Generating these prodigious luminosities is a problem for almost any power source. One of the reasons that the relativistic jet models are so popular is that they ease the lumin-

Table 2. Luminosities and their variations for 15 BL Lac objects.

(1) Object	(2) z	(3) d(Mpc)	(4) L^{IR}_{MAX} (erg/sec)	(5) L_{MAX}/L_{MIN}	(6) No. Obsn.
0235+164	0.852	3410	4.8×10^{46}	5.0	12
0735+178	0.424	1700	2.3×10^{46}	3.4	18
0851+202	0.306	1220	1.4×10^{46}	2.5	11
1156+295	0.729	2920	9.6×10^{46}	6.5	9
1253-055	0.538	2150	7.5×10^{45}	3.2	6
1308+326	0.996	3980	3.2×10^{46}	11.7	10
1400+162	0.244	970	1.2×10^{45}	2.4	3
1514-241	0.049	200	4.8×10^{44}	1.1	2
1641+399	0.595	2380	2.8×10^{46}	2.9	11
1652+398	0.034	140	3.1×10^{44}	1.2	2
1727+503	0.055	220	7.9×10^{43}	1.6	7
1921-293	0.353	1410	$4.5\ c\ 10^{45}$	1.6	7
2155-304	0.170	680	6.9×10^{45}	3.2	8
2200+420	0.069	280	4.2×10^{45}	3.1	4
2223-052	1.404	5620	5.1×10^{46}	2.0	5

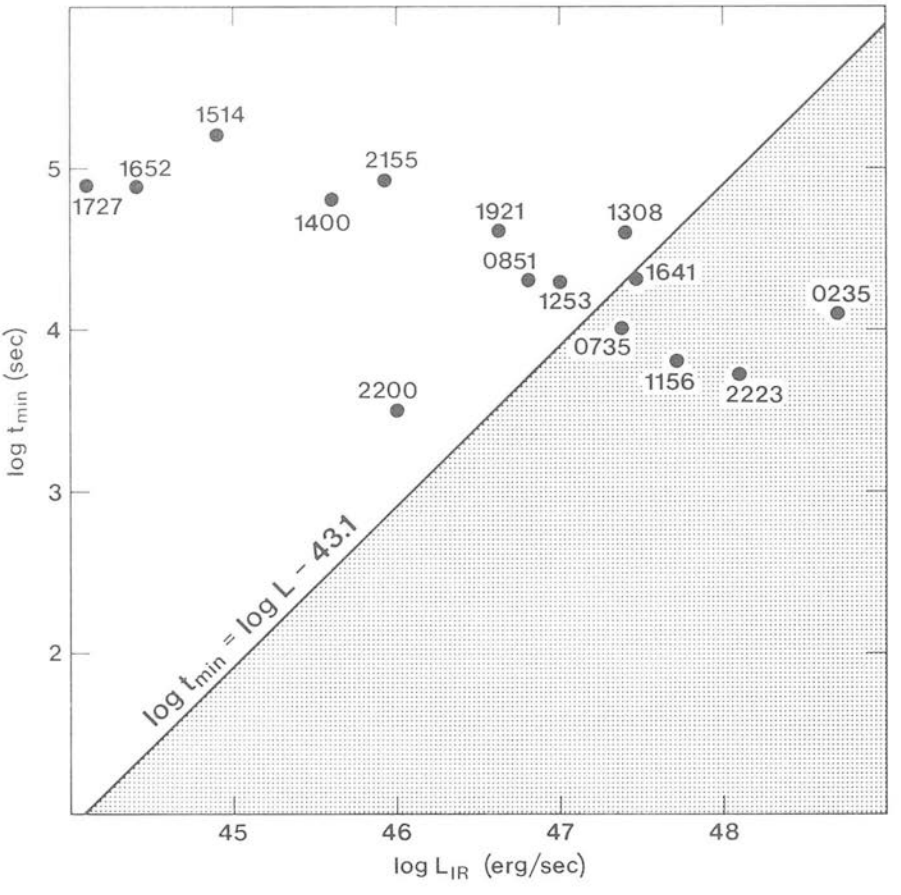


Figure 2. The maximum infrared luminosity is plotted against the minimum variability time-scale in the source rest frame. The shaded area corresponds to the forbidden zone of variability for a spherically accreting supermassive object (from Elliot & Shapiro 1974). A bandwidth of 1–10 μm is used for the infrared luminosity. Values of t_{MIN} are normalized to variations of 10 per cent in intensity. Objects are labelled by the first four digits of their Parkes-type position designators.

osity requirements in compact sources. The size of the power source is inferred from the minimum variability time-scale in the BL Lac rest frame $t_{VAR}/(1+z)$, where $t_{VAR} = L/(dL/dt)$. A plot of $t_{VAR,min}$ versus the 1–10 μm luminosity (data combined with paper I) now has *five* objects which violate the spherical accretion limit providing support for non-spherical accretion and jet emission (Fig. 2). The objects in question are BL Lac, A0 0235 + 164, B2 1308 + 326, 3C 446 and 1641 + 399. Note that in Figure 2, the values of t_{MIN} are normalized to variations of 10 per cent in intensity, so that the different rates and amplitudes of variability in BL Lac objects are dealt with uniformly.

We can investigate the dependence of source polarization and variability properties on luminosity. The sample is quite small, because it includes only those sources with redshifts and must exclude Mk 501, 1727 + 503 and 2155 – 304 in order to rule out objects with thermal dilution. All the measured quantities (flux, polarization, position angle, spectral index) vary by large amounts for a *given* object, and this can swamp the trends between objects unless the parameters are carefully chosen. For example, maximum and mean luminosities may probe different physical regions (if relativistic jet models apply, the

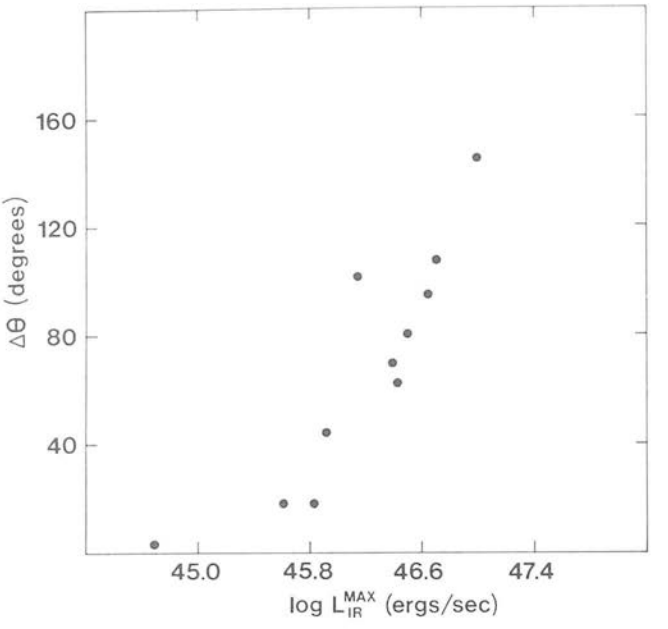


Figure 3. Range of position angles from the accumulated monitoring plotted against the peak infrared luminosity.

mean luminosity might be emission from the bulk of the jet, while the peak luminosity is emission from knots or condensations within the jet). There are no significant correlations between the maximum 1–2 μm luminosity and either the (1) maximum degree of polarization, (2) the range of polarization, (3) the variability index (defined by L_{MAX}/L_{MIN}), (4) the maximum 1–2 μm spectral index or (5) the mean 1–2 μm spectral index. (There is a possible correlation between mean spectral index and mean luminosity, but only at the 1.8 σ level in a sample of 12 objects.). However, somewhat surprisingly, the range of position angles for an object does correlate with the luminosity. Fig. 3 shows the plot, and the Spearman non-parametric estimator gives a rank correlation of 0.86, which means that the parameters are related at the 99 per cent confidence level. Removing any subset of BL Lacs does not destroy the correlation; nor does plotting mean luminosity instead of peak luminosity. Only part of the correlation is due to the selection effect of more frequently observed objects showing larger values of position angle rotation. The relative error in the 1–10 μm luminosity of Fig. 2 can be shown to be little larger than the relative error in the photometry, given in Table 1. This assumes no dramatic breaks in the spectral shapes of these objects.

This relationship is important because it connects the luminosity of a BL Lac object with a luminosity-independent parameter. The range of position angle is a purely geometrical quantity. For synchrotron radiation in the infrared, the source is optically thin and plasma propagation and Faraday effects can be ignored. In this regime, the position angle of polarization reflects the component of the magnetic field perpendicular to the line-of-sight. There are two ways to interpret large position angle variations. Either there is a large scale rotation of the field embedded in the emitting volume; or the field makes a small angle to the line of sight, so that the projection of small changes in a field direction results in large rotations of position angle.

4.3 RELATIONSHIP BETWEEN POLARIZATION AND SPECTRAL INDEX

A general link between high degree of polarization, large amplitude variability and steep infrared spectrum was discussed in Paper I. There is a clear separation between the properties of BL Lac objects and radio-loud quasars, since most radio-loud quasars do not have high polarization (Stockman & Angel 1978), large amplitudes of variability (Bonoli *et al.* 1979), or steep infrared spectra (Neugebauer *et al.* 1979). As part of a unified picture of active nuclei, Blandford & Königl (1979) have proposed a model where much of the non-thermal radiation comes from a relativistic jet. They envisage a sequence of increasing activity (radio-quiet quasar, radio-loud quasar, OVV quasar, BL Lac object) corresponding to smaller and smaller angles of the jet axis wrt. the line-of-sight. In Paper I, a good correlation was

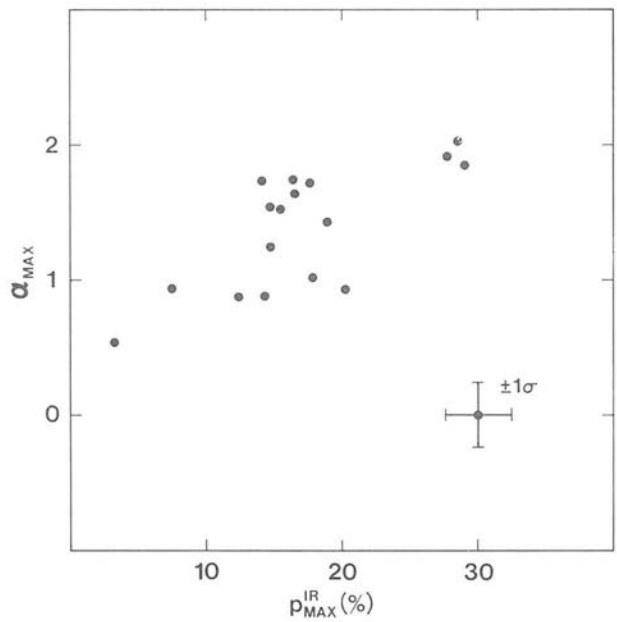


Figure 4. Infrared spectral index plotted against maximum 2.2 μm polarization.

found between maximum infrared polarization and maximum 2.2 μm spectral index for 14 BL Lac objects. With more data on each object, and a slightly increased sample of 17 (it also becomes possible to standardize on K polarization), the correlation is still evident (see Fig. 4). The expanded data set gives a Spearman correlation of 0.68, and so the variables are related at the 95 per cent confidence level. The data are listed in Table 3, which contains the object name (1), the maximum 2.2 μm polarization (2), the maximum optical polarization (3), the optical variability amplitude (4), and the mean 1–2 μm spectral index (5). The numbers in column 4 are taken from Angel & Stockman (1980).

The quantities p^{IR}_{MAX} and α_{MAX} may be indirectly related, through a third parameter (N), which is the number of observations. For any observable with an arbitrary distribution of values about the mean, the maximum measured value increases as the number of measurements increases (in the case of a Gaussian parent distribution, it increases as $\sqrt{N}$). Plots of N vs. p^{IR}_{MAX} and N vs. α_{MAX} do in fact show a very slight correlation, but the correlation between p^{IR}_{MAX} and α_{MAX} is too strong to be the result of a selection effect (the correlation still has a similar scatter if $\overline{p_{IR}}$ and $\bar{\alpha}$ are used instead of peak values).

A relationship between p^{IR}_{MAX} and α_{MAX} with such a steep slope is intriguing because it connects two parameters closely related to the geometry of the source. We note that the

Table 3. Maximum polarization, variability, spectral index.

(1)	(2)	(3)	(4)	(5)
Object	p_{MAX}^{IR}	p_{MAX}^{OPT}	ΔB	$\bar{\alpha}_{IR}$
0235+164	36.2	44	5.2	2.1
0306+103	3.6	–	2.2	–
0735+178	32.6	31	2.5	1.2
0754+101	18.9	26	1.0	0.95
0818-131	21.4	36	2.9	0.78
0851+202	19.7	32	4.0	1.0
0912+297	12.2	13	1.9	0.99
1147+245	14.8	13	1.4	1.4
1156+295	15.3	9	2.6	1.2
1253-055	17.6	19	6.7	1.5
1308+326	20.1	25	5.6	0.92
1335-127	–	–	–	1.7
1400+162	–	14	–	1.2
1418+546	17.9	19	4.8	1.0
1514-241	7.4	7	2.5	0.86
1538+149	–	22	2.8	0.92
1641+399	16.2	16	2.0	1.6
1652+398	–	4	0.5	0.36
1727+503	–	6	1.9	1.0
1749+096	–	9	–	2.0
1921-293	13.9	8	2.5	1.5
2155-304	3.0	7	–	0.35
2200+420	15.1	23	4.0	1.6
2223-052	16.3	17	3.4	1.5
2254+074	17.4	21	1.6	–

slope is too steep to be due simply to the variation of p with α in simple partially aligned field models.

p_{MAX}^{IR} is a measure of the maximum degree of alignment of the magnetic field in the synchrotron-emitting region ($p_{MAX} \propto (B_{\parallel}^2 - B_{\perp}^2)/\underline{B}^2$, where $B_{\parallel}$ and $B_{\perp}$ are the field components resolved parallel and perpendicular to the line-of-sight). For a single synchroton component with a uniform magnetic field, $p_{MAX} = (\alpha + 1)/(\alpha + 5/3)$. In some objects like 0235 + 164 and 0735 + 178 the values of p_{MAX} are a large fraction of the theoretical maximum value, implying that the integrated field over the whole emitting volume is nearly uniform. The steepest spectral index α_{MAX} indicates strong synchrotron losses, but could

also relate to the geometry of the source. Marscher (1980) has shown that the spectrum of an unresolved relativistic jet steepens as its angle to the line-of-sight decreases. One way of interpreting Fig. 4 is as a sequence of BL Lac objects with jets, where the steep spectrum, highly polarized sources have jets nearly pointing at the observer.

4.4 CHANGES IN POLARIZATION, VARIABILITY AND *IR*-EXCESS

There are 85 data points covering night-to-night variations ($\Delta t = 20-28$ hr) in the flux at 2.2 μm and changes in the 1–2 μm spectral index. Some of the transitions are highly significant, and they can all be dealt with as an ensemble because there are no selection effects biasing the distributions $n(\Delta S_K)$ and $n(\Delta\alpha)$. No systematic effects of the kind discussed by O'Dell *et al.* (1978) are found. Although these variations probe the smallest dimensions of the compact source, the synchrotron life-times of the highest energy electrons are short enough to affect the spectrum on a time-scale of a day. The spectral index is therefore probably determined by a balance between acceleration mechanism and synchrotron losses for the relativistic electrons.

In addition, changes in 2.2 μm polarization and changes in the spectral index can be considered, giving 61 data points for night-to-night variations. A relationship has already been demonstrated between the quantities $p^{\rm IR}_{\rm MAX}$ and $\alpha_{\rm MAX}$. However, there is no correlation between changes in polarization and changes in spectral index. This null result is interesting because it indicates that the relationship between $p^{\rm IR}_{\rm MAX}$ and $\alpha_{\rm MAX}$ is not simply a result of short-term energy injection into the source, but is connected with some fundamental (geometric) property of each BL Lac object.

The database of polarimetry and photometry from this paper and from Paper I is sufficient to investigate the amplitude of variability on different time-scales. The data was collected during seven separate runs covering an eighteen month period, and night-to-night changes are particularly well sampled. The evidence indicates that substantial variations on time-scales of less than a day are rare (but see Puschell *et al.* 1979; Stockman & Angel 1978; Snyder *et al.* 1980). After a large polarization monitoring programme, Angel *et al.* (1978) concluded that the spectrum of fluctuations falls off above $\nu \sim (\text{day})^{-1}$. Therefore, the night-to-night observations probe the smallest unit of emission in the active core of the objects. Additional run-to-run measurements have an average spacing of one month, which is a typical time-scale for the larger flares measured in long-term monitoring programmes (Epstein *et al.* 1982).

First, the distribution of polarization changes at $\Delta t \sim 1$ day can be compared with the polarization changes over the longer baseline $\Delta t \sim 1$ month. Although some of the polarization changes ($\Delta p, = p_2 - p_1$) are numerically small, an analysis of the relative error distribution shows that (1) the errors are distributed normally, and (2) the error distributions for night-to-night and run-to-run are identical. We use the Mann-Whitney U-Test, which is a non-parametric test and does not presume a form for the parent distribution. This test is then used to compare the distributions of Δp on day and on month time-scales. The resulting rankings show that the two distributions are identical at the 87 per cent confidence level. Second, the flux changes (ΔI) on day and month scales can be compared. The Mann-Whitney Test reveals that the monthly flux changes are longer than the night-to-night flux changes, at the 99.94 per cent confidence level. The mean level of ΔI for $\Delta t \sim 1$ month is a factor of 2.4 higher than the mean ΔI for $\Delta t \sim 1$ day. The relatively constant value of Δp is not due to p being a bounded parameter [i.e., $0 < p < (\alpha + 1)/(\alpha + 5/3)$ for synchrotron radiation], since the values of Δp cover a small fraction of the allowed range. While the amplitude of intensity variation in BL Lac objects increases from day to month time-scales, the amplitude of polarization variation is constant.

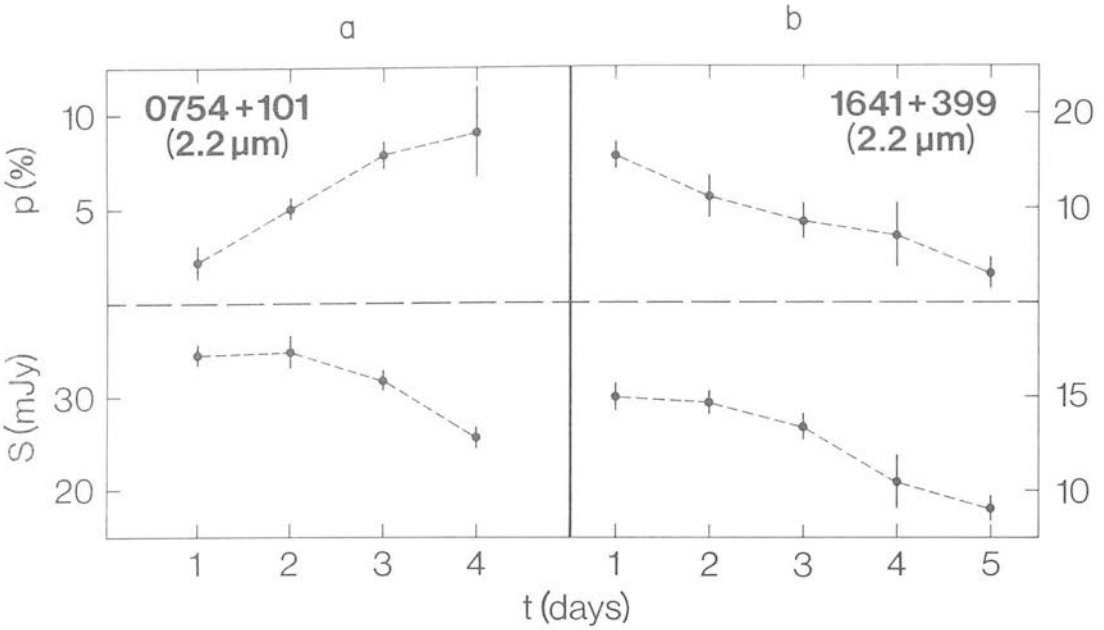


Figure 5. Degree of polarization and total flux at 2.2 μm are plotted for two objects monitored on several consecutive nights (a) 0754 + 101, (b) 1641 + 399.

4.5 CHANGES IN POLARIZED AND TOTAL FLUX

In Paper I it was found that the changes in polarized and unpolarized flux for two BL Lac objects could be understood in terms of a synchrotron component of variable flux and fixed polarization. On a diagram of ΔI_{Tot} vs. ΔI_{Pol}, these objects had a strongly peaked or correlated distribution (where $\Delta I_{\mathrm{TOT}} = I_2 - I_1$ and $\Delta I_{\mathrm{POL}} = p_2 I_2 - p_1 I_1$). Other objects had changes in flux and polarization that were completely uncorrelated. Examples of monotonic behaviour are given in Figs 5a and b, where the flux and polarization are plotted for 0754 + 101 and 1641 + 399 over the period of a monitoring run. The two objects show an opposite sense of polarization and flux changes. This distribution can be reinvestigated now with much improved statistics: 85 data points compared to 41 in Paper I. Fig. 6 shows inter-night changes in the polarized and total flux plotted in absolute units. All objects with measurements on consecutive nights are included in the diagram and measurements in the *J*, *H* and *K* bands are combined.

One problem with interpreting the plot of ΔI_{TOT} vs. ΔI_{POL} is that the numerical values of ΔI_{POL} are low and the relative errors in the values of ΔI_{POL} are much higher than the relative errors in ΔI_{TOT}. Study of the distribution of individual errors shows that the errors in ΔI_{TOT} and ΔI_{POL} both have (a) similar, normal distributions, and (b) similar mean values of 0.3 mJy. A linear regression straight line can be fitted to the data for OJ 287 and 0735 + 178: $\Delta I_{\mathrm{POL}} = 0.16\ \Delta I_{\mathrm{TOT}} - 0.03$ with coefficients of determination $R^2 = 0.57$. The value of

$$\chi^2 = \sum_i [(\Delta I_{\mathrm{POL}})_i - 0.16\ (\Delta I_{\mathrm{TOT}})_i + 0.03]^2 = 24.0$$

for the straight-line fit can be compared with the value of

$$\chi^2 = \sum_i (\Delta I_{\mathrm{POL}})^2 = 55.2$$

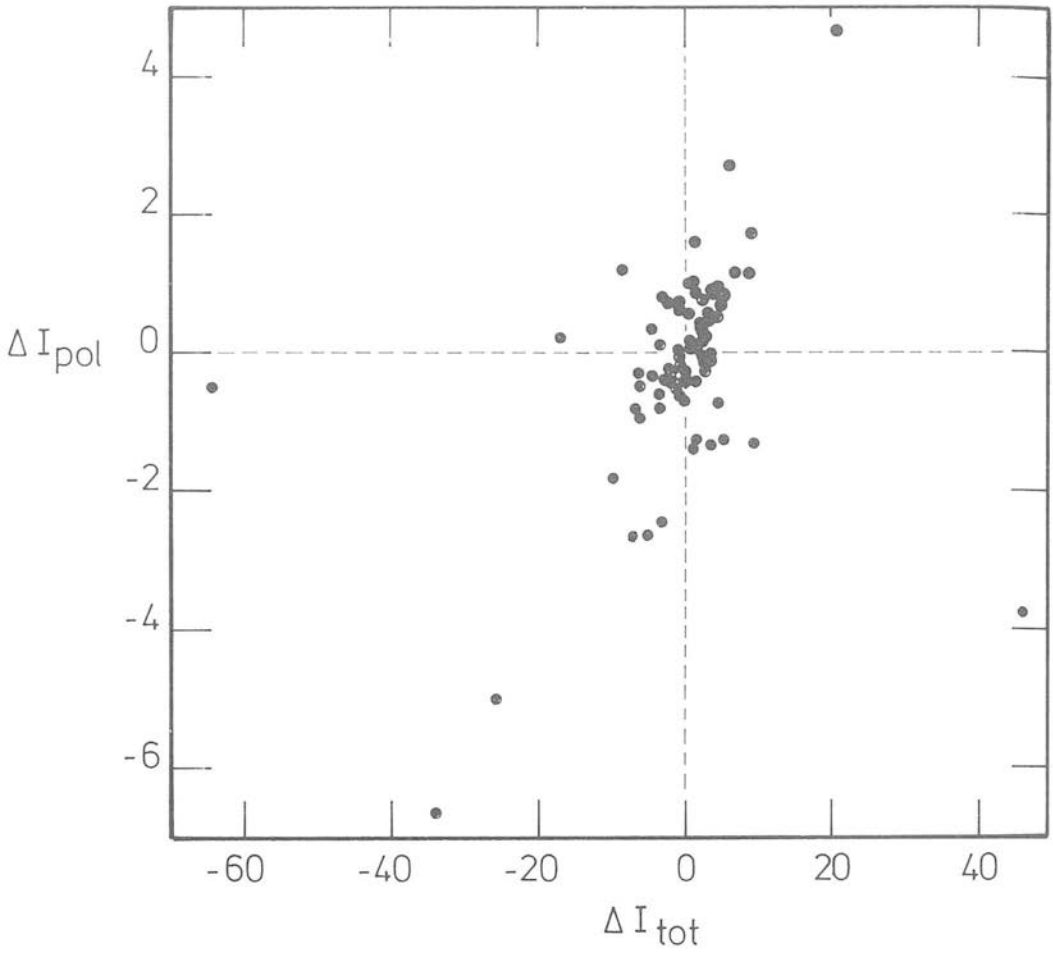


Figure 6. Inter-night changes in polarized flux are plotted against inter-night changes in total flux (both in absolute units). The mean standard errors in each axis are shown.

for the assumption that the values of ΔI_{POL} are scattered about the origin with no correlation (for 50 degrees of freedom). Therefore the data is inconsistent with

$$\frac{1}{n}\sum_{i}^{n} \Delta I_{\mathrm{POL}} = 0$$

or a constant at the 95 per cent confidence level. Applying a non-parametric test gives the same conclusion. The Spearman Rank Correlation coefficient gives a probability of 99.5 per cent that the mean value of ΔI_{POL} is not zero, and depends on ΔI_{TOT}.

What produces the types of variation? Blandford & Rees (1978) have discussed consequences of polarization produced by n randomly oriented subunits (each ~ 70 per cent polarized). The degree of polarization is ~ 50 per cent/$\sqrt{n}$ and as one subunit is added or subtracted, the relative intensity, percentage polarization, and polarization angle change by $1/n$, ~ $50/n$, and $1/\sqrt{n}$ radians respectively. The subunits vary on a time-scale of days, so their dimensions are ~ 10^{15} cm. Moore *et al.* (1982) have suggested such a model for behaviour observed in BL Lac itself during an intensive observing campaign in 1979. One possibility is that the subunits are emission knots in a relativistic jet. High polarized knots

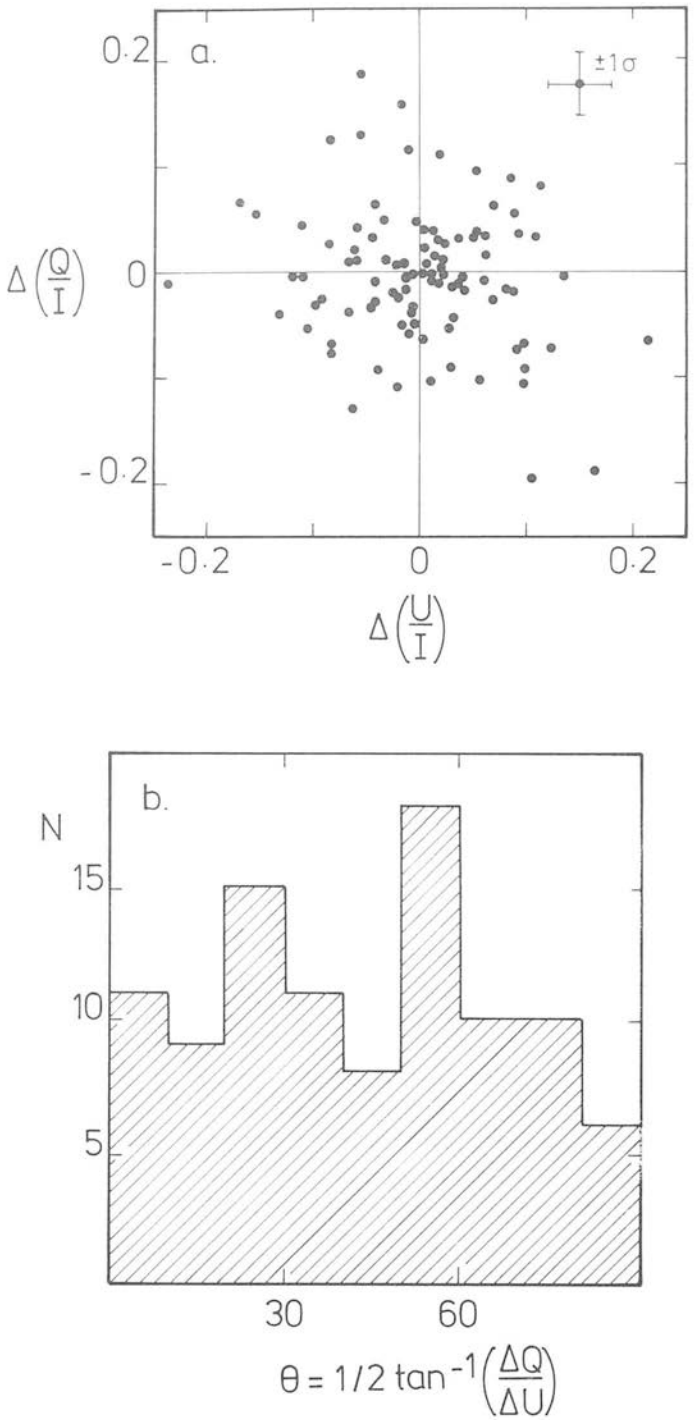


Figure 7. (a) Internight changes in $\Delta(Q/I)$ and $\Delta(U/I)$ are plotted. The uniform distribution of points around the origin indicates that the distribution of p_{Var} is unaffected by instrumental biasses; (b) The distribution of angle $\theta = 1/2 \tan^{-1} (\Delta Q/\Delta U)$ is plotted.

have been resolved optically in the M87 jet (Sulentic, Arp & Lowe 1979) and they have been seen in *VLBI* maps of many radio sources (Miley 1980). The non-variable component might then correspond to the uniform volume emission of the jet, or alternatively, to an accretion disc surrounding the central power source. However, changes in polarization and flux appear totally uncorrelated for four of the objects, and there are many significant individual examples of large changes in polarization with little or no corresponding change in total flux. This has no interpretation in the multiple sub-unit model where the flux and polarization changes should always be correlated. It is also difficult to explain using the incoherent synchrotron model. Propagation and optical depth effects will be small at infrared wavelengths, so polarization should reflect the magnetic field geometry. Changes in Faraday optical depth cannot be responsible because they would be accompanied by large swings in position angle, which are not observed.

4.6 WAVELENGTH-DEPENDENT POLARIZATION

In Paper I, 0735 + 178 and 0235 + 164 were found to have wavelength-dependent polarization, with the polarization increasing towards shorter wavelengths. Further observations of 0735 + 178 have confirmed the existence of wavelength-dependent polarization, with variable slope. Bailey, Hough & Axon (1982) have recently found evidence for wavelength-dependent polarization where the amount of $p(\lambda)$ depends on the degree of polarization; their data having better polarimetric precision but poorer wavelength resolution and photometric precision than the data discussed here. This important result will be a key to the synchrotron modelling of BL Lac objects. The most likely explanation for strong $p(\lambda)$ behaviour involves the superposed spectral variations of physically separate emitting regions (Puschell *et al.* 1983; Impey, Brand & Tapia 1982b), so both polarimetric and photometric observations will be needed in future.

4.7 THE $Q-U$ PLANE

Another way of studying the variable polarization of BL Lac objects is a $Q-U$ diagram in which the normalized Stokes parameters Q/*I* and U/*I* are plotted on the x and y axes. In a $Q-U$ plot, contours of constant polarization are circles and contours of constant position angle are radii. The distance from a data point on the diagram to the origin is the polarization p, and the vector joining the two data points on successive nights represents $p_{\rm Var}$. Position angles on the sky are measured with respect to the Q axis. In Paper I, we observed that the polarization trajectories for monitoring runs on several objects seemed to follow characteristic paths of constant position angle and polarization. With a much improved database we can investigate this behaviour statistically. There are 97 vectors each representing a change in polarization between successive nights. There are no instrumental effects biasing the orientation of the $p_{\rm Var}$ vectors, since the values of $p_{\rm Var}$ are uniformly distributed around the origin of a plot of $\Delta(Q/I)$ vs. $\Delta(U/I)$ (Fig. 7). Therefore, the $p_{\rm Var}$ vectors are uniformly distributed in angle on the $Q-U$ plane.

To confirm the effect found in Paper I, the coordinate system for measuring position angles must be transformed. Fig. 8 shows the two reference frames. In the normal (celestial) system, angles are measured between the vector $p_{\rm Var}$ and the Q-axis. In the new system, angles are measured between the vector $p_{\rm Var}$ and a radius drawn to connect the origin and $\bar{p}$ (the midpoint of $p_{\rm Var}$). If this angle is labelled ϕ then $\phi = 0^\circ$ represents a change in polarization at a constant position angle, and $\phi = 90^\circ$ represents a change in position angle at a constant polarization. The distribution of ϕ is plotted in Fig. 9a and it is clearly non-uniform. There is a large peak near $\phi = 90^\circ$, at the 3σ level.

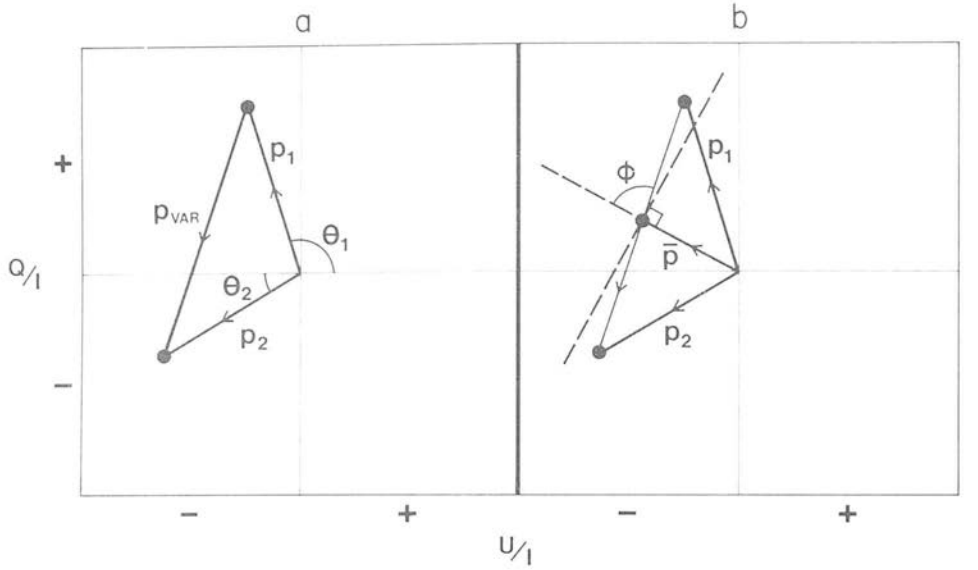


Figure 8. (a) Normalized Stokes parameters Q/I and U/I are plotted in the celestial system. The vectors $\mathbf{p}_1$ and $\mathbf{p}_2$ have degrees of polarization p_1 and p_2 and position angles θ_1 and θ_2. The inter-night change in polarization is given by $\mathbf{p}_{Var} = \mathbf{p}_2 - \mathbf{p}_1$. (b) Q/I and U/I are plotted in a reference frame defining the angle ϕ, which is the angle measured between the vectors $\bar{\mathbf{p}}$ and $\mathbf{p}_{Var}$.

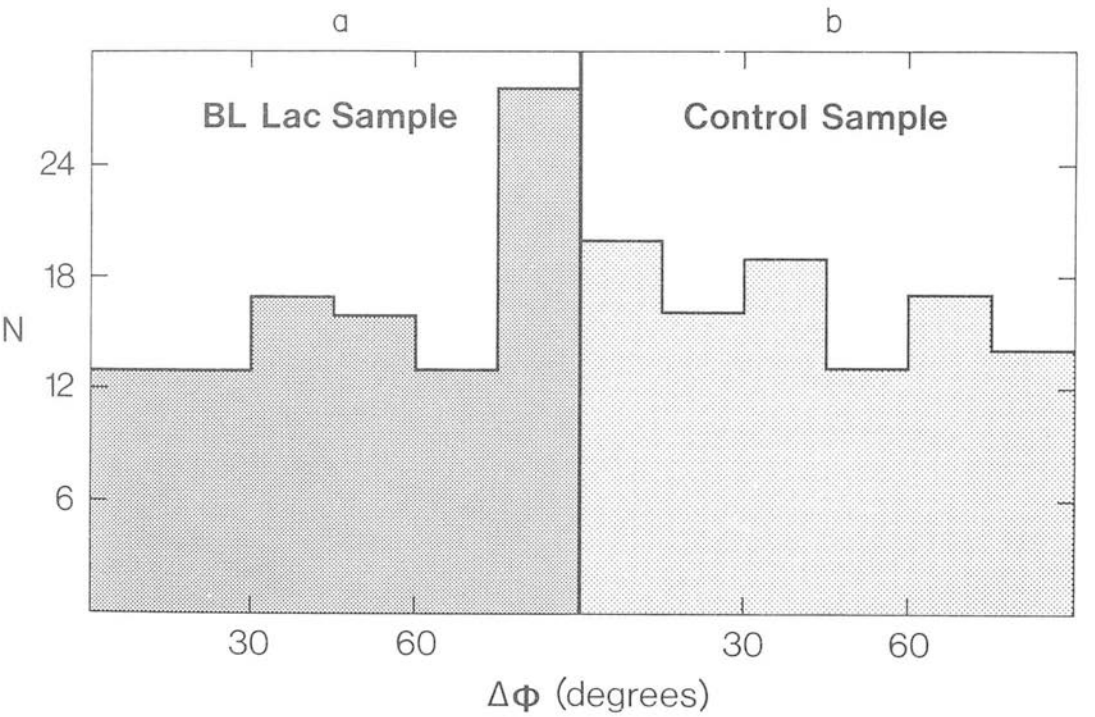


Figure 9. (a) The distribution of ϕ is plotted for 100 inter-night observations of polarization. The measurements involve the J, H and K band and include all objects in the sample; (b) This distribution of ϕ is plotted for a control sample of polarized, variable stars. The observations were made by K. Serkowski.

Since many of the individual measurements have poor signal-to-noise (in terms of ΔQ and ΔU), it is possible that the peak at 90° is an artifact of noisy data. Therefore, the same test was performed on a control sample of data: optical polarimetry of red variable stars by K. Serkowski (unpublished data). Many of the stars have long monitoring runs and the same number of data points were used. Although the physics of the polarization variations is completely different from BL Lac objects, the mean signal-to-noise of the two sets of measurements is similar. The Serkowski data is shown in Fig. 9b, and it is uniform in ϕ to well within the statistical errors.

The implication of the ϕ distribution is that position angle rotations and polarization changes are 'decoupled': rotations in position angle are not generally accompanied by polarization changes. There are physical reasons for preserving the distinction between p and ϕ. In a synchrotron model, changing p may be due to the addition of several variable components or to the compression of the magnetic field in a high surface brightness region. Changing ϕ may be due to rotating field lines or to the addition of several components; therefore, separate mechanisms may be at work when p and ϕ vary.

5 Conclusions

The important characteristics of individual BL Lac objects are summarized in Table 4 by a check against the corresponding property. Column (2) indicates a thermal component from the two-colour plot, and column (3) indicates an estimated 1–10 μm luminosity near or above the Eddington limit. Columns (4) and (5) show the presence of wavelength-dependent polarization and rapid position angle rotations (> 20 per cent/day). Column (6) indicates whether the internight changes in total and polarized flux are correlated or uncorrelated.

There has been much theoretical interest recently in relativistic jets as the source of con-

Table 4. Checklist of properties of BL Lac objects.

(1)	(2)	(3)	(4)	(5)	(6)
Object	Thermal component	$L_{MAX} \gtrsim L_{Edd}$	p(λ)	θ(t)	$\Delta p \propto \Delta I$
0235+164	–	✓	✓	✓	X
0735+178	–	–	✓	✓	✓
0754+101	–	–	–	–	–
0851+202	–	–	–	✓	✓
1156+295	–	✓	–	✓	X
1308+326	–	✓	–	–	X
1418+546	–	–	–	✓	–
1641+399	–	✓	–	–	X
1652+398	✓	–	–	–	–
1727+503	✓	–	–	–	–
2155-304	✓	–	–	–	–
2223-052	–	✓	–	–	–

tinuum energy in many types of radio sources (Blandford 1981). There have also been ambitious attempts to link many types of active nuclei in terms of a relativistic jet seen at different angles to the observer's line-of-sight (Marscher 1980, Blandford & Königl 1979, Königl 1981, Blandford & Rees 1978, Scheuer & Readhead 1979). These models try to account for the differences in radio, optical, and X-ray properties between radio-quiet quasars, radio-loud quasars, low frequency radio variables, OVV quasars, and BL Lac objects. Geometry plays an important role in these scenarios, and in all cases BL Lac objects are thought to be sources with the smallest viewing angle to the jet axis. Unfortunately, the evidence for this unified picture is mostly circumstantial, since most of the VLBI maps of jets have been of low luminosity sources, and it is a large jump to apply these structures to high luminosity quasars and BL Lac objects. However, recent theoretical work has shown that polarization data can be a powerful discriminant between various types of relativistic jet (Bjornsson 1982). These polarization observations offer a good change to test jet models by comparing luminosity and polarization properties of the data with the expected viewing angle and magnetic field geometry of the jet.

Acknowledgments

We thank the panel for the allocation of telescope time, and Terry Lee and the UKIRT staff for excellent support. In particular, John Clark and Malcolm Stewart contributed greatly to the implementation of the hardware and software for the polarimeter. We are grateful to Bill Zealey for able assistance at the telescope. We also thank the referee for suggestions which considerably improved the clarity of this paper. CDI acknowledges the receipt of an SERC/NATO Fellowship, and part of the work was carried out while in receipt of an SERC Studentship.

References

Aller, H. D., Aller, M. F. & Hodge, P. E., 1981. *Astr. J.*, **86**, 325.

Angel, J. R. P., Boroson, T. A., Adams, M. T., Duerr, R. E., Giampapa, M. S., Gresham, M. S., Gural, P. S., Hubbard, E. N., Kopriva, D. A., Moore, R. L., Peterson, B. M., Schmidt, G. D., Turnshek, D. A., Wilkerson, M. S. & Zotov, N. V. 1978. *Pittsburgh Conference on BL Lac Objects*, p. 117, ed. Wolfe, A. M., University of Pittsburgh.

Angel, J. R. P. & Stockman, H. S., 1980. *Ann. Rev. Astr. Astrophys.*, **18**, 321.

Bailey, J., Hough, J. H. & Axon, D. J., 1984. *Mon. Not. R. astr. Soc.*, **203**, 339.

Becker, W., 1937. *Z. fur Astrophys.*, **15**, 225.

Beichman, C. A., Prardo, S. H., Neugebauer, G., Soifer, B. T., Matthrews, K. & Wooten, H. A., 1981. *Astrophys. J.*, **247**, 780.

Bjornsson, C.-I., 1982. *Astrophys. J.*, **260**, 855.

Blandford, R. D. & Rees, M. J., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 328, ed Wolfe, A. M., University of Pittsburgh.

Blandford, R. D. & Konigl, A., 1979. *Astrophys. J.*, **232**, 34.

Blandford, R. D. *XXIII COSPAR Conference*, Budapest 1981, Reidel, Dordrecht, Holland, in press.

Bonoli, F., Braccesi, L., Frederici, L., Zitelli, V. & Formiggini, L., 1979. *Astr. Astrophys. Suppl.*, **35**, 391.

Borra, E. F., 1983. *Astrophys. J.*, **273**, L55.

Chanan, G. A., Margon, B. Helfand, D. J., Downes, R. A. & Chance, D., 1981. *Bull. Am. Astr. Soc.*, **13**, 848.

Condon, J. J., Hicks, P. D. & Jauncey, D. L., 1977. *Astr. J.*, **82**, 692.

Conway, R. G., Haves, P., Kronberg, P. P., Stannard, C., Vallee, J. P. & Wardle, J. F. C., 1974. *Mon. Not. R. astr. Soc.*, **168**, 137.

Cotton, W. D., Counselman, C. C., Geller, R. B., Shapiro, I. I., Wittels, J. J., Hinteregger, H. F., Knight, C. A., Rogers, A. E. E., Whitney, A. R. & Clark, T. A., 1979. *Astrophys. J.*, **229**, L115.

Craine, E. R., Duerr, R. & Tapia, S., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 99, ed Wolfe, A. M., University of Pittsburgh.
Eachus, L. J. & Liller, W., 1975. *Astrophys. J.* **200**, L61.
Elliott, J. L. & Shapiro, S. L., 1974. *Astrophys. J.*, **192**, L3.
Elvius, A., 1968. *Lowell Obs. Bull.*, **142**, 75.
Epstein, E. E., Fogarty, W. G., Mottmann, J. & Schneider, E., 1982. *Astr. J.*, **87**, 44.
Fogel, J. A., Persson, S. E., Aaronson, M. A. & Matthews, K., 1978. *Astrophys. J.*, **220**, 75.
Glassgold, A. E., Bregman, J. N., Huggins, P. J., Kinney, A. L., Pica, A. J., Pollock, J. T., Leacock, R. J., Smith, A. G., Webb, J. R., Wiśniewski, W. Z., Jeske, N., Spinrad, H., Henry, R. B. C., Miller, J. S., Impey, C., Neugebauer, G., Aller, M. F., Aller, H. D., Hodge, P. E., Balonek, T. J., Dent, W. A. & O'Dea, C. P., 1983. *Astrophys. J.*, **274**, 101.
Hawkins, M. R. S., 1983. Proc. *Quasars and Gravitational Lenses*, 24th Liege Astrophysical Symposium, p. 36, Université de Liège.
Impey, C. D. & Brand, P. W. J. L., 1981. *Nature*, **292**, 814.
Impey, C. D. & Brand, P. W. J. L., 1982. *Mon. Not. R. astr. Soc.*, **200**, 19.
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1982a. *Mon. Not. R. astr. Soc.*, **200**, 19 (Paper I).
Impey, C. D., Brand, P. J. W. L. & Tapia, S. 1982b. *Mon. Not. R. astr. Soc.*, **198**, 1.
Kellerman, K. I., Jauncey, D. L., Cohen, M. H., Shaffer, B. B., Clark, B. G., Broderick, J., Ronnang, B., Rydbeck, O. E. H., Matreyenko, L., Moiseyev, I., Vitkevitch, V. V., Cooper, B. F. C. & Batchelor, R., 1971. *Astrophys. J.*, **179**, 1.
Kinman, T. D., 1976. *Astrophys. J.* **205**, 1.
Knacke, R. F., Capps, R. W. & Johns, M., 1976. *Astrophys. J.*, **210**, L69.
Konigl, A., 1981. *Astrophys. J.*, **243**, 700.
MacLeod, J. M. & Andrew, B. H., 1968. *Astrophys. Lett.*, **1**, 243.
Marscher, A. P., 1980. *Astrophys. J.*, **235**, 386.
Miley, G., 1980. *Ann. Rev. Astro. Astrophys.*, **18**, 165.
Miller, J. S., 1978. *Comm. Astrophys.*, **7**, 175.
Miller, J. S. & Hawley, S. A., 1977. *Astrophys. J.*, **212**, L47.
Miller, J. S., French, H. B. & Hawley, S. A., 1978. *Astrophys. J.*, **219**, L85.
Moore, R. L. & Stockman, H. S., 1981. *Astrophys. J.*, **243**, 6.
Moore, R. L., McGraw, J., Angel, J. R. P., Duerr, R., Lebofsky, M., Rieke, G., Wisniewski, M., Axon, D., Bailey, J., Hough, J., Thompson, I., Breger, M., Schulz, H., Clayton, J., Martin, P., Miller, J., Schmidt, G., Africano, J. & Miller, H., 1982. *Astrophys. J.*, **260**, 415.
Neugebauer, G., Oke, J. B., Becklin, E. E. & Matthews, K., 1979. *Astrophys. J.*, **230**, 79.
O'Dell, S. L., Puschell, J. J., Stein, W. A. & Warner, J. W., 1978. *Astrophys. J. Suppl.*, **38**, 267.
Puschell, J. J. & Stein, W. A., 1980. *Astrophys. J.*, **237**, 331.
Puschell, J. J., Jones, T. W., Phillips, A. C., Rudnick, L., Simpson, E., Sitko, M., Stein, W. A. & Moneti, A., 1983. *Astrophys. J.*, **265**, 625.
Puschell, J. J., Stein, W. A., Jones, T. W., Warner, J. W., Owen, F., Rudnick, L., Aller, H. & Hodge, P., 1979. *Astrophys. J.*, **227**, L11.
Rieke, G. H. & Kinman, T. D., 1974. *Astrophys. J.*, **192**, L115.
Rieke, G. H., Grasdalen, G. L., Kinman, T. D., Hintzen, P., Wills, B. J. & Wills, D., 1976. *Nature*, **260**, 754.
Rieke, G. H., Lebofsky, M. J., Kemp, J. C., Coyne, G. V. & Tapia, S., 1977. *Astrophys. J.*, **218**, L37.
Rieke, G. H. & Lebofsky, M. J., 1979. *Ann. Rev. Astr. Astrophys.*, **17**, 477.
Rudnick, L., Owen, F. N., Jones, T. W., Puschell, J. J. & Stein, W. A., 1978. *Astrophys. J.*, **225**, L5.
Scheuer, P. A. G. & Readhead, A. S. C., 1979. *Nature*, **277**, 182.
Schmidt, M., 1974. *Astrophys. J.*, **193**, 505.
Schmitt, J. L., 1968. *Nature*, **218**, 663.
Seielstad, G. A., Cohen, M. H., Linfield, R. P., Moffett, A. T., Rommey, J. D. & Schitizzi, R. T., 1979. *Astrophys. J.*, **229**, 53.
Snyder, W. A., Davidson, A., Wood, K., Kinzer, R., Smathers, H., Shulman, S., Meekins, J., Yentis, D., Evans, W., Bryan, E., Chubb, T., Freidman, H., & Margon, B. 1980. *Astrophys. J.*, **237**, L11.
Stein, W. A., O'Dell, S. L. & Strittmatter, P. A., 1976. *Ann. Rev. Astr. Astrophys.*, **14**, 173.
Stocke, J., Liebert, J., Stockman, H., Danziger, J., Lub, J., Maccacarro, T., Griffiths, R., & Giommi, P. 1982. *Mon. Not. R. astr. Soc.*, **200**, 27P.
Stockman, H. S. & Angel, J. R. P., 1978. *Astrophys. J.*, **220**, L67.
Sulentic, J. W., Arp, H. & Lorre, J., 1979. *Astrophys. J.*, **233**, 44.

Tananbaum, H., Avni, Y., Branduard, G., Elvis, M., Fabbiano, G., Feigelson, E., Giacconi, R., Henry, J. P., Pye, J. P., Soltan, A. & Zamorani, G., 1980. *Astrophys. J.*, **234**, L9.
Tapia, S., Craine, E. R., Gearhart, M. R., Pacht, E. & Kraus, J., 1977. *Astrophys. J.*, **215** L71.
Usher, P. D., 1975. *Astrophys. J.*, **198**, L57.
Veron, P., 1978. In *Pittsburgh Conference on BL Lac Objects*, p. 20, ed Wolfe, A. M., University of Pittsburgh.
Visvanathan, N., 1973. *Astrophys. J.*, **179**, 1.
Wade, R., 1980. PhD dissertation, California Institute of Technology, California.
Weiler, K. W. & Johnston, K. J., 1980. *Mon. Not. R. astr. Soc.*, **190**, 269.
Wills, D., Wills, B. J., Breger, M. & Hsu, J. C., 1980. *Astr. J.*, **85**, 1555.

Mon. Not. R. astr. Soc. (1984) 210, 961–977

Infrared polarimetry and photometry of BL Lac objects – III

P. A. Holmes and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

C. D. Impey* *Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA*

P. M. Williams *United Kingdom Infrared Telescope Unit, 900 Leilani Street, Hilo, Hawaii 96720, USA*

Accepted 1984 May 10. Received 1984 May 10; in original form 1984 February 1

Summary. The data presented here is part of a continuing monitoring programme of BL Lac objects with *J*, *H* & *K* photometry and polarimetry. A total of 30 BL Lac objects have now been observed photometrically. We also have infrared polarimetry for 24 of these objects. The sample is sufficiently large to examine statistically, and several important correlations have emerged. Internight variations and wavelength dependence of polarization indicate that BL Lac objects, as a class, may be understood in terms of a relatively simple two-component model. Luminosity calculations suggest that relativistic beaming or non-spherical accretion is important in the most luminous objects. The $p(\lambda)$–p relation allows us to examine the magnetic-field structure of the emitting region.

1 Introduction

BL Lac objects are the most extreme example of active nuclei. Their study is vital to the understanding of all active galactic nuclei, because they allow us to probe the central power source directly. A large infrared database is necessary to examine their polarization properties statistically in a region where they emit most of their energy. Most polarimetry has previously been at optical wavelengths (e.g. Angel *et al.* 1978). Thirty objects have now been observed in our programme, and many of these for the first time with infrared polarimetry. This paper is the third resulting from the monitoring programme; see also Paper I (Impey *et al.* 1982) and Paper II (Impey *et al.* 1984).

* Present address: California Institute of Technology, Division of Astronomy and Physics, Pasadena, California 91125, USA.

2 Observations

All the observations were carried out on the 3.8-m United Kingdom Infrared Telescope on Mauna Kea, using a photovoltaic InSb detector and a rotating HR polaroid followed by a Lyot depolarizer. All the data were collected in two runs in 1982 November and 1983 January. Of the first run, only November 25 was not lost to bad weather. Of the six consecutive nights in January only half a night was lost to cloud; the weather was otherwise excellent. Three-colour (*JHK*) photometry was collected for 15 objects. Three-colour polarimetry was also collected for eight objects in total, and for six objects on more than one occasion. The *f*/35 chopping secondary was used with a chopping frequency of 7–8 Hz. The instrumental sensitivity and polarization parameters were calculated at least twice per night. The broad-band filters used and their effective wavelengths were J (1.25 μm), H (1.65 μm), K (2.2 μm) and L (3.4 μm).

The calibration and reduction procedures were described in Paper 1. The analyser efficiency and position angle zero point were redetermined each time the polarimeter was mounted on the telescope. The polarimetry was calibrated using nearby unpolarized stars, and the instrumental polarization, of less than 0.5 per cent, was repeatable over the run to better than 0.25 per cent. Flux calibrations have taken into account the effect of differing flux distributions of the objects and the calibrators in the broad pass bands.

The 'patchiness' of the interstellar medium makes the extinction difficult to determine. Estimates of the interstellar extinction correction were taken from the maps of reddening derived from H I and galaxy counts (Burstein & Heiles 1982), and by using the extinction values of Savage & Mathis (1979). Unlike the sample in Paper I, six of the objects observed here required significant correction ($A_K \geqslant 0.03$). BL Lac (2200 + 420) is particularly reddened ($A_{J,H,K} = 0.35, 0.25, 0.15$), and only 1156 + 295 and 1308+ 326 required no correction at all.

The quoted errors in the flux are the formal errors due to noise in the signal. They do not contain a factor due to the inhomogeneity in galactic extinction, but do contain a factor associated with the degree of repeatability of offset magnitudes of standard stars over one night.

The objects were acquired either directly on the Scanco TV system or by offsetting from a nearby bright star with accurate coordinates and then peaking on the 2.2 μm signal of the object. All observations were made with an 8 arcsec aperture, and near dark of moon.

3 Results

3.1 INDIVIDUAL OBJECTS

(1) 0215 + 015. This is the highest-redshift and most luminous BL Lac object known (Pettini *et al.* 1983). We observed 0215 + 015 on January 10 and 12. Although the polarization was high (16–17 per cent), the object was faint ($H = 13$) which precluded further observation. There was no evidence for variability over a time-scale of two days.

(2) 0219 + 428. We observed 3C 66A on three consecutive nights in January and found no evidence for variability in flux or position angle. Wavelength dependence of polarization ($dp/d\lambda > 0$) was seen on the first two nights but was absent on the third.

(3) 0235 + 164. The optical polarization has a range of 5–44 per cent (Angel & Stockman 1980) and from 2–36 per cent in the infrared (Paper I; see also Impey, Brand & Tapia 1982). Although our data show no evidence for flux variability in January, this object brightened by around 1 mag between November 25 and January 8. There was a strong wavelength

dependence of polarization in November, with p increasing with frequency, but this dependence was absent in January. There appeared to be stochastic changes in the polarization behaviour from night to night, the position angle remaining relatively stable whilst the polarization degree varied by a factor of 2 between January 8 and 9. Between the two observing runs in November and January, there was a rotation in position angle of approximately 70° and a steepening of the spectral index from 1.1 to 1.6.

(4) 0422 + 004. Two nights' observations showed the source to have dimmed $\Delta m = 0.3$ mag with an increase in polarization from $p_K = 13.6$ to 19.8 per cent. No rotation in position angle was observed. We believe our observations provide the first infrared polarimetric data on this object.

(5) 0735 + 178. This object has a high and variable polarization in the optical (Carswell *et al.* 1974) and in the infrared (Papers I and II). Five nights observing in January showed relative stability of flux, polarization and position angle. Paper I reported that polarization increased towards shorter wavelengths, but such behaviour was not observed in this run. A large rotation of position angle was observed between November and January, $\Delta\theta \sim 80°$.

(6) 0754 + 101. This Ohio radio source OI 090.4 has a range in optical polarimetry of 3–26 per cent, and in infrared polarimetry of 2–19 per cent (Papers I and II). It has an archival variability of $\Delta m = 2.5$ mag and given the rapid variability of polarization, it would be surprising if rapid flux variations were not eventually observed. During this observing run, only small fluctuations in flux and polarization were seen and the position angle remained fairly stable.

(7) 0829 + 046. This Parkes and Ohio Source has an optical polarization range of $p = 1.7$–22 per cent (Zekl, Klare & Appenzeller 1981). Our observations, which are the only infrared polarimetry data for the object, reveal a high and variable polarization, and variable flux. 0829 + 046 was observed on three consecutive nights in January. The object brightened by $\Delta m \sim 0.5$ mag corresponding to an increase in the polarization by a factor of 2, and a small rotation in position angle of $\Delta\theta \sim 10$–15°. This behaviour is consistent with a highly polarized power-law slope 'turning on' (see later discussion).

(8) 0851 + 202. The optical polarization of OJ 287 has a range of 0.4 − 29 per cent (Holmes *et al.* 1984; Angel & Stockman 1980). Our data reveal large variations in flux and polarization properties, with the flux decreasing by 7 per cent in 1 hr from a peak brightness which corresponded to a minimum in the polarization degree. A model for this behaviour is described in Holmes *et al.* (1984).

(9) BL Lac. This prototype object is one of the best studied of all blazars, and its properties have been reviewed by Stein, O'Dell & Strittmatter (1976), Miller (1978) and in Paper II. It was observed on only two nights on these observing runs, on November 25 and January 9. The polarization rose from 9–12 per cent while the flux dropped by more than 1 mag. In November, the position angle was $\theta \sim 100°$ with a rotation of $\Delta\theta \sim 80°$ between the two dates.

(10) Other objects.

(a) 0736 + 017. This object is a BL Lac-like object as it has variable optical polarization 0–6 per cent (Moore & Stockman 1983). Our data are the first infrared polarimetric observations.

(b) 0912 + 297 was observed on one night with *JHK* photometry and polarimetry. There was marginal evidence for wavelength dependence of polarization; this object has a stable polarization position angle. The infrared spectral index $\alpha_{JK} = 0.7$, ($S_\nu \propto \nu^{-\alpha}$) was distinctly different from that of $\alpha = 0.99$ (Paper II) observed in 1981 April and $\alpha = 1.8$ (Paper I) observed in 1980 April.

(c) 1156 + 295 is a strong-lined quasar at a redshift $z = 0.728$ (Schmidt 1974) with an

optical polarization of up to 28 per cent (D. Wills, personal communication). Our observations show the highest infrared polarization yet in this object at $p = 18$ per cent.

(d) 1253−055 (3C 279) is an OVV quasar, and at peak brightness is one of the most luminous known objects in the Universe. The polarization was high at $p_K = 14.5$ per cent but the source was too faint to make further polarimetry worth while.

(e) 1308 + 326. The polarimetry ranges are 0–28 per cent in the optical (Zekl *et al.* 1981) and 7–20 per cent at infrared wavelengths (Paper I). Although the object was highly polarized at $p_K \sim 17.6$ per cent, it was too faint for further polarimetry.

(f) 1418 + 546 was confirmed a BL Lac by Craine, Duerr & Tapia (1978), and has an archival variability of $\Delta M_B \sim 4.8$ mag (Paper I). The infrared polarization has been seen to vary from 9 to < 1 per cent in just three months, with a maximum of $p_K = 18$ per cent.

4 Discussion

4.1 INFRARED BL LAC DATABASE

The data presented in Table 1 extend the infrared BL Lac database considerably. The polarimetry taken in this observing run was far more accurate than that presented in Papers I and II, due in part to the insertion of a Lyot depolarizer in the polarimeter optics.

Table 1. Flux and polarization properties of the blazars in the sample.

OBJECT	DATE	w/b	F_ν(mJy)	$\sigma(F_\nu)$	p%	σ(p%)	θ°	$\sigma(\theta^\circ)$
0215+015	10 Jan 83	K	8.6	0.2	15.9	3.6	145	7
		H	6.2	0.1	17.4	1.1	145	2
		J	4.2	0.1	-		-	
	12 Jan 83	K	8.6	0.2	-		-	
		H	6.3	0.1	-		-	
		J	4.6	0.1	-		-	
0219+428	8 Jan 83	K	14.6	0.4	16.5	1.5	17	3
		H	11.6	0.2	18.2	1.6	25	2
		J	8.6	0.2	22.5	4.8	20	6
	9 Jan 83	K	12.5	0.1	15.7	1.1	23	2
		H	10.9	0.4	16.2	1.2	27	2
		J	7.1	0.1	20.8	3.9	24	5
	10 Jan 83	K	13.9	0.2	15.6	1.4	24	3
		H	10.9	0.1	15.4	1.0	22	2
		J	8.3	0.1	14.5	2.1	28	4
0235+164	25 Nov 82	K	5.8	0.1	14.7	2.0	26	4
		H	3.9	0.1	17.6	3.2	40	5
		J	3.2	0.1	23.8	2.6	31	3
	8 Jan 83	K	16.3	0.4	17.8	1.8	106	3
		H	11.6	0.2	13.5	0.7	103	1
		J	6.6	0.1	18.1	1.5	103	2
	9 Jan 83	K	15.9	0.2	8.5	1.2	117	5
		H	10.8	0.4	7.1	2.3	116	9
		J	6.3	0.1	-		-	
	10 Jan 83	K	14.3	0.3	6.9	1.5	107	7
		H	10.0	0.1	9.6	2.8	111	9
		J	6.1	0.1	7.1	3.5	114	14
	12 Jan 83	K	15.4	0.3	11.6	0.4	105	1
		H	10.8	0.2	11.5	2.2	105	6
		J	7.0	0.2	16.2	2.2	109	4
0422+004	8 Jan 83	K	15.7	0.5	13.6	1.3	166	3
		H	11.7	0.2	-		-	
		J	8.4	0.2	-		-	
	12 Jan 83	K	11.9	0.2	19.8	1.5	164	3
		H	8.6	0.2	13.7	1.6	168	3
		J	6.1	0.2	-		-	
0735+178	25 Nov 82	K	28.1	0.6	13.2	6.4	21	6
		H	21.4	0.2	20.8	0.4	7	1
		J	15.1	0.3	19.8	1.0	9	2
	7 Jan 83	K	21.5	0.2	14.2	1.1	136	8
		H	17.5	0.6	12.6	1.4	135	8
		J	11.5	0.4	14.6	2.9	141	10
	8 Jan 83	K	22.7	0.6	12.4	1.2	129	3
		H	16.9	0.3	11.3	0.7	133	2
		J	11.2	0.2	11.8	1.2	136	3

Table 1 – *continued*

OBJECT	DATE	B/b	F (mJy)	σ(F)	p%	σ(p%)	θ°	σ(θ°)
	9 Jan 83	K	22.3	0.2	13.7	1.3	137	3
		H	17.7	0.7	12.7	1.5	138	3
		J	10.6	0.2	17.0	1.4	127	2
	10 Jan 83	K	20.9	0.4	10.8	1.3	131	4
		H	15.5	0.2	14.2	1.2	131	2
		J	11.3	0.1	12.6	0.6	134	1
	12 Jan 83	K	21.9	0.4	12.2	0.8	141	2
		H	16.3	0.3	12.0	1.5	132	4
		J	11.9	0.3	10.0	2.0	135	6
0736+017	10 Jan 83	K	6.6	0.1	7.3	4.3	44	19
		H	4.6	0.1	–		–	
		J	3.2	0.1	–		–	
0754+101	8 Jan 83	K	18.4	0.5	8.7	1.1	169	4
		H	13.3	0.3	8.9	1.4	179	5
		J	9.1	0.2	10.2	0.6	177	2
	9 Jan 83	K	15.9	0.2	6.2	1.2	0	7
		H	11.9	0.4	8.4	1.0	176	4
		J	8.7	0.2	6.2	1.8	176	9
	11 Jan 83	K	18.4	0.4	8.5	0.8	4	3
		H	13.7	0.2	5.3	1.0	163	5
		J	10.2	0.1	10.2	2.5	179	7
0829+046	10 Jan 83	K	21.3	0.4	7.8	0.2	130	1
		H	15.4	0.2	7.4	0.9	129	4
		J	10.5	0.1	6.7	2.3	118	10
	11 Jan 83	K	31.4	0.6	13.8	1.0	116	2
		H	23.3	0.2	13.8	0.9	114	2
		J	16.7	0.2	16.2	0.5	114	1
	12 Jan 83	K	28.6	0.6	14.0	0.8	129	2
		H	19.9	0.4	13.6	0.5	122	1
		J	14.0	0.4	12.1	1.0	122	2
0851+202	25 Nov 83	K	32.2	0.6	9.1	0.4	6	2
		H	24.2	0.2	9.1	0.5	10	2
		J	17.3	0.3	8.7	1.3	10	4
	7 Jan 83	K	88.8	0.8	9.4	0.3	99	1
		H	77.1	2.8	9.2	0.2	78	1
		J	60.0	2.2	8.3	0.9	71	3
	8 Jan 83	L	183.3	5.0	–		–	
		K	133.2	3.6	3.5	0.3	78	4
			123.7	3.4				
		H	103.6	1.9	2.6	0.2	71	2
			98.0	1.8				
		J	77.6	1.4	1.5	0.2	98	3
			74.8	1.4				
	9 Jan 83	L	168.7	3.4	–		–	
		K	121.4	1.2	6.8	0.6	111	3
		H	85.4	3.1	4.6	0.1	114	1
		J	67.0	1.2	4.4	0.4	116	3
	10 Jan 83	L	159.6	4.8	–		–	
		K	106.8	1.9	5.7	0.4	114	2
		H	80.8	0.8	4.4	0.1	120	1
		J	62.2	0.6	3.6	0.3	119	2
	11 Jan 83	L	152.5	3.1	–		–	
		K	101.0	1.8	10.2	0.3	109	1
		H	75.0	0.8	8.4	0.3	107	1
		J	58.3	0.5	8.4	0.3	109	1
	12 Jan 83	L	173.4	1.7	–		–	
		K	113.9	2.1	6.4	0.2	120	1
		H	87.7	1.6	6.0	0.2	118	1
		J	67.0	1.8	6.1	0.3	118	1
0912+297	12 Jan 83	K	7.7	0.2	10.5	1.9	175	6
		H	6.4	0.1	12.1	1.0	179	2
		J	5.2	0.2	13.5	2.8	0	6
1156+295	9 Jan 83	K	3.5	0.1	18.2	5.6	64	10
		H	2.6	0.1	–		–	
		J	1.8	0.1	–		–	
1253–055	11 Jan 83	K	3.9	0.1	14.5	3.5	49	8
		H	2.5	0.1	–		–	
		J	1.8	0.1	–		–	
1308+326	12 Jan 83	K	3.2	0.1	17.6	3.0	136	6
		H	2.0	0.1	–		–	
		J	1.3	0.1	–		–	
1418+546	11 Jan 83	K	13.4	0.3	7.6	2.8	92	12
		H	10.1	0.2	–		–	
		J	7.4	0.2	–		–	
2200+420	25 Nov 83	K	66.1	1.2	9.0	1.7	103	6
		H	–		5.8	1.0	99	5
		J	45.9	0.9	6.2	1.9	99	9
	9 Jan 83	K	24.0	0.2	–		–	
		H	21.4	0.8	–		–	
		J	14.0	0.3	–		–	

In the total programme we have now observed 30 BL Lac objects photometrically, and, in most cases, polarimetrically also. There are now more than 80 known BL Lac and BL Lac-like objects, and the number is increasing rapidly. 30 objects, however, constitute a sufficiently large sample to be considered representative of known BL Lacs, but there will be certain selection effects. Most BL Lacs have been discovered from positions of flat-spectrum radio sources. There is no known example of a radio quiet BL Lac. This may not be a selection effect, as BL Lacs are intrinsically strong radio emitters compared to quasars.

Paper II employs a two-colour plot to illustrate the power-law nature of spectral energy distributions for BL Lac objects. Nearly all of our current observations lie within photometric errors of the power-law line, so there is no need to invoke a thermal component.

4.2 LUMINOSITY

The luminosity calculations are described in detail in Papers I and II. The blazars in our sample cover a wide range of luminosities, a factor of 120 in the data presented here, and a factor of 1.2×10^3 in the total programme. The huge luminosities and short variability time-scales constrain the possible power sources very tightly, and in some cases, spherical accretion on to a supermassive black hole is ruled out.

Table 2 shows data from this paper only Column (1) shows object name, (2) redshift, (3) luminosity distance, (4) near-infrared spectral index at maximum flux density, (5) integrated 1–2 μm luminosity. The flux of OJ 287 was seen to vary by 7 per cent in just 1 hr. Defining the variability time-scale in the object's rest frame as

$$\tau_{\min} = 10 \text{ per cent } [\Delta t/(1+z)]\,(L/\Delta L)$$

then we find that the integrated 1–10 μm luminosity violates the Elliot & Shapiro (1974) spherical accretion criterion

$$\log \tau_{\min}(\text{secs}) \geqslant \log L\ (\text{erg s}^{-1}) - 43.1.$$

Four other objects have so far been seen to violate this criterion (see Papers I and II) namely, BL Lac, 0235 + 164, 1308 + 326 and 3C 446, suggesting that jet emission is an important mechanism in blazars. We also find that the 1–2 μm luminosity of OJ 287, together with the

Table 2. Luminosity and variability of blazars in the current sample.

Object	z	d(Mpc)	α_{IR}	$L_{1-2\mu m}$ (erg s^{-1})
0215+015	1.686	6740	1.11	4.3×10^{46}*
0219+428	0.444	1780	0.94	4.8×10^{45}*
0235+164	0.852	3410	1.61	2.2×10^{46}
0735+178	-0.424	1700	1.25	6.6×10^{45}
0736+017	0.191	760	1.30	3.7×10^{44}*
0851+202 (OJ287)	0.306	1220	0.96	2.1×10^{46}*
1156+295	0.729	2920	1.22	3.2×10^{45}
1253-053	0.538	2150	1.32	1.9×10^{45}
1308+326	0.996	3980	1.59	6.0×10^{45}
2200+420	0.069	276	0.65	6.0×10^{44}

* indicates highest luminosity observed.

Fabian & Rees (1978) luminosity criterion, $L \leqslant 2\times10^{41}(\eta/0.1)\tau_{\min}\,\mathrm{erg\,s^{-1}}$ requires an efficiency of $\eta \gtrsim 20$ per cent, which approaches the limit even for matter falling onto a maximally rotating black hole.

With this increased sample we can test for correlations between luminosity and other source parameters. The surprising correlation, discovered in Paper II, between the range of position angles and maximum infrared luminosity is still evident. Such a relationship, if confirmed, would be very important since it connects blazar luminosity with a purely geometrical quantity. Position-angle rotations reflect the magnetic field behaviour. Large rotations are caused by either (1) large rotations of the magnetic field in the emitting volume or (2) small rotations in a field that makes a small angle to the line-of-sight. The updated data, however, show a weaker correlation than before, yielding a Spearman coefficient $r_s = 0.538$ for 12 data points i.e. correlated at the 5 per cent significance level. It is clear that both the position-angle range and the maximum luminosity observed will scale with the number of observations (N); to test whether the relationship is real, or merely due to their dependence upon N (see Fig. 1), we use the Spearman partial rank correlation coefficient.

This test may be formally stated as follows (see Macklin 1982). For three variables (A, X, Y) we wish to test the null hypothesis that the ($A-X$) correlation arises entirely from the ($A-Y$) and ($X-Y$) correlations separately. The partial rank correlation coefficient is defined as

$$r_{AX,\,Y} = \frac{r_{AX} - r_{AY}\cdot r_{XY}}{[(1-r_{AY}^2)(1-r_{XY}^2)]^{1/2}}$$

where r_{ij} is the Spearman coefficient for the correlation between the two variables ij. The significance level is defined by

$$D_{AX,\,Y} = (1/2)(N-4)^{1/2}\ln\left(\frac{1+r_{AX,\,Y}}{1-r_{AX,\,Y}}\right)$$

$D_{AX,\,Y}$ is distributed normally about zero with unit variance. Applying this test to the parameters we find the following coefficient values

$r_{\theta L} = 0.538$ $\quad r_{\theta L,\,N} = 0.349$ $\quad D_{\theta L,\,N} = 1.03$

$r_{\theta N} = 0.629$ $\quad r_{\theta N,\,L} = 0.503$ $\quad D_{\theta N,\,L} = 1.57$

$r_{LN} = 0.476$ $\quad r_{LN,\,\theta} = 0.210$ $\quad D_{LN,\,\theta} = 0.60$

for $N = 12$.

We conclude that while each pair of parameters is correlated at approximately 5 per cent level of significance, there is insufficient evidence to accept any given correlation as intrinsic. More observations should delineate between real and apparent relations.

A correlation between maximum infrared polarization and maximum luminosity has emerged from the updated data, with a Spearman coefficient $r_s = 0.582$ for 15 points, i.e. correlated at the 5 per cent significance level. As before, we are unable to say whether this relationship is intrinsic or wholly due to correlations with the number of observations.

One of the most fascinating set of correlations to have emerged is the redshift, luminosity, spectral-index set. Applying the partial rank test, as before, we are able to reject the $\alpha z, L$ null hypothesis at 95 per cent confidence level. It is more likely that the redshift-spectral index relation is intrinsic since $D_{\alpha z,\,L} > D_{\alpha L,\,z}$.

There is a selection effect operating in the α-luminosity relation. Let us consider a standard blazar, that is, one which emits a constant luminosity, irrespective of spectral

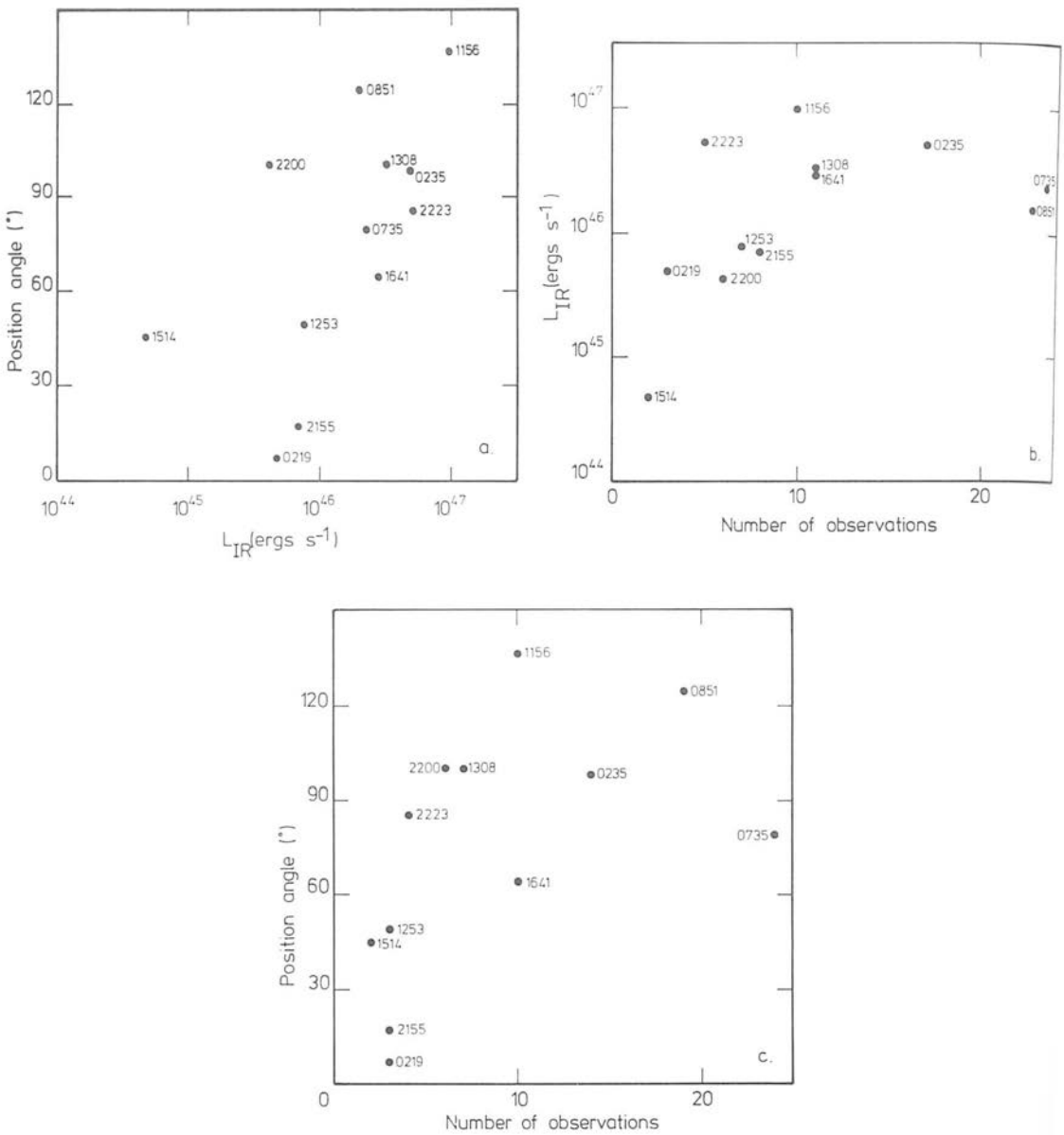


Figure 1. (a) Position angle range ($\Delta\theta$) plotted against $\log_{10}$ maximum IR luminosity. Spearman coefficient $r_S = 0.538$, for 12 data points. (b) Log L (IR) max plotted against number of observations, $r_S = 0.476$. (c) $\Delta\theta$ plotted against number of observations, $r_S = 0.629$.

index. For objects at greater redshifts, we would preferentially detect those with shallower, rather than steeper, spectral indices. If this were the only effect in operation, we would see an α-luminosity anti-correlation. This is not observed.

We consider two possible causes for the α–z relation. Firstly, there is some evidence to suggest that spectral curvature may be a general property of blazars. Many multiwavelength observations have revealed spectra which steepen with increasing frequency in the infrared-

Table 3. Luminosity, polarization and spectral index of all blazars.

Object	z	$L_{1-2\mu m}$ (erg s^{-1})	$P_{IR}^{max}(\%)$	ΔP(range %)	$\Delta\theta^{0}$(range)	$\bar{\alpha}_{IR}$
0215+015	1.686	4.3×10^{46}	H 17.4	-	-	1.18
0219+428	0.444?	4.6×10^{45}	J 22.5	8.0	7	0.96
0235+164	0.852	4.8×10^{46}	J 36.2	32.3	98	1.97
0306+103	-	-	K 3.6	1.4	69	-
0422+004	-	-	K 19.8	6.2	2	1.16
0735+178	0.424	2.3×10^{46}	J 32.6	22.6	79	1.19
0736+017	0.191	3.7×10^{44}	K 7.3	-	-	1.30
0754+101	-	-	K 18.9	16.6	123	1.02
0818-131	-	-	J 21.4	15.4	37	0.78
0829+046	-	-	J 16.2	9.5	14	1.21
0851+202	0.306?	2.1×10^{46}	H 19.7	18.2	124	0.98
0912+297	-	-	J 13.5	8.4	55	0.91
1147+245	-	-	K 14.8	4.7	4	1.40
1156+295	0.729	9.6×10^{46}	K 18.2	14.6	136	1.20
1253+055	0.538	7.5×10^{45}	K 16.2	6.5	49	1.46
1308+326	0.996	3.2×10^{46}	K 20.1	12.9	100	1.03
1335-127	-	-	-	-	-	1.7
1400+162	0.244	1.2×10^{45}	-	-	-	1.2
1418+546	-	-	K 17.9	17.0	128	1.01
1514-241	0.049	4.8×10^{44}	H 7.4	4.6	45	0.86
1538+149	-	-	-	-	-	0.92
1641+399	0.595	2.8×10^{46}	K 16.2	12.9	64	1.6
1652+398	0.034	3.1×10^{44}	-	-	-	0.36
1727+503	0.055	7.9×10^{43}	-	-	-	1.0
1749+096	-	-	-	-	-	2.0
1921-293	0.353	4.5×10^{45}	K 13.9	-	-	1.5
2155-304	0.170	6.9×10^{45}	K 3.0	2	17	0.35
2200+420	0.069	4.2×10^{45}	H 15.1	10.9	100	1.28
2223-052	1.404	5.1×10^{46}	K 16.3	10.8	85	1.5
2254+074	-	-	K 17.4	-	-	-

optical-ultraviolet regions [e.g. Bregman *et al.* 1984 (0735 + 178); Glassgold *et al.* 1983 (1156 + 295)]. Also, non-simultaneous observations at optical and infrared wavelengths show a difference in mean spectral index ($\Delta\alpha = \bar{\alpha}_{opt} - \bar{\alpha}_{IR}$) of typical value $\Delta\alpha \simeq 0.4$ for many blazars. Spectral curvature will be discussed in detail in a future paper. If such positive ($d\alpha/d \ln \nu > 0$) curvature is general, then for higher-redshift objects, we would be looking at a steeper part of the flux spectrum.

A second possibility is evolution of the population with cosmic time. A given power-law spectrum maintains its spectral index irrespective of recessional velocity. If verified by further observation, the α–z relation would therefore indicate intrinsic evolution of the blazar energy source.

The parameter values for the whole programme are given in Table 3. It shows in column (1) object, (2) redshift, (3) the maximum 1–2 μm luminosity, (4) the maximum infrared polarization, (5) position-angle range, and (6) the mean infrared spectral index.

4.3 RELATION BETWEEN POLARIZATION, SPECTRAL INDEX, AND POSITION ANGLE RANGE

For pure synchrotron radiation from a power-law electron-energy distribution, the maximum polarization (II) possible is from a perfectly aligned magnetic field, and is given by

$$\mathrm{II} = (1 + \alpha)/(5/3 + \alpha).$$

It is feasible, therefore, that a plot of α_{IR} against $p_{\mathrm{IR}}^{\mathrm{max}}$ would reveal a relation of this form. It must be remembered, however, that the members of the sample will not generally all achieve the same degree of magnetic-field alignment. Neither may we suppose that the maximum p_{IR} observed is related to the maximum ever achieved for any individual objects, particularly where only a small number of observations are available.

A strong correlation between mean infrared spectral index and maximum polarization for 14 objects was reported in Paper I. The expanded data set of 17 objects in Paper II showed a somewhat diluted correlation, with a Spearman coefficient $r_s = 0.68$. The further expanded and updated sample of 22 objects presented here is shown in Fig. 2. As the number of objects in the sample and the number of observations have increased, so the correlation has faded. It appears as though a second population is emerging from the observations, where $p_{\mathrm{IR}}^{\mathrm{max}}$ and α are anti-correlated. It is more feasible, however, that the majority of blazars in our sample have typical spectral indices in the range $\alpha = 1.2 \pm 0.4$, and a maximum infrared polarization in the range $p_{\mathrm{max}} = 12$–24 per cent. The apparent correlation in the early data was probably largely due to the fortuitous inclusion of extreme objects, such as 0235 + 164, 0306 + 102 and 2155 − 052. We find no significant correlation for $p_{\mathrm{IR}}^{\mathrm{max}}$ or Δp with either $\bar{\alpha}_{\mathrm{IR}}$ or α_{max}.

It is also interesting to note that $p_{\mathrm{IR}}^{\mathrm{max}}$ is unrelated to the range of position angles ($\Delta\theta$). These two parameters depend only upon the magnetic-field geometry; polarization varies with the degree of field alignment, whereas position-angle range depends upon field rotations only. So we see that for the magnetic field in the emitting region, the degrees of maximum alignment and rotation are not strongly correlated.

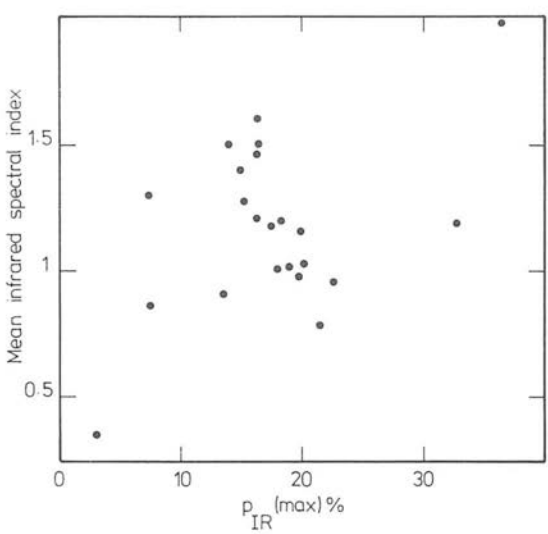


Figure 2. Mean spectral index ($\bar{\alpha}_{1-2\,\mu\mathrm{m}}$) plotted against $p_{\mathrm{IR}}^{\mathrm{max}}$, suggesting a possible emergence of a second anti-correlated population $r_s = -0.021$.

We do find, however, that the range of polarizations and $\Delta\theta$ are correlated at the 1 per cent significance level ($r_S = 0.662$ for 20 points). The Spearman partial rank coefficients show that this relationship may be due entirely to a strong correlation of each parameter with the number of observations.

4.4 WAVELENGTH DEPENDENCE

4.4.1 $p(\lambda)$

In this section, we use the notation $p(\lambda)$, $\theta(\lambda)$ to represent wavelength-dependence of polarization and position angle respectively. One of the defining characteristics of blazars as a class is their high and variable linear polarization. The canonical explanation for this phenomenon is emission of incoherent synchrotron radiation from relativistic electrons. For a single undiluted synchrotron source, we would expect wavelength-independent polarization and position angle. Furthermore, the continuity in flux from optical to infrared suggests the same emission mechanism in the same emitting region, and thus we would expect similar polarimetric properties. For most objects, this is found to be the case. Some objects, however, exhibit significant rotations of position angle (Moore *et al.* 1980; Holmes *et al.* 1984) and changes in the degree of polarization with wavelength (Paper I; Puschell & Stein 1980).

Nordsieck (1976) has shown that polarization is a function of local spectral index. Björnsson & Blumenthal (1982) demonstrated that the maximum deviation from the relation $p \propto (\alpha + 1)/(\alpha + 5/3)$ is approximately 7.5 per cent for concave downward spectra ($d\alpha/d \ln \nu > 0$). This variation with spectral index is too small to account for the observed $p(\lambda)$ of some objects and we are forced to resort to several-component models for an explanation.

In Paper I, only two objects, AO 0235+164 and PKS 0735+178, were found to have $p(\lambda)$. Wavelength dependence of polarization or position angle seems to be a transient phenomenon (e.g. BL Lac, Knacke, Capps & Johns 1976; Puschell & Stein 1980). No blazars have yet been observed to show consistently $p(\lambda)$ or $\theta(\lambda)$. 0235 + 164 showed a strong $p(\lambda)$ with polarization increasing into the blue, and high polarization degree in our November observations, but no such dependence in January. 0735 + 178 showed no evidence for wavelength dependence in either November or January (*cf.* Paper I). As we shall show, such transience is entirely consistent with a model consisting of two or more variable components; the components differ in spectral index and polarization properties, rotating and/or varying with respect to one another.

Although dilution by an underlying galaxy of a pure synchrotron source would result in $p(\lambda)$, it has never been convincingly demonstrated to be the cause of the observed wavelength dependence of polarization for any object. The effect of dilution will be weak, since the non-thermal flux usually completely dominates any galactic component. Also, variation arguments and the transience of $p(\lambda)$ eliminate the possibility that galactic dilution is entirely responsible (Bailey *et al.* 1981).

4.4.2 $p(\lambda)$–p

Bailey, Hough & Axon (1983, hereafter BHA) first noted the trend of increasing wavelength dependence with increasing polarization, $p(\lambda)$–p. From observations of seven blazars (6 BL Lac objects and one highly polarized quasar), they find that the ratio of optical to infrared polarizations increases with optical polarization for $p_{opt} > 10$ per cent, and is not significantly different from unity for $p < 10$ per cent. Their 25 data points yield a Spearman

Rank Correlation coefficient $r_s = 0.765$, revealing a correlation at the 99.9 per cent confidence level. Their data are clearly inconsistent with wavelength-independent polarization, although a large χ^2 value ($\chi^2/N = 2.1$) reveals a genuine scatter of the data about the best-fit line, not due to measurement error. χ^2 is evaluated from the equation

$$\chi^2 = \sum_i [(y_i - mx_i - c_i)^2/(\sigma_y^2 + m^2\sigma_x^2)].$$

It should be noted here that two unrelated variables (x, y) would produce a weak correlation when plotted as x/y versus x. This 'stretching effect' is small and is not the cause of the observed correlation.

The $p(\lambda)$–p relationship is also seen in our narrower-band observations in the infrared. At the time the observations in Papers I and II were made, the polarimetric accuracy was poor and yet a correlation is still evident. For Paper I data, a plot of p_H/p_K versus p_H yields a Spearman coefficient $r_s = 0.663$ (correlated at 99 per cent confidence level). The small χ^2 value ($\chi^2/N = 0.6$) results from the poor accuracy and prevents us inferring intrinsic 'cosmic' scatter. A similar result is found by plotting p_J/p_K against p_J for Paper II data.

The most marked confirmation of the $p(\lambda)$–p phenomenon came with the data presented in this paper (see Fig. 3). We choose the longest wavelength range from J (1.25 μm) to K (2.2 μm). The best-fit line yields a chi-square value $\chi^2 = 29.5$ for 28 data points, compared with the value for the null hypothesis ($m = 0$, $c = 1$, i.e. wavelength-independent polarization) of $\chi_0^2 = 169$. The Spearman coefficient $r_s = 0.81$ reveals a correlation at the 99.9 per cent confidence level. Unlike BHA we find that polarizations at two wavelengths are *not* equal within errors for values below $p = 10$ per cent. If we omit the OJ 287 data (for which we already have a model), however, we find that $p_J \sim p_K$ for $p_{J,K} < 10$ per cent, in agreement with BHA's result, for which we offer a possible explanation in the following section. We also draw attention to the $p(\lambda)$–p relationship, not only for the sample, but also for the individual members of the sample. All six objects, for which there are three or more points, exhibit a $p(\lambda)$–p relation.

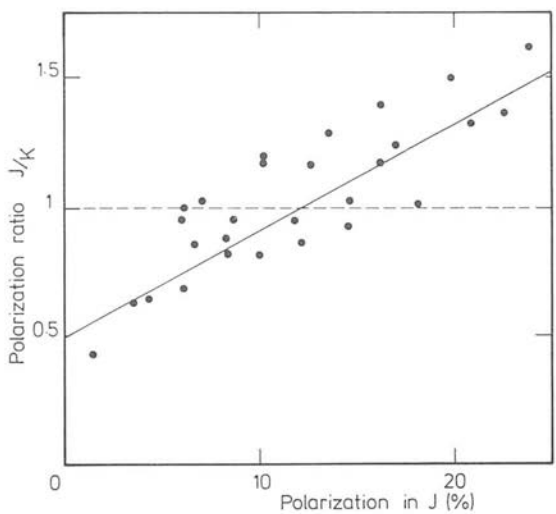


Figure 3. Polarization ratio in J and K wavebands plotted against J polarization. The ensemble of points is clearly inconsistent with canonical assumption of wavelength independence $r_s = 0.811$ $N = 28$. $\chi^2 = 29.5$, χ^2 [$p \neq p(\lambda)$] = 169, Pearson coefficient $r = 0.849$. The solid line is the best-fit line defined by minimum χ^2, which is clearly inconsistent with wavelength-independent polarization.

4.4.3 $\theta(\lambda)$

Most noticeable in both our data and BHA data is the lack of significant rotation with wavelength $[\Delta\theta(\lambda) = 0]$ which we might have expected to accompany the strong $p(\lambda)$–p effect.

The unusual $p(\lambda)$ and $\theta(\lambda)$ seen in OJ 287 are explained in terms of a two component model. Each component is polarized and has a power-law slope, and they differ only in spectral index and position angle. However, a general property of such a model is $\theta(\lambda)$, which is not usually observed. One way to circumvent this problem, while retaining the simplicity of the model, is to consider a similar two-component model, where the polarizations differ, but where the position-angle difference is very small.

Such a model would result in $p(\lambda)$ (see Fig. 4) and a wavelength-independent position angle. We define ν_p as the frequency at which the polarized fluxes are equal for each component. ν_1 is the frequency at which component 1 is totally dominant.

Furthermore, brightening/dimming of one component relative to another would shift ν_p, so for observation frequencies centred on $\nu_0 < \nu_p$, we would see the $p(\lambda)$–p relation described by BHA.

The $p(\lambda)$–p phenomenon may thus be described in one of three ways. Fig. 3 may be occupied by:

(1) three distinct populations: (a) $p \neq p(\lambda)$, $p \sim 10$ per cent; (b) $dp/d\lambda > 0$, $p < 10$ per cent; (c) $dp/d\lambda < 0$, $p > 10$ per cent;

(2) by the two-component model described by Fig. 4, with a separate population occupying the remaining area (i.e. $dp/d\lambda \gtrless 0$ if $p_2 \gtrless p_1$); or

(3) the OJ 287 type two-component model being general, with polarization minima frequencies (ν_p) that vary from object to object (and also vary with time for any given object).

The $p(\lambda)$–p graph alone is insufficient to distinguish between these several cases. Position-angle data eliminate the OJ 287 model as being general, as $\theta(\lambda)$ is seldom observed.

Multiwavelength observations for several objects will constrain models further, and will ultimately determine whether or not a single general model may be employed to describe the polarization behaviour of blazars as a class.

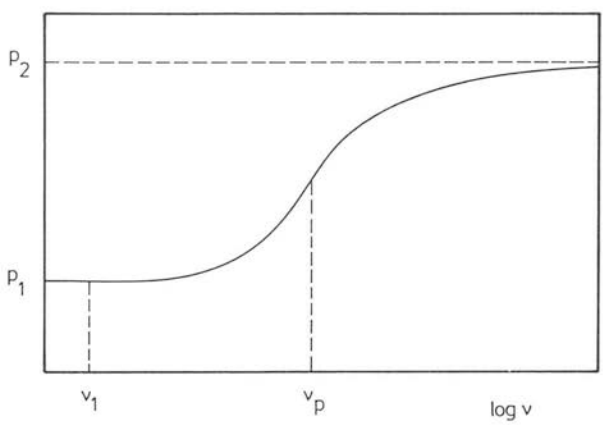


Figure 4. Polarization curve for two components (p_1 and p_2) with similar position angles but differing spectral indices. υ_p is the frequency at which the polarized fluxes are equal.

4.5 INTERNIGHT VARIATIONS

The observations presented in this paper extend the data base of Papers I and II to 125 data pairs. This data, combining internight changes ΔI_{tot}, ΔI_{pol}, in the total and polarized flux in the *J*, *H*, *K* wavebands, has neither slope nor intercept significantly different from zero, a Pearson correlation coefficient of $r = -0.032$ and a Spearman Rank cc of $r_s = 0.168$. The positive correlation between ΔI_{tot} and ΔI_{pol} seen in Papers I and II seems to have vanished. This result, however, is strongly affected by the unusual behaviour of OJ 287 in 1983 January, when this object showed a marked anti-correlation. We can reject the null hypothesis that those variables are uncorrelated in OJ 287, at the 99 per cent confidence level, although a large χ^2 value reveals a large intrinsic scatter of the data from the best-fit line.

Excluding this OJ 287 data from the analysis, we find a Spearman coefficient of $r_s = 0.473$ from 110 data pairs, and thus reject the null hypothesis at the 99.95 per cent confidence level (see Fig. 5). In Papers I and II, only OJ 287 and 0735+178 show a definite correlation between ΔI_{tot} and ΔI_{pol}. The 1983 January variations for 0735 + 178 show no such correlation, neither is there sufficient data on any other object to reveal a similar correlation.

What can cause the different types of internight variation? The shaded area of Fig. 5 will be populated by objects whose variability is caused by a polarized flux ($p \leqslant 75$ per cent) 'turning on or off' e.g. the polarized sub-units model of Blandford & Rees (1978) (see Paper I). To populate the unshaded area requires a more sophisticated model, for example, a change in the degree of alignment of the magnetic field, or two or more polarized components rotating relative to one another while varying in flux.

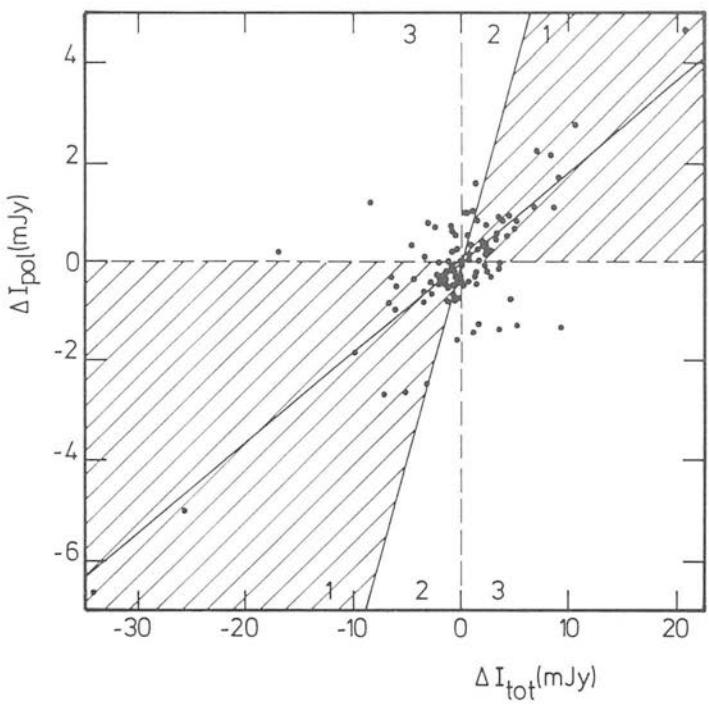


Figure 5. Internight variations in total and polarized flux, except for OJ 287 (1983 January) and BL Lac (1981 May). BL Lac has been omitted for convenience of graph presentation, but is included in the data analysis. Only the shaded area may be populated by objects that vary by 'turning on/off' polarized flux.

Högbom (1979) has shown that all astrophysical fields are essentially tangled. However, Laing (1980) suggests that these fields may be compressed or stretched to form 'local alignment' along a line-of-sight, to produce regions of high polarization. It is therefore possible to populate area 2 (Fig. 5) by the increased flux in a jet, for example, convecting the field to increased alignment. Area 3 may be populated by (1) an OJ 287-type two-component model, or (2) tangling/untangling a magnetic field. Such a field, given a twist, will decrease alignment and so decrease polarization degree, while increasing the local field density and therefore, total flux.

The data of Fig. 5 has a best-fit line yielding $\chi^2/N = 2.7$, and a Spearman coefficient of $r_s = 0.473$, indicating a correlation at the 99.95 per cent confidence level, but with significant scatter of the data from the best-fit line. Of these 110 data pairs, only eight lie beyond 2σ of the unshaded area. The data are thus consistent with internight variations produced by sub-units polarized at ~ 18 per cent.

Detailed internight polarization behaviour can only be evaluated for objects with large ($\Delta m \gtrsim 0.5$ mag) internight flux variations. 0829 + 046 brightened by $\Delta m \sim 0.5$ mag from January 10th to 11th, with a corresponding increase in polarization by a factor of 2. This can be understood in terms of a synchrotron source turning on, with $p \sim 30$ per cent $\theta \sim 109°$, and spectral index $\alpha = 0.86$ [*cf.* α_{JK} (January 10th) = 1.25] (see Fig. 6).

OJ 287 brightened by $\Delta m \sim 0.25$ mag from January 7th to 8th. The increase in flux is a power-law spectrum of index $\alpha = 1.6$, over the whole spectral range $U(0.365\,\mu m)$ to $K(2.2\,\mu m)$ (Holmes *et al.* 1984). The other internight flux changes are consistent with this slope.

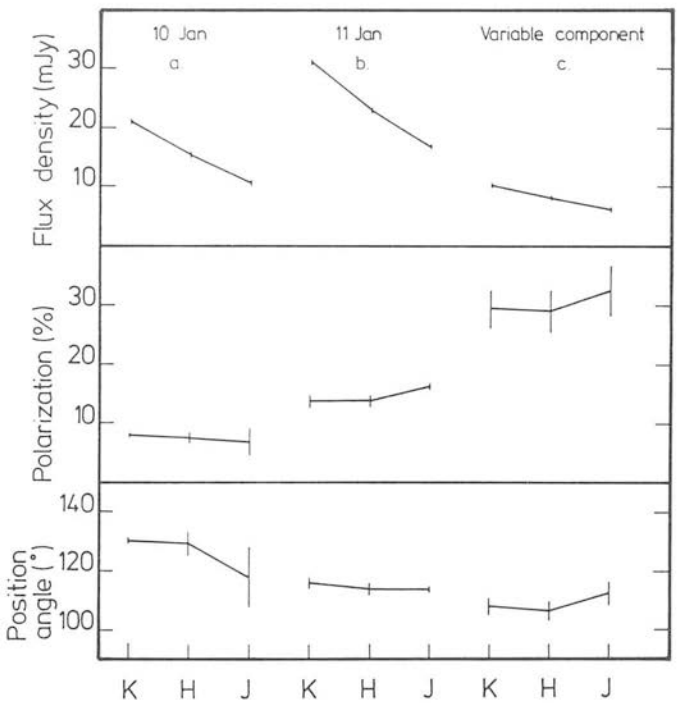


Figure 6. Polarization properties of 0829+046 on 1983 January 10th and 11th and of the variable component.

5 Conclusions

The most recent results of the BL Lac object monitoring programme are presented. The most important results emerging from our infrared observations may be summarized as follows:

(1) Five of the most luminous sources violate the luminosity–variability criterion, implying non-spherical accretion and/or relativistic beaming.

(2) Internight variations show that the variability of BL Lacs is consistent with a model consisting of two or more variable, polarized components.

(3) The intriguing $p(\lambda)$–p result, together with position-angle data, allow us to probe the magnetic-field structure of the emitting region.

(4) The range of position angles seen in each source shows a possible correlation with maximum infrared luminosity, which would link luminosity with a geometrical parameter.

(5) The correlation of spectral index with luminosity and/or redshift is indicative either of blazar evolution or of spectral curvature.

All our observations to date show that BL Lac objects as a class may be well described by a simple two-component model where relativistic beaming is important.

Acknowledgments

We thank Terry Lee and the UKIRT staff for excellent support. In particular we thank John Clark and Malcolm Stewart who contributed to the hardware and software for the polarimeter respectively. PAH acknowledges receipt of the Robert Cormack Bequest Fellowship, and part of this work was funded by the SERC. Our thanks are also due to Mel Dyck for the use of his Lyot depolarizer.

References

Angel, J. R. P., Boronson, T. A., Adams, M. T., Duerr, R. E., Giampapa, M. S., Gresham, M. S., Gural, P. S., Hubbard, E. N., Kopriva, D. A., Moore, R. L., Peterson, B. M., Schmidt, G. D., Turnshek, D. A., Wilkerson, M. S., Zotov, N. V., Maza, J. & Kinman, T. D., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 117, ed. Wolfe, A. M., University of Pittsburgh.

Angel, J. R. P. & Stockman, H. S., 1980. *Ann. Rev. Astr. Astrophys.*, **18**, 321.

Bailey, J., Cunningham, E. C., Hough, J. H. & Axon, D. J., 1981. *Mon. Not. R. astr. Soc.*, **197**, 627.

Bailey, J., Hough, J. H. & Axon, D. J., 1983. *Mon. Not. R. astr. Soc.*, **203**, 339.

Björnsson, C-I. & Blumenthal, G. R., 1982. *Astrophys. J.*, **259**, 805.

Blandford, R. D. & Rees, M. J., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 328, ed. Wolfe, A. M., University of Pittsburgh.

Bregman, J. N., Glassgold, A. E., Huggins, P. J., Aller, H. D., Aller, M. F., Hodge, P. E., Rieke, G. H., Lebofsky, M. J., Pollock, J. T., Pica, A. J., Leacock, R. J., Smith, A. G., Webb, J., Balonek, T. J., Dent, W. A., O'Dea, C. P., Ku, W. H-M., Schwartz, D. A., Miller, J. S., Rudy, R. J. & Levan, P. D., 1984. *Astrophys. J.*, **276**, 454.

Burstein, D. & Heiles, C., 1982. *Astrophys. J.*, **87**, 1165.

Carswell, R. F., Strittmatter, P. A., Williams, R. E., Kinman, T. D. & Serkowski, K., 1974. *Astrophys. J.*, **190**, L101.

Craine, E. R., Duerr, R. & Tapia, S., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 99, ed. Wolfe, A. M., University of Pittsburgh.

Elliot, J. L. & Shapiro, S. L., 1974. *Astrophys. J.*, **192**, L3.

Fabian, A. C. & Rees, M. J., 1978. *IAU/COSPAR Symposium on X-ray Astronomy*, eds Baity, W. A. & Peterson, L. E., Pergamon Press, Oxford.

Glassgold, A. E., Bregman, J. N., Huggins, P. J., Kinney, A. L., Pica, A. J., Pollock, J. T., Leacock, R. J., Smith, A. G., Webb, J. R., Wisniewski, W. Z., Jeske, N., Spinrad, H., Henry, R. B. C., Miller, J. S., Impey, C. D., Neugebauer, G., Aller, M. F., Aller, H. D., Hodge, P. E., Balonek, T. J., Dent, W. A. & O'Dea, C. P., 1983. *Astrophys. J.*, **274**, 101.

Högbom, J. A., 1979. *Astr. Astrophys. Suppl.*, **36**, 173.
Holmes, P. A., Brand, P. W. J. L., Impey, C. D., Williams, P. M., Smith, P., Elston, R., Balonek, T., Zeilik, M., Burns, J., Heckert, P., Barvamis, R., Kenney, J., Schmidt, G. & Puschell, J., 1984. *Mon. Not. R. astr. Soc.*, in press.
Impey, C. D., Brand, P. W. J. L. & Tapia, S., 1982. *Mon. Not. R. astr. Soc.*, **198**, 1.
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1982. *Mon. Not. R. astr. Soc.*, **200**, 19 (Paper I).
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1984. *Mon. Not. R. astr. Soc.*, **209**, 245 (Paper II).
Knacke, R. F., Capps, R. W. & Johns, M., 1976. *Astrophys. J.*, **219**, L69.
Laing, R. A., 1980. *Mon. Not. R. astr. Soc.*, **193**, 439.
Macklin, J. T., 1982. *Mon. Not. R. astr. Soc.*, **199**, 1119.
Miller, J. S., 1978. *Comm. Astrophys.*, **7**, 175.
Moore, R. L., Angel, J. R. P., Rieke, G. H., Lebofsky, M. J., Wisniewski, W. Z., Mufson, S. L., Vrba, F. L., Miller, H. R., McGimsey, B. Q. & Williamson, R. M., 1980. *Astrophys. J.*, **235**, 717.
Moore, R. L. & Stockman, H. S., 1983. *Standard Observatory Preprint No. 476.*
Nordsieck, K. H., 1976. *Astrophys. J.*, **209**, 653.
Pettini, M., Hunstead, R. W., Murdoch, H. S. & Blades, J. C., 1983. *Astrophys. J.*, **273**, 436.
Puschell, J. J. & Stein, W. A., 1980. *Astrophys. J.*, **237**, 331.
Savage, B. D. & Mathis, J. S., 1979. *Ann. Rev. Astr. Astrophys.*, **17**, 73.
Schmidt, M., 1974. *Astrophys. J.*, **193**, 505.
Stein, W. A., O'Dell, S. L. & Strittmatter, P. A., 1976. *Ann. Rev. Astr. Astrophys.*, **14**, 173.
Zekl, H., Klare, G. & Appenzeller, I., 1981. *Astr. Astrophys.*, **103**, 342.

Mon. Not. R. astr. Soc. (1990) 243, 640–655

The optical and infrared emission of blazars

K. R. Ballard, A. R. G. Mead and P. W. J. L. Brand
Department of Astronomy, Edinburgh University, Edinburgh EH9 3HJ

J. H. Hough
Division of Physical Sciences, Hatfield Polytechnic, Hatfield, Herts AL10 9AB

Accepted 1989 October 24. Received 1989 September 7; in original form 1989 June 6

SUMMARY

We have carried out a polarization study of 44 blazars and candidate blazars observed at optical and near-infrared frequencies. The spectrum of the radiation is discussed in terms of shock acceleration models. Frequency dependence is a common feature of the polarization behaviour, with the polarization and spectral index generally increasing towards higher frequencies in those BL Lac objects which were observed. No evidence was found to support the previous claim that frequency dependence is related to high levels of polarization. No characteristic form of blazar variability was evident from the data. The intrinsic frequency-dependent polarization of the BL Lac objects is the result of spectral curvature, due to source inhomogeneity. We suggest that the observed spectral, temporal and polarization behaviour is the result of a polarized cut-off component, tentatively identified with emission from a shock and an unpolarized steep spectrum component. This can explain the observations of 1253−055 (3C 279) and many of the other observed cases of frequency-dependent polarization.

1 INTRODUCTION

In a previous paper (Mead *et al.* 1990; hereafter MEA) we presented the results of a programme to monitor the flux and polarization behaviour of a sample of blazars. The observations were made at the United Kingdom Infrared Telescope during four observing runs (1986 August, 1987 July, 1987 September and 1988 February). The observations, covering the *UBVRIJHK* frequency range were carried out simultaneously at one infrared and all five optical wavebands. A summary of the observations is given in Table 1. These data significantly enlarge the database of infrared polarimetry and photometry established by the observations of Impey *et al.* (1982, 1984), Holmes *et al.* (1984a,b) and Brindle *et al.* (1986). The data were collected with the MKI and MKII Hatfield polarimeters (see MEA for a description of the instruments) and have smaller polarimetric errors than previous data; they allow statistical analysis at unprecedented levels.

This paper will address the interpretation of the observed optical emission of the 44 blazars in terms of specific models of the compact emission region. The term blazar will be taken to include BL Lac objects and also highly polarized quasars which exhibit strong emission-line spectra. The optical continua of both can usually be regarded together, since both contain synchrotron emission from relativistic charged particles. A full treatment of this process is given by, for example, Pacholczyk (1970, 1977). The canonical synchrotron model, which has had considerable success in explaining the emission of extended radio sources, consists of a uniform magnetic field, a power-law distribution of electron energies and an isotropic distribution of electron velocities. This predicts a frequency-independent form of the polarization. It is well established (see Puschell & Stein 1980; Holmes *et al.* 1984a; Brindle *et al.* 1986; Smith *et al.* 1987 etc.) that frequency dependence of both polarization and position angle is a common occurrence among the BL Lac objects, dominated by synchrotron emission. These properties of blazar emission indicate that a more complicated, inhomogeneous model for the emission region is required. The polarization of inhomogeneous sources has been the subject of works by Nordsieck (1976), Björnsson & Blumenthal (1982) and Björnsson (1985); their results will be applied in detail in this paper.

The frequency distribution of the radiation has received much attention. Due to the fact that variability is observed at all wavelengths, accurate multifrequency spectra can be obtained only through simultaneous observations. At radio frequencies blazars generally have flat spectra, and the optically thick synchrotron spectral index of −2.5 is not seen – the 'Cosmic Conspiracy' of Cotton *et al.* (1980). At higher frequencies the spectrum is steep ($\alpha \sim 1.0$) and is often of a

Table 1. Summary of the polarimetric observations at UKIRT.

IAU Name	1986 Jul./Aug.	1987 Jul.	1987 Sep.	1988 Feb.
0048 − 097	HJIRVB	KHJIRVBU	HIRVBU	
0106 + 013	HIB	HIRVBU	HIRVBU	
0109 + 224	HJIRVB	HIRVBU	HIRVBU	
0118 − 272	HJIRVB	HIRVBU	HIRVBU	
0138 − 097	HIB	HIRVBU	KHIRVBU	
0219 − 164		KHJIRVBU		
0219 + 428	HIB	HIRVBU	HIRVBU	
0235 + 164		KHJIRVBU	KHJIRVBU	
0300 + 470			HIRVBU	
0323 + 022	HIB		HIRVBU	
0336 − 019	HIB			
0338 − 214	HIB		HIRVBU	
0414 + 009			HIRVBU	HJIRVBU
0422 + 004				KHIRVBU
0735 + 178			HIRVBU	HIRVBU
0736 + 017			HIRVBU	JIRVBU
0754 + 100				KHJIRVBU
0818 − 128				KHJIRVBU
0851 + 202				KHIRVBU
0906 + 015				HIRVBU
1101 + 384				KHJIRVBU
1147 + 245				HJIRVBU
1156 + 295				HIRVBU
1253 − 055	HJIRVBU	KHJIRVBU		HIRVBU
1413 + 135	HIB			
1418 + 546	HIB	HJIRVBU	HIRVBU	KHJIRVBU
1424 + 240				HIRVBU
1510 − 089	HIB			
1514 − 241	HIB	KHJIRVBU		
1538 + 149	HIB			
1641 + 399	KHJIRVBU	KHJIRVBU	HIRVBU	
1652 + 398			KHJIRVBU	
1717 + 178	HIB	HIRVBU	HIRVBU	
1727 + 502	HIB		HIRVBU	
1749 + 096	HIB	JIRVBU		
1921 − 293	HIB	JIRVB		
2032 + 107		JIRVBU	HIRVBU	
2155 − 304	HJIRVB	KHJIRVBU	KHJIRVBU	
2200 + 420	HJIRVB	KHJIRVBU	HIRVBU	
2208 − 137	HIB			
2223 − 052	HIB	HJIRVBU	HIRVBU	
2230 + 114	HIB			
2251 + 158	HIB	HIRVBU	HIRVBU	
2254 + 074	HJIRVBU	HIRVBU	HIRVBU	

power-law form (Allen, Ward & Hyland 1982). There has been some discussion as to whether the spectrum is best described by a power law with breaks (e.g. Gear *et al.* 1985, who have breaks in their 1 μm–2 mm spectra at 10 μm). Breaks at 3×10^{14} Hz are suggested by Cruz-Gonzalez & Huchra (1984). In contrast Landau *et al.* (1986) fit all their (1400 Å–20 cm) multifrequency spectra with parabolæ. There is some observational evidence for a cut-off to the infrared/optical spectrum (e.g. Rieke, Lebofsky & Kinman 1979; Bregman *et al.* 1981). Evidently, the origin of the non-thermal distribution of particle energies is a question of crucial importance. A rough estimate of the magnetic field ($B\sim10^{-5}$ *T*; Gear *et al.* 1985) yields electron lifetimes of the order of days at optical frequencies, implying that *in situ* acceleration of electrons must take place. To date, the most commonly advanced theory is first-order Fermi acceleration at magneto-hydrodynamic shocks. The basic process is outlined by Longair (1981). At a non-relativistic strong shock front this process yields a spectral index of $\alpha=0.5$. The spectral index can steepen by 0.5 when electron losses become important (Kardashev 1962) and a cut-off is present at a frequency where the synchrotron loss time becomes equal to the acceleration time (Blandford 1979). Meisenheimer & Heavens (1986) use shock acceleration to model multi-frequency spectra of 3C273. The arguments for relativistic motion in blazars were reviewed by Impey (1987). On the basis that many superluminal sources are blazars, he argues for a connection between superluminal motion and the blazar phenomenon, which suggests that if the acceleration is due to a shock then it is relativistic.

Flux variability can be an important tool in understanding the emission region. For example, Marscher & Gear (1985) showed that the millimetric behaviour of 3C273 could be understood by a simple model of a shock passing along a jet close to the line-of-sight, characterizing the evolution in terms of the energy loss mechanism (Compton, synchrotron and 'adiabatic'). On the basis of a week of monitoring of BL Lac, Moore *et al.* (1982) advocated a model of randomly oriented sub-components turning on at a rate of about ten per day and decaying over about five days. Brindle *et al.* (1985), reporting observations of the same object when fainter, proposed a fixed underlying component with superimposed sub-components.

Alternative models for the emission region are summarized by Begelman, Blandford & Rees (1984); these include the possibility that the emission may be associated with orbiting hotspots on an accretion disc. It should be noted that thick accretion discs, like relativistic flows, may be able to radiate at super-Eddington luminosities (Abramowicz, Calvani & Nobili 1980). However, given the uncertainty as to whether such discs are stable (Rees 1984), these ideas will not be developed here.

In Section 2 we will discuss the possibility that the emission from several objects may be contaminated by unpolarized emission, in particular the 'blue bump' emission reported for several quasars. In Section 3 we discuss the properties of the data; this has been partitioned into three subsections, discussing the spectral index, the frequency dependence of the polarization and the variability, each of which is a theoretically separable area. Section 4 presents a model to explain the observed phenomena of spectral curvature and frequency dependent polarizations. Section 5 presents the conclusions of the paper.

Following Björnsson (1985), frequency dependence of the degree of polarization will be referred to as FDP and frequency dependence of the polarization position angle as FDθ. As with MEA the shape of the spectrum will be characterized by use of the (local) spectral index

$$\alpha(\nu) = -\frac{d\log S_\nu(\nu)}{d\log\nu}. \qquad (1)$$

A frequency-dependent spectral index is referred to as spectral curvature.

2 CONTAMINATION BY UNPOLARIZED COMPONENTS

As was first noted by Shields (1978) several notable highly polarized quasars have excess flux at optical and ultraviolet frequencies, which takes the form of contamination of the

Table 2. Objects possibly containing contaminating flux.

Object Name		Date of observation	Significance Level (%)
0106 + 013	(2)	1987 July 28	2.5
0109 + 224	(1)	1986 August 4	1.0
0118 − 272	(2)	1987 July 27	0.01
0138 − 097	(1)	1986 August 5	1.0
0735 + 178	(2)	1988 February 15	5.0
0754 + 100	(1)	1988 February 18	2.5
0818 − 128	(2)	1988 February 15,16	0.01
1418 + 546	(1), (2)	1986 August 6	0.01
1418 + 546	(2)	1988 February 18	0.01
1424 + 240	(2)	1988 February 16	0.01
1717 + 178	(1)	1986 August 7	0.01
1749 + 096	(2)	1986 August 5	0.01
1921 − 293	(1)	1986 August 6	0.1
1921 − 293	(2)	1987 July 27	0.01
2032 + 107	(2)	1987 July 27	0.01
2208 − 137	(2)	1986 August 3,4 & 7	0.01
2251 + 158	(1)	1987 September 20	3.0

(1) ≡ FDP with $dp/d\nu < 0$, (2) ≡ concave flux spectra ($d\alpha/d\nu < 0$).

optical emission by unpolarized flux. This results in either FDP with $dp/d\nu < 0$ or spectral curvature with $d\alpha/d\nu < 0$. Malkan & Sargent (1982) fitted this 'blue-bump' emission by blackbody radiation, associated with radiating accretion discs in the centres of active galactic nuclei. The observations of 1641 + 399, described in a previous paper (Mead *et al.* 1988), represent the most extreme example of contamination of the synchrotron component seen in the UKIRT data set. There exists the possibility that other data may similarly contain 'non-blazar' emission. In order to select objects where this may have affected the spectra, objects showing either $d\alpha/d\nu < 0$ or not showing monotonic polarization variation were selected, and are listed in Table 2. They represent a very small subset of the observations of blazars in MEA. Several of these showed little or no significant polarization.

The χ^2 levels of significance in Table 2 are quoted for a fit to a frequency-independent polarization (for the FDP cases) and to a power-law flux distribution (for the abnormal spectra cases). An alternative explanation of this behaviour is superposition of misaligned polarized components. This is equivalent to the Holmes *et al.* (1984b) two-component model, where two polarized components, of approximately equal polarized fluxes, produce a polarized signature which decreases to zero at the frequency where the polarized fluxes are equal. This will also produce FDθ, and indeed the FDθ seen in the behaviour of 0109 + 224 on 1986 August 4 and 1418 + 546 on 1988 February 18 could lend weight to this explanation.

One of the important potential causes of contaminating flux is the starlight of the host galaxy. Since this will be expected to peak in the near-IR, such emission would be similar to that observed for blazars, being curved ($d\alpha/d\nu > 0$), very steep and showing FDP with $dp/d\nu > 0$ although largely unpolarized. Consequently care must be taken to ensure that behaviour which will be interpreted as being characteristic of the blazar emission process is not the result of contaminating starlight. This implies that all those blazars known to be located in low-redshift galaxies (0521 − 365, 0548 − 022, 1101 + 384, 1133 + 704, 1514 − 241, 1652 + 398 and 2200 + 420) must be studied with multi-aperture photometry or imaging, in order to separate out the galaxy component (*cf.* Kikuchi & Mikami 1987). Other objects may also exhibit continuum properties affected by their relatively weaker underlying galaxies.

3 THE DATA

Three aspects of the data will be considered in turn. First the shape of the spectrum which can be related to current models of shock acceleration and energy losses will be discussed. The polarization behaviour and the correlations derived from the polarization can provide evidence on magnetic field geometry. Finally, variability will be considered as this can provide limits to the size of the emission region, and can be used to constrain models of the emission region.

There are a number of points about the nature of the data which are studied in this and the following sections. The UKIRT data form an inhomogeneous sample, subject to known (but unquantifiable) selection effects. The strategy adopted was to observe as many blazars as possible. Repeat observations were made for those objects which showed variability, high polarizations, FDP or FDθ. Consequently there are in-built biases in this sample of observations which will necessarily affect the conclusions if each observation were to be treated as an independent point. Even if the observing strategy has avoided these problems, it would still be questionable whether the individual observations were independent. This problem is related to the time-scales associated with variations in the flux and polarization properties. It is known that these properties can be constant over periods of several nights, so that observations over one run would not be independent. To try to avoid the effects of these biases, the following results refer to the median and maximum values of the parameters for each object although this loses some information. In these sections ν_f will refer to the effective frequency of the waveband f, quoted in the observer's frame. The lack of complete redshift information for the blazar sample makes it impossible to transform these measurements to the emission frame.

3.1 Flux distribution

3.1.1 Spectral curvature and the distribution of spectral index

The results presented in this and the following sections are obtained by testing the data against a variety of null hypotheses (e.g. a power-law flux distribution). The fits were achieved by minimizing the χ^2 statistic, the use of which is valid if the errors concerned are distributed normally. The errors on the photometry, however, include systematic uncertainties in the photometric calibration. Consequently, where the data concerned include flux information, the true level of the significance is likely to be different from that calculated in the case of normally distributed errors. The use of the χ^2 statistic is justified on the basis that the aim is to obtain some idea of the relative goodness of fit between

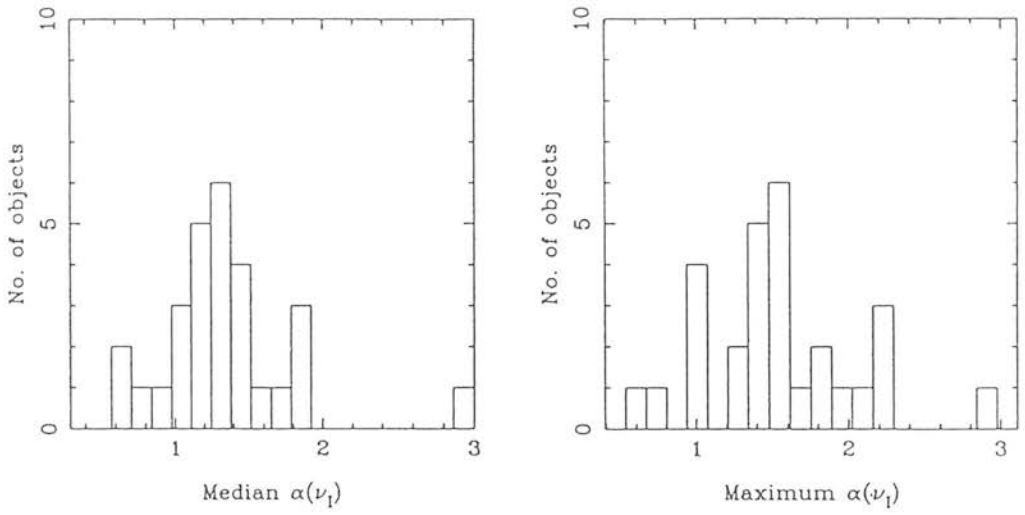


Figure 1. The median and maximum values of the spectral indices at I.

observations rather than to reject a power law fit for an individual object.

The distribution of spectral indices at waveband I is tabulated in Table 3 and shown in Fig. 1. Both median and maximum values for each object are shown. The data used are all the observations in the sample for which polarized observations with calibrated flux data were obtained. These are the objects given in Tables 3 and 4. The fits are separated into polarized and unpolarized samples, because only when polarizations in excess of the 3 per cent limit (used to distinguish highly polarized objects from low polarizations, e.g. Moore & Stockman 1984) are seen, is it certain the

Table 3. Spectral indices at I and spectral curvature parameters.

Object Name	$\alpha(\nu_I)$ Median	$\alpha(\nu_I)$ Maximum	No. of Obs.	$\Delta\alpha_{(B-H)}$ Median	$\Delta\alpha_{(B-H)}$ Maximum	No. of Obs.
0048 − 097	1.32	2.29	9	0.02	0.42	7
0106 + 013	1.36	1.36	1	−0.43	−0.43	1
0109 + 224	1.24	1.55	7	0.42	0.85	5
0118 − 272	1.18	1.98	5	0.00	0.62	4
0138 − 097	1.25	1.45	7	0.45	0.70	5
0219 − 164	1.00	1.00	1	0.35	0.35	1
0219 + 428	1.21	1.44	3	0.01	0.72	3
0235 + 164	1.84	2.23	2	2.44	2.49	2
0414 + 009	1.01	1.03	2	1.09	1.09	2
0422 + 004	1.19	1.22	2	1.25	1.39	2
0735 + 178	1.55	1.62	3	0.41	0.47	3
0754 + 100	1.39	1.51	4	0.40	0.56	4
0818 − 128	1.20	1.52	3	1.29	3.37	3
0851 + 202	1.44	1.47	3	0.31	0.45	3
1101 + 384	0.62	0.62	1	1.03	1.03	1
1147 + 245	1.36	1.36	1	0.27	0.27	1
1156 + 295	1.25	1.29	3	0.06	0.11	3
1253 − 055	1.16	1.62	9	0.42	0.95	8
1418 + 546	1.15	1.59	4	0.80	1.28	4
1424 + 240	0.76	0.95	3	0.45	1.04	3
1514 − 241	0.99	0.99	1	1.99	1.99	1
1641 + 399	1.87	2.07	6	−1.46	−0.62	6
1717 + 178	1.83	1.83	1	–	–	–
1749 + 096	2.89	2.89	1	0.28	0.28	1
2155 − 304	0.62	0.71	2	0.02	0.42	2
2200 + 420	1.38	1.52	5	2.73	2.87	5
2223 − 052	1.71	1.78	3	0.28	0.50	3
2254 + 074	0.92	2.23	5	2.39	2.69	4

Table 4. The polarized flux spectral indices.

Name	$\alpha_p(\nu_I)$ Median	$\alpha_p(\nu_I)$ Maximum	No. of Obs.	$\Delta\alpha_{p,(B-H)}$ Median	$\Delta\alpha_{p,(B-H)}$ Maximum	No. of Obs.
0048 − 097	1.16	1.27	9	1.28	1.68	7
0106 + 013	0.64	0.64	1	−2.67	−2.67	1
0109 + 224	1.40	1.59	5	−0.03	−0.03	4
0118 − 272	1.06	1.11	5	−0.39	1.08	4
0138 − 097	1.17	2.08	7	0.33	0.76	5
0219 − 164	1.07	1.07	1	0.04	0.04	1
0219 + 428	1.19	1.38	5	0.35	0.36	3
0235 + 164	2.32	3.41	2	1.60	2.50	2
0338 − 214	2.03	2.03	1	–	–	–
0414 + 009	1.71	2.19	2	−0.71	−0.38	2
0735 + 178	1.50	1.50	1	0.74	0.74	1
0754 + 100	1.51	1.55	2	0.05	0.92	2
0818 − 128	1.63	1.92	3	1.32	3.38	3
0851 + 202	1.36	1.53	2	−0.24	0.53	2
1101 + 384	0.92	0.92	1	0.44	0.44	1
1156 + 295	1.28	1.40	3	−0.08	0.34	3
1253 − 055	1.59	1.78	3	0.43	1.11	2
1418 + 546	1.35	1.61	3	0.43	0.82	2
1424 + 240	1.03	1.05	3	0.64	1.03	3
1514 − 241	1.79	1.79	1	1.33	1.33	1
1641 + 399	1.76	2.29	7	0.06	0.46	6
1717 + 178	1.62	1.62	1	–	–	–
1727 + 502	3.45	3.45	1	–	–	–
1749 + 096	1.77	2.33	4	1.86	1.86	1
1921 − 293	2.57	2.77	3	–	–	–
2200 + 420	2.66	2.66	1	–	–	–
2223 − 052	1.66	1.43	3	1.96	2.00	3
2254 + 074	1.68	1.96	6	0.59	0.93	3

observed flux contains a synchrotron contribution. If the spectrum is fitted by a power law at the 5 per cent level, then the value of the frequency-independent spectral index is used. Otherwise, a parabola in the $\log S_\nu(\nu) - \log \nu$ plane is fitted, and is used to derive the spectral index at I.

$$\log S_\nu(\nu) = \log S_{14} - a \log \nu_{14} + b(\log \nu_{14})^2. \quad (2)$$

S_{14} is the flux at 10^{14} Hz. There is no theoretical justification for choosing such a form, it is simply an empirical fit to the data. There is a problem with those observations made with the Mark I Hatfield polarimeter at H, I and B only. Three-point data sets are fitted with zero degrees-of-freedom for a three-parameter fit, and as the data are subject to observational errors the fitted parabolae may be quite different from the true flux distribution. Consequently, if a power-law fit to a three-point data set was rejected, then no information from this data set was included in Fig. 1. Fig. 2 shows the degree of spectral curvature for all the UKIRT data, again showing the median and maximum values for each object. The statistic used, $\Delta\alpha_{(B-H)}$, is the difference in the spectral indices at B and H,

$$\Delta\alpha_{(B-H)} = \alpha(\nu_B) - \alpha(\nu_H). \quad (3)$$

This quantity is estimated using the aforementioned parabolic fits.

3.1.2 *The distribution of polarized flux*

A partial indication of whether the range of spectral indices given in the previous section is truly representative of the synchrotron components can be obtained from polarized flux distributions. The polarized spectral indices α_p and the degree of curvature $\Delta\alpha_{p(B-H)}$ (analogous to $\Delta\alpha_{(B-H)}$ in the previous section) are shown in Table 4. The distributions of these quantities are shown in Figs 3 and 4. These are calculated in an exactly analogous way to the similar quantities of the previous section. The spectral parameters so derived represent the behaviour of the synchrotron component, as no other sources of polarized flux are thought to be present. Only in the case of frequency-independent polarization will the polarized flux spectral index be equal to the flux spectral index. For an object displaying convex curvature in its total flux spectrum, the polarized flux spectral index $\alpha_p(\nu)$ will be marginally smaller than $\alpha(\nu)$.

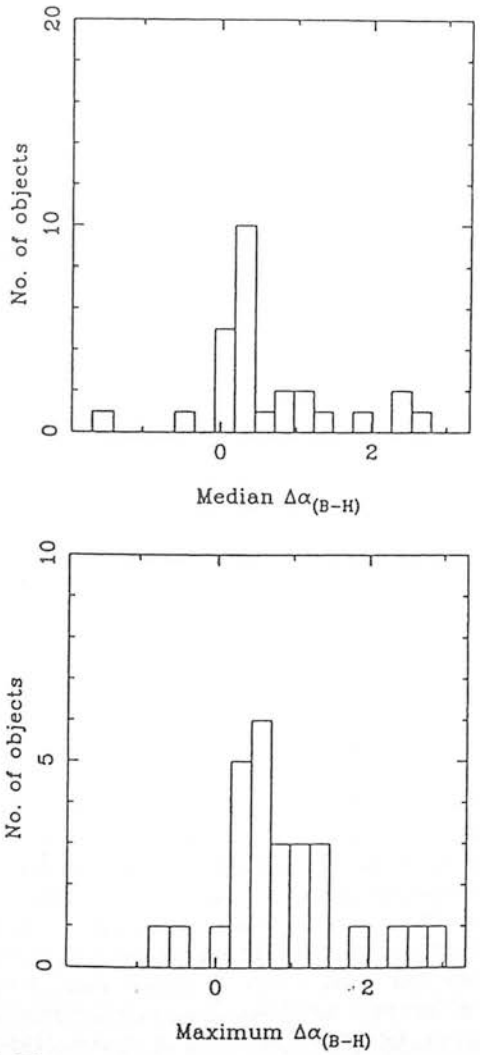


Figure 2. The median and maximum values of the spectral curvature, $\Delta\alpha_{(B-H)}$.

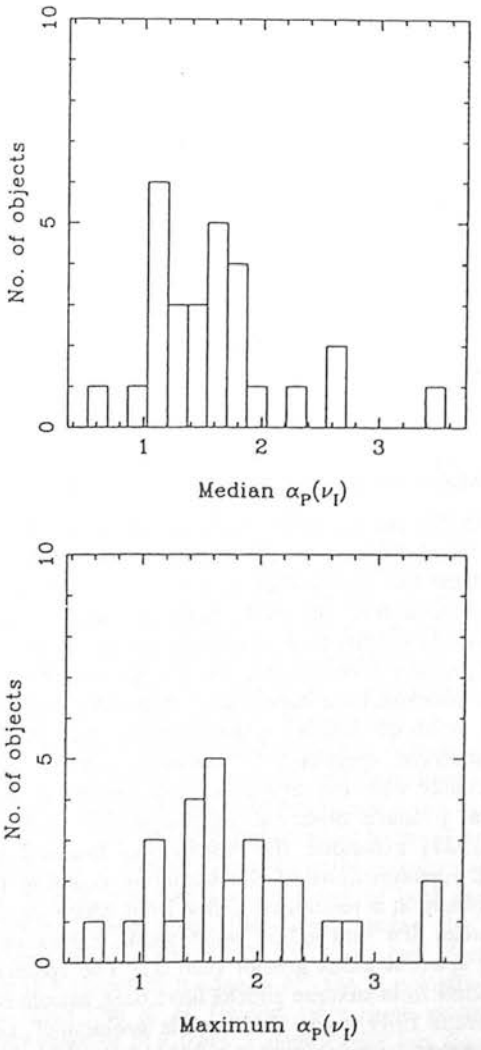


Figure 3. The median and maximum values of the polarized spectral indices at I.

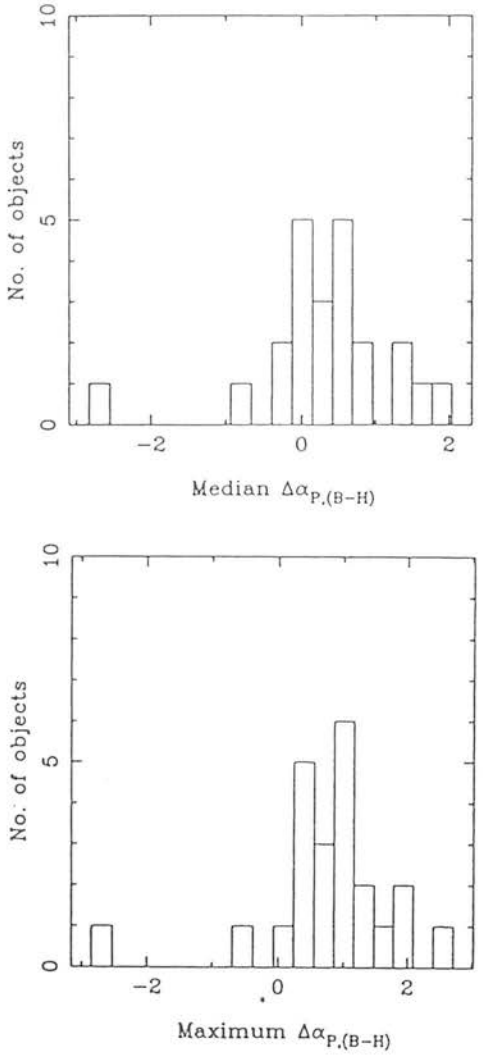


Figure 4. The median and maximum values of the polarized spectral curvature, $\Delta\alpha_{P,(B-H)}$.

3.1.3 *The interpretation of the spectral flux distribution*

It is useful to discuss the implications of the observations described in the above sections, although to a certain extent a discussion of the flux distribution cannot be held independently of a discussion of any FDP, as this constrains most explanations of the origin of the continuum flux properties of blazars. The primary candidate for the process whereby the electrons (or positrons) are accelerated, first-order particle acceleration at shock fronts, makes specific predictions about the observed spectrum. In summary, strong non-relativistic shocks with the magnetic field parallel to the shock normal produce observed spectra of 0.5. Kirk & Schneider (1987) expanded the distribution function in terms of the eigenfunctions of the scattering equation to solve the problem at a relativistic shock front which yields spectra between 0.4 and 0.65, while weak shocks can produce any spectral index greater than 0.5. The spectral indices expected from oblique shocks have been calculated (Kirk & Heavens 1989) when the shock is 'subluminal', i.e. the magnetic field lines intersect the shock front at a velocity less than c. For shock speeds $v \geq 0.1c$ they find spectral indices harder than in the parallel shock case, tending to an unphysical value of $\alpha \sim 0$ as v increases. These spectral indices should be constant over many decades of frequency. In the high-frequency domain, synchrotron energy losses cause the observed spectral index to increase by 0.5 (Kardashev 1962), with a high-energy cut-off beyond which the Fermi process cannot accelerate electrons (Blandford 1979).

Weak shocks can explain any spectral index, but such shocks are generally rejected for two reasons. First, the observed spectral index is a strong function of shock speed; this implies that even the large range of spectral indices shown in Fig. 1 corresponds to a narrow range of shock speeds, and it is unlikely that the range of speeds should be small. Second, strong shocks are expected to be a common feature of the hydrodynamic flows which are thought to be the origin of the blazar emission. As these amplify the emissivity of the fluid, by accelerating particles and compressing the magnetic field, it may be reasonable to expect such shocks to play an important part in the origin of the observed emission. As most blazar identifications are made on the basis of their optical emission, this objection can be overcome if selection effects are important; there is an obvious bias against very steep optical spectra. To resolve this question a complete survey of a sample of compact radio sources is needed.

If arbitrarily weak shocks are rejected, then the range of spectral indices observed is not consistent with the power-law spectra predicted by Fermi acceleration at strong shocks. Spectral indices of the order of 3 are too large.

The high-energy cut-off inherent in acceleration models implies a steepening of the spectrum over a relatively small frequency interval, where the models have steep spectra and spectral curvature but not steep ($\alpha > 1.2$) *power-law* spectra. Such curvature is observed but is not limited to sources with the steepest spectra. For example the spectral index (at I) of the observation of 0818 − 128 is high (1.52) but not extreme, but the curvature ($\Delta\alpha_{(B-H)} = 3.37$) is the highest seen. Checking for a correlation (using the Spearman rank correlation statistic) between either the median or maximum spectral indices and curvature parameters produced no result which was significant even at the 10 per cent level.

Heavens (1988) considered advection of electrons away from a shock front where the magnetic field decays arbitrarily with distance from the shock. He showed that the resulting spectrum may have spectral indices higher than those predicted by the simple shock theory and gradually curved though approximately described by a power law for single decades of frequency if these are well below the cut-off. This is not the only inhomogeneous source model which produces spectral indices much higher than that derived from the electron spectral index. For example, the synchrotron-self-Compton (SSC) model of Ghisellini, Maraschi & Treves (1985) gives rise to steep optical-IR continua as a result of the integration of the luminosity of an inhomogeneous jet. There exist a series of inhomogeneous models, including those of Marscher (1980) and Königl (1981) with steep power-law flux distributions. However, there are a large number of free parameters which go into determining the spectral index. Ghisellini, Maraschi & Treves (1985) attempt fits to the IR/optical continua of two blazars (2155 − 304 and 0537 − 441) using this model by varying

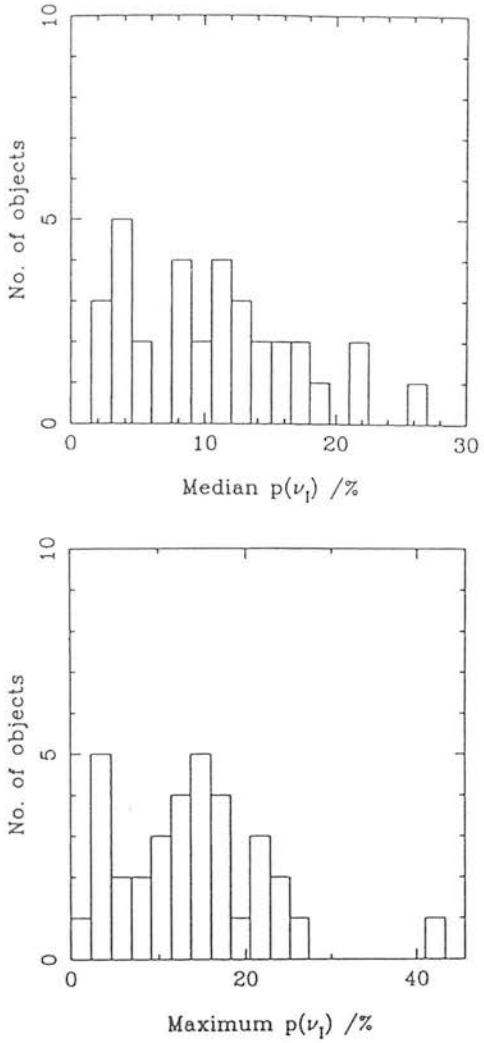


Figure 5. The median and maximum values of the polarizations at I.

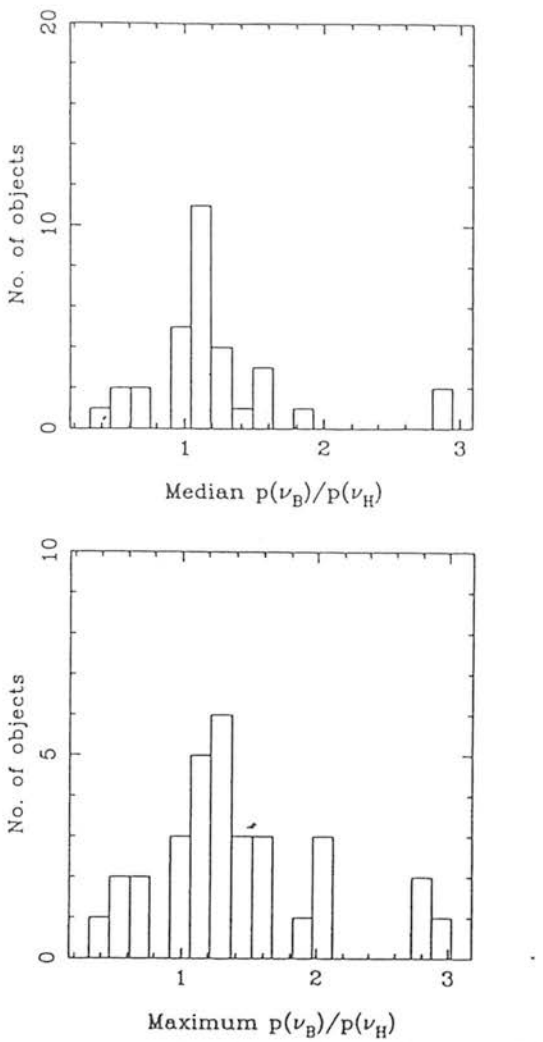


Figure 6. The median and maximum values of the degree of the frequency dependence of polarization, $p(\nu_B)/p(\nu_H)$.

some of these parameters. Fitting such models to the UKIRT data is not possible, given the limited number of frequency points.

We conclude that the spectral indices shown in Fig. 1 are too high to be consistent with a simple model of shock acceleration at a strong shock with a parallel, or oblique, homogeneous magnetic field. The fact that position angles tend to align with the VLBI axis (Impey 1987) suggests the oblique case may be more realistic. Inhomogeneous jet models produce higher spectral indices although, given the restricted frequency range, these cannot be tested fully here. Alternatively a cut-off to the spectrum can produce a greater range of spectral indices. All the data were obtained with simultaneous polarimetry and, as this strongly constrains models of the emission region, we postpone further discussion until after the polarization properties have been considered.

3.2 Analysis of the polarization properties

3.2.1 Frequency dependence of the degree of polarization

All data with two or more polarization measurements were tested for frequency dependence of polarization and polarization position angle. Power law forms [$p(\nu) \propto \nu^{\beta}$] were fitted to the data using the χ^2 statistic to test the goodness of fit. This empirical fit allows an unconstrained minimization to be performed with respect to log p_{14} (p_{14} is the polarization at 10^{14} Hz) and β. Those objects not fitted (at the 5 per cent level) by a constant polarization are described as showing FDP; those not fitted by a power law with constant β at the 5 per cent level are described as having 'complex' polarization behaviour. This procedure will flag behaviour, such as that seen in 2200+420, where the FDP is only significant over a narrow range of the observed frequencies. Fig. 5 shows the distribution of the median and maximum degrees of polarization at I and Fig. 6 the median and maximum changes in polarization over the observed frequency range, expressed in terms of the ratio of the B polarization to the H polarization [$p(\nu_B)/p(\nu_H)$]. The data displayed in these figures are listed in Table 5. Fig. 5 shows a wide range of both median and maximum polarizations, the largest being the series of observations of 1253−055 (3C 279) (see Section 4.2). There is a common tendency for the *B* polarization to be higher than that at *H*. Indeed, only three objects have a median value of $p(\nu_B)/p(\nu_H)$ significantly less than one [the highly polarized quasars 1641+399 (3C 345), 1921−293

Table 5. Polarization data for Figs 5 and 6. Numbers in brackets are 1 σ errors.

Object Name	$p(\nu_I)$/% Median		$p(\nu_I)$/% Maximum		No. of Obs.	$p(\nu_B)/p(\nu_H)$ Median		$p(\nu_B)/p(\nu_H)$ Maximum		No. of Obs.
0048 − 097	14.55	(0.61)	21.17	(1.02)	11	1.24	(0.13)	1.64	(0.17)	10
0109 + 224	8.91	(0.85)	14.04	(0.90)	9	1.09	(0.10)	1.20	(0.05)	7
0118 − 272	16.49	(0.68)	17.82	(0.58)	7	1.17	(0.04)	1.29	(0.08)	7
0138 − 097	22.16	(1.13)	24.80	(1.08)	6	1.09	(0.15)	1.42	(0.05)	8
0219 − 164	12.45	(0.42)	12.45	(0.42)	1	0.93	(0.04)	0.93	(0.04)	1
0219 + 428	11.46	(0.84)	15.46	(0.55)	10	1.08	(0.13)	1.29	(0.06)	10
0235 + 164	13.28	(0.90)	15.52	(1.59)	2	1.52	(0.38)	2.02	(0.24)	2
0300 + 470	9.44	(1.52)	9.44	(1.52)	1	0.69	(0.52)	0.69	(0.52)	1
0323 + 022	3.84	(1.02)	3.84	(1.02)	1	–	–	–	–	–
0338 − 214	10.76	(1.05)	11.07	(1.56)	2	1.00	(0.25)	1.08	(0.13)	2
0414 + 009	7.82	(2.25)	7.82	(2.25)	1	0.63	(0.42)	0.63	(0.42)	1
0422 + 004	16.67	(0.65)	20.79	(1.12)	2	1.08	(0.13)	1.11	(0.05)	2
0735 + 178	16.57	(0.65)	21.90	(0.98)	4	1.07	(0.11)	1.15	(0.06)	4
0754 + 100	11.16	(0.39)	12.39	(0.48)	4	1.10	(0.08)	1.27	(0.03)	4
0818 − 128	23.32	(1.33)	24.73	(1.66)	2	1.29	(0.05)	1.29	(0.05)	1
0851 + 202	18.06	(0.66)	18.36	(1.22)	3	1.07	(0.05)	1.19	(0.06)	3
1101 + 384	3.73	(0.13)	3.73	(0.13)	1	1.45	(0.03)	1.45	(0.03)	1
1147 + 245	2.71	(0.67)	2.71	(0.67)	1	2.82	(0.21)	2.82	(0.21)	1
1156 + 295	26.42	(2.87)	27.27	(2.08)	3	0.91	(0.08)	0.99	(0.05)	3
1253 − 055	31.30	(0.36)	41.58	(0.55)	9	1.08	(0.01)	1.29	(0.06)	9
1418 + 546	4.62	(0.31)	15.37	(2.82)	9	1.51	(0.09)	2.93	(0.16)	7
1424 + 240	4.98	(0.66)	5.14	(0.55)	3	0.91	(0.09)	0.95	(0.09)	3
1514 − 241	3.31	(0.21)	4.76	(0.30)	2	1.53	(0.14)	1.66	(0.07)	2
1641 + 399	15.00	(0.50)	16.09	(1.14)	6	0.48	(0.16)	0.52	(0.15)	6
1652 + 398	1.56	(0.16)	1.56	(0.16)	1	2.85	(0.11)	2.85	0.11	1
1717 + 178	16.11	(1.85)	17.82	(3.62)	2	0.93	(0.50)	1.35	(0.27)	2
1727 + 502	2.48	(0.49)	2.51	(0.82)	2	–	–	–	–	–
1749 + 096	8.89	(0.58)	16.53	(0.95)	4	1.21	(0.21)	1.39	(0.13)	4
1921 − 293	7.56	(1.23)	8.13	(1.65)	2	0.41	(0.79)	0.41	(0.79)	1
2155 − 304	3.10	(0.12)	10.29	(0.23)	5	1.14	(0.06)	1.99	(0.04)	6
2200 + 420	12.02	(0.50)	14.15	(0.37)	13	1.21	(0.06)	1.88	(0.13)	13
2223 − 052	11.64	(1.87)	11.90	(1.21)	3	1.12	(0.23)	1.57	(0.14)	4
2251 + 158	4.18	(0.96)	4.18	(0.96)	1	0.47	(0.29)	0.47	(0.29)	1
2254 + 074	9.65	(1.45)	12.06	(1.26)	11	1.82	(0.16)	2.09	(0.19)	6

and 2251 + 158]. The probable reason for this is contamination of the blazar flux by unpolarized optical components, discussed previously. The data were also tested for FDθ which is much rarer than FDP (see MEA, table 4). There is no evidence suggesting any preference for clockwise over counter-clockwise variations with frequency.

3.2.2 Correlation of frequency dependence with other properties

It has been suggested (Bailey, Hough & Axon 1983; Holmes *et al.* 1984a; Holmes 1985; Brindle *et al.* 1986) that there is a tendency for frequency dependence to be associated with high polarization [the '$p(\lambda)-p$' effect]. The appearance of this effect has been approached in two ways. Fig. 7 shows histograms of the I polarizations wherein objects displaying either FDP (Fig. 7a) or FDθ (Fig. 7b) have been flagged. The distributions of the two samples were then tested using the Kolmogorov–Smirnov test (e.g. Conover 1980) under the null hypothesis that the frequency-independent data had the same distribution as the frequency-dependent data. When considering FDP, the Kolmogorov–Smirnov statistic was $T=0.112$, which corresponds to the 75.2 per cent level of significance. When considering FDθ, the Kolmogorov–Smirnov result was $T=0.150$, which corresponds to the 50.0 per cent level of significance. An alternative is to use non-parametric correlation tests to test directly for a correlation between the amount of frequency dependence and the polarization. Fig. 8 shows plots of $\log[p(\nu_B)/p(\nu_H)]$ versus the I polarization. Whereas all the observations were plotted in Fig. 7, only median or maximum values for each object are shown in Fig. 8. Neither was significant at the 5 per cent level. We conclude that there is no evidence that the amount of frequency dependence is dependent on the degree of polarization.

This data set is larger and covers a wider frequency range than those for which the $p(\lambda)-p$ effect was claimed; thus it is surprising that, if real, this effect does not appear. A possible explanation is provided by a selection effect, given that the polarimetric errors obtained on these data are certainly smaller than those on the data of Bailey, Hough & Axon (1983), Holmes *et al.* (1984a) and Holmes (1985). It seems possible that these authors simply missed examples of FDP (and FDθ) in objects of moderately low polarization that

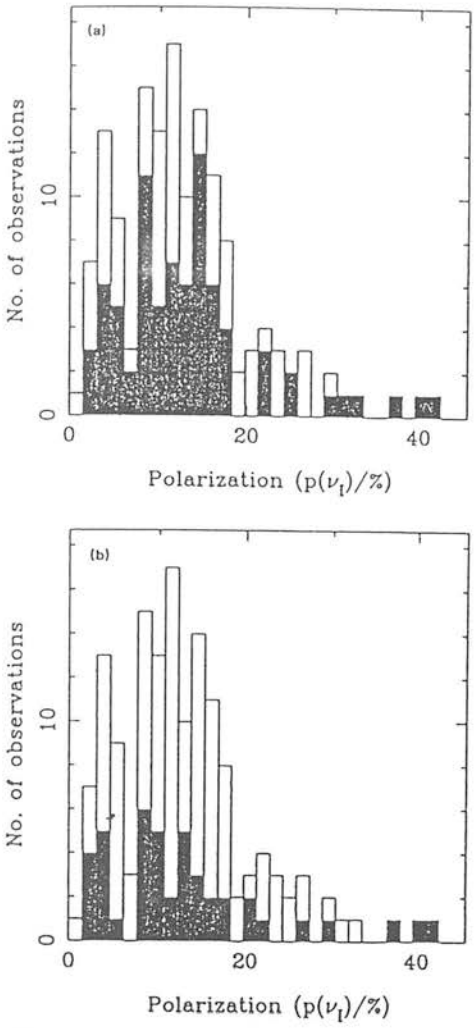


Figure 7. The distributions of the I polarizations. The shaded region represents the distribution of those objects showing FDP (a) and FDθ (b) significant at the 5 per cent level. The distributions are not significantly different.

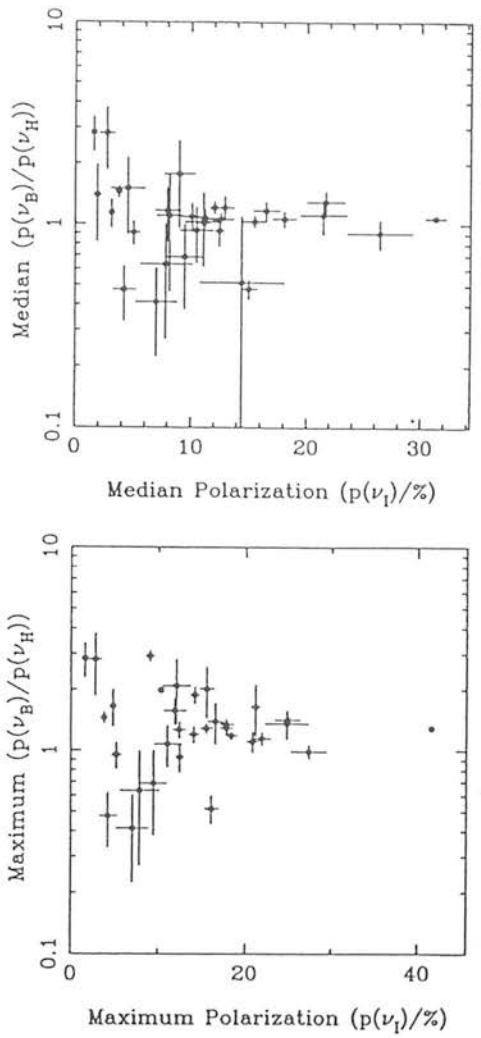


Figure 8. Plots of the median and maximum values of $p(\nu_B)/p(\nu_H)$ versus $p(\nu_I)$. No correlation was present here at the 5 per cent level.

would have been detected in this work. Further, Holmes (1985) notes that, in some of his sample, this correlation is markedly affected by specific blazars observed more frequently than others. This indicates that his result may be susceptible to the kinds of systematic biases discussed earlier.

In Section 3.1.3 cut-offs were mentioned as a possible explanation for the steep and curved spectra observed in blazars. If this were generally so for blazars, then a correlation between the steepness of the spectrum and the occurrence of FDP should be observed. This has been tested using the Spearman rank correlation test statistic for the correlation between $\alpha(\nu_I)$ and $[p(\nu_B)/p(\nu_H)]$. Also tested for were correlations between $\alpha(\nu_I)$ and $p(\nu_I)$, $\Delta\alpha_{(B\text{-}H)}$ and $p(\nu_I)$ and $\Delta\alpha_{(B\text{-}H)}$ and $p(\nu_B)/p(\nu_H)$. No correlations were found at the 5 per cent level.

3.3 Variability

Before discussing variability it is necessary to differentiate between the different time-scales being tested. Data were taken over several nights at four epochs (1986 August, 1987 July, 1987 September and 1988 February). As a consequence there are three inter-epoch and numerous inter-night time-scales to be considered. The χ^2-statistic was used to test the null hypothesis that there was no difference between measurements for any pair of observations. The critical level of significance used was 1.0 per cent, any lower value being regarded as evidence for variability. Again it must be stressed that possible systematic errors may compromise the assumption of normal errors in the photometry.

Inter-epoch variability in the flux densities was seen in the majority of objects, and ranged from 1253−055 and 1749+096 which varied by a factor of almost three, to 0118−272 and 0138−097 which remained constant between 1987 July and September. Night-to-night variations were also commonly seen, though exceptions were more common than above. One object, 2254+074, varied from 1.18 mJy to 0.44 mJy in one day at I, but this was atypical. Three objects (0109+224 in 1987 September, 1156+295 in 1988 February and 1253−055 in 1986 August) varied by up to 33 per cent whilst retaining a constant spectral index. No observations were seen which repeated the pattern of Gear, Robson & Brown (1986), where spectral indices decreased with increasing flux. Three objects showed *steepening* spectral indices with increasing flux (0048−097

and 0109+224, both in 1986 August and 1253−055 in 1988 February).

As with the flux variability, inter-epoch and inter-night variations in the degree of polarization were seen in the majority of blazars. The only example of a polarized object displaying constant inter-epoch polarization degree was 2254+074. In addition seven other objects failed to show any inter-night variations when repeat observations were performed. These were 0118−272, 0138−097, 1156+295, 1418+546, 1424+240, 1641+399 and 2223−052. Changes in the polarization did not always coincide with significant changes in the observed flux of these objects. The variations in the polarization data were not all of a similar form; for example 1253−055 showed a generally increasing degree of polarization with the onset of FDP ($dp/d\nu > 0$) whereas 0109+224 showed (during 1986 August) FDP of both senses ($dp/d\nu < 0$ and $dp/d\nu > 0$).

There is evidence for variability on time-scales of tens of minutes in the polarization position angle in the data of 1253−055 (3C 279) during 1986 August. All these data were collected with the Mark I instrument where the flux and polarization spectrum was constructed from a series of simultaneous measurements at three frequencies. Only on 1986 August 2 did the individual sets show strong evidence for FDθ, on all other nights during this period the data were consistent with frequency-independent position angles but with variations of order half a degree over the approximately ten-minute time-scale of the integrations. Table 6 shows the average position angle for each run during this period. Inspection shows that for any pair of runs the evidence for variability is marginally significant. Nevertheless this was capable of inducing artificial FDθ in the summed data. The data given in MEA are the summed values which are the best estimates of the position angle during the whole series of observations, although the errors are underestimated. There was no evidence for similar rapid changes in the polarization degree nor did any other object show such rapid variability of the position angle.

Table 6. The position angle data for 1253−055 during 1986 August.

U.T. Date	Wavebands	U.T. Time (hours)	Position angle (deg.)	
1986 August 1	HIB	6.165	120.46	(0.35)
	HIB	6.358	120.84	(0.38)
1986 August 2	HIB	6.260	FDθ	
	JVB	6.415	FDθ	
1986 August 4	HIB	5.825	125.40	(0.20)
	HIB	5.958	125.20	(0.20)
	JVB	6.072	126.06	(0.21)
	JRB	6.166	126.77	(0.16)
	JRU	6.313	128.00	(0.20)
1986 August 5	HIB	5.910	132.57	(0.29)
	HIB	5.999	131.87	(0.23)
	JVU	6.154	132.16	(0.20)
	JRU	6.280	131.45	(0.36)
1986 August 6	HIB	5.798	136.23	(0.32)
	HVB	5.997	136.85	(0.42)
	JRU	6.161	135.66	(0.20)

Only one object (0235+164) was observed with a polarization position angle constant over the inter-epoch time-scale. Only two high signal-to-noise observations of this object were made (in 1987 July and 1987 September). One of the more extreme cases was 1749+096 where the position angle varied from 5° to 35° in one day in 1986 August. Only six objects failed to show inter-night variations during any of the epochs. These were 0118−272, 0138−097, 0818−128, 1156+295, 1418+546 and 1424+240. Note that these objects also failed to display any inter-night variations of the degree of polarization. However, the other objects which similarly failed to show inter-night changes in the polarization degree did show inter-night changes in position angle. Indeed, position angle variations were marginally more common than variations in the degree of polarization.

The fact that variability in the position angle is the rule rather than the exception should not be taken as invalidating the results of Rusk & Seaquist (1985) and Impey (1987) who showed that preferred polarization position angles were generally along the milliarcsecond radio-structure axis. Impey (1987) regarded any object where two-thirds of the position angles were consistent to within 40° as having a preferred (range of) position angle. The variations which were detected in the UKIRT data were generally of a few to ten degrees in amplitude, but six objects exhibited variability around or in excess of the 40° limit. These were 0109+224, 0118−272, 0219+428, 1418+546, 1749+096 and 2254+074. Three other objects (0138−097, 1727+502 and 2155−304) showed FDθ which came close to violating the limit. Since the majority of the detected variations were of much smaller amplitude, this data set should not invalidate Impey's result.

4 MODELS OF THE FREQUENCY DEPENDENCE OF POLARIZATION

4.1 Theoretical background

Björnsson & Blumenthal (1982) set out a formalism whereby the polarization properties of an inhomogeneous source can be evaluated. Their result was

$$p(\nu) = \Pi(\nu)\,\frac{\alpha(\nu)+1}{\alpha(\nu)+5/3}, \tag{4}$$

where $\Pi(\nu)$ is composed of integrals over the magnetic field geometry, and represents the degree of ordering of the magnetic field. This provides a description of the properties of an inhomogeneous synchrotron source. There are several possible specific models of the emission region which are summarized below.

Section 3.1.3 discussed the effect of cut-offs in the energy spectrum of the radiating electrons on the total flux spectrum. Such a cut-off would also have a dramatic effect on the observed polarization behaviour. Strong FDP occurs around the cut-off frequency but no FDθ is seen. Fig. 9

shows the expected polarization signature of an *ad hoc* two-component model consisting of a polarized cut-off component and an unpolarized power law component. The cut-off was assumed to be a sharp upper limit in the electron energy distribution. The rise in polarization results from the fact that synchrotron emission from an isotropic electron distribution tends to 100 per cent polarization above the critical frequency, the fall-off in polarization resulting from the dominance of the unpolarized emission at high frequencies.

Björnsson (1985) considers the effect of anisotropic particle velocity distributions on the polarization behaviour of the synchrotron radiation which can result in significant FDP and FDθ. This explanation for the observed FDθ has not been pursued, as the naïve two-component model can adequately explain the observations. Nevertheless relativistic shock acceleration models do produce anisotropic particle distributions (e.g. Kirk & Schneider 1987; Heavens & Drury 1988) and shock acceleration is implicitly assumed in the cut-off models applied in Sections 4.2 and 4.3.

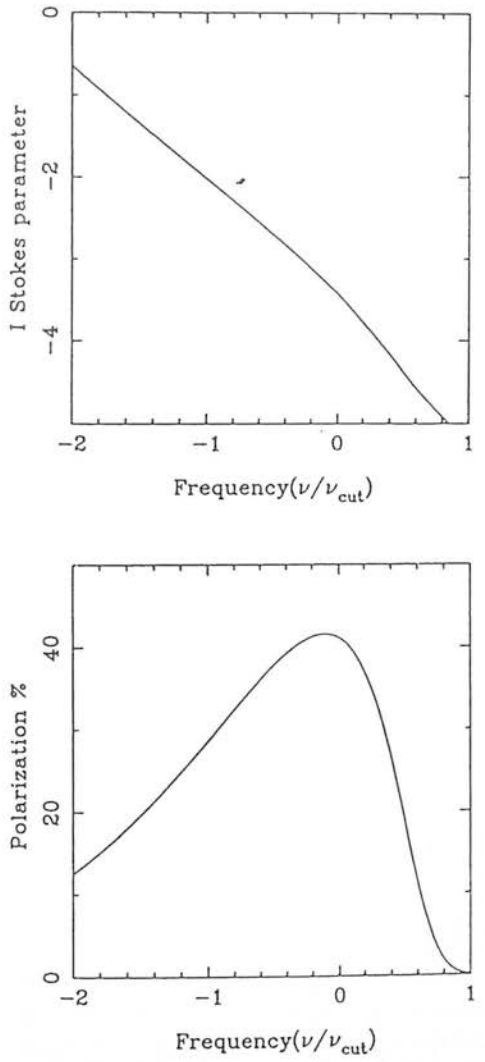


Figure 9. The flux and polarization behaviour of two superimposed synchrotron components. One component has a spectral index of 1.0, a perfectly ordered magnetic field and a sharp upper cut-off to the energy distribution. The second has a spectral index of 1.5 and is unpolarized.

4.2 Application to the observations of 1253 − 055

The best example of FDP in the UKIRT data is 1253 − 055 (3C 279) which was extremely polarized and was observed at high signal-to-noise ratios. The fluxes were fitted by power laws on all nights in 1986 August. The fluxes decreased from 1986 August 4 to August 6, but were otherwise constant. The photometric errors were too large to permit any description of the spectral index behaviour during the latter period. The flux varied by a factor of almost three between 1986 August and 1987 July 28 when the flux could again be expressed as a power law. Finally, it was observed three times during 1988 February. On February 16 and February 18 the flux could be described as a power law, while on February 17 a parabola gave the best fit. There was a decrease in the observed flux, followed by an increase in flux and spectral index. In the following discussion the data obtained on the night of 1986 August 1 will be excluded as these consist of only three frequency points. The polarization varied more often than the flux, between 45.92 ± 0.98 per cent (the highest optical polarization yet seen in a blazar) on 1986 August 5 to 19.65 ± 0.81 per cent on 1988 February 16. During 1986 August FDP developed with $dp/d\nu > 0$ as the polarization rose to a maximum. The frequency dependence of position angle was subject to the systematic errors discussed in Section 3.3. The significant FDθ seen on 1986 August 2 appears to be real. Otherwise the data appear to be consistent with variable but frequency-independent position angles and so the explanation of the behaviour on these nights is much easier. The 1987 July 28 data showed FDP ($dp/d\nu > 0$) but no FDθ. The polarization rose each night in 1988 February. Significant FDP with $dp/d\nu > 0$ was present on the final night only, while FDθ was present on the first two nights.

The Björnsson & Blumenthal (1982) formalism (hereafter the α-parameterization) will be used to parameterize the emission initially. The lack of any FDθ on most nights justifies a frequency-independent form of $\Pi(\nu)$ as an initial approximation. This has been used to fit all the data of 1253 − 055 (including the data of 1986 August 2, 1988 February 16 and 1988 February 17 for which it is not strictly applicable). Table 7 shows the results of these fits. The

Table 7. The results of fits to the 1253 − 055 data. $P(\chi^2)$ is the probability of obtaining the fitted value of χ^2 or greater if the model is correct and normal statistics are assumed.

	86/8/2	86/8/4	86/8/5	86/8/6	87/7/28	88/2/16	88/2/17	88/2/18
α-parameterisation								
$\log S_{14}$	1.57	1.40	1.35	1.25	1.16	1.86	1.86	1.86
a	0.50	−0.17	−0.08	−0.08	−0.28	1.47	1.51	1.27
b	−0.63	−1.21	−1.21	−1.10	−1.93	0.09	−0.08	−0.33
Π	0.39	0.48	0.54	0.55	0.40	0.26	0.33	0.39
χ^2	2.96	11.7	9.60	12.0	19.5	10.7	7.28	9.27
$P(\chi^2)$	0.81	0.31	0.48	0.28	0.08	0.22	0.51	0.32
Cut-off Model								
$\log S_1$	1.62	1.52	1.51	1.35	1.36	1.74	1.73	1.69
α_1	0.96	0.78	0.93	0.79	0.93	1.06	1.37	1.35
Π	0.42	0.54	0.60	0.64	0.44	0.30	0.42	0.57
ν_c	14.0	12.8	11.6	12.8	3.78	6.80	6.97	6.86
$\log S_2$	0.79	1.10	1.06	0.97	1.07	0.64	1.19	1.46
α_2	1.50	1.72	1.89	1.55	2.49	0.77	1.47	1.59
χ^2	3.92	7.50	7.21	7.27	7.51	4.71	4.16	3.28
$P(\chi^2)$	0.42	0.48	0.51	0.51	0.68	0.58	0.65	0.77

spectral flux distribution has been fitted using the parabolic form of equation (2). Fig. 10 shows the application of this parameterization to the 1986 August 5 data. The α-parameterization essentially separates the polarization behaviour into two parts, i.e. the degree of magnetic field ordering and the (potentially coupled) distribution of the projected magnetic field and electron Lorentz factor. The former factor determines the *degree* of polarization while the latter determines the spectral flux distribution. This in turn determines the FDP through the spectral index. Consequently the FDP supplies no more information about the nature of the source than is contained in the spectral flux distribution and the level of polarization. This has no relation to any physical model of the emission region. The α-parameterization implies that the FDP is entirely a result of the process which produces the spectral shape of the continuum flux, which remains unexplained. The α-parameterization cannot account for the FDθ of 1986 August 2 or 1988 February 16 and 17. FDθ implies that the functions $q(\nu)$ and $u(\nu)$ in Björnsson & Blumenthal (1982) are frequency-dependent. This gives rise to a potentially frequency-dependent form of $\Pi(\nu)$. Despite this the α-parameterization is still able to fit the observed FDP on this night, and so, although not strictly valid, it is a reasonable first approximation, suggesting that the frequency dependence due to $\Pi(\nu)$ was small compared to that arising from the α term. The change in position angle over the observed frequency range is consistent with this hypothesis.

In addition, the cut-off polarization model was fitted to these data. Again it was assumed that the magnetic field in the region was partially ordered and was such that no FDθ would result. It was not possible to fit the data using a single cut-off component. Therefore the cut-off model fits presented in Table 7 include unpolarized power-law components which permit reasonable fits. The observed FDP results from a combination of the intrinsic FDP of the cut-off component and the dilution of the polarized flux by an unpolarized component of differing spectral shape. The fits are characterized by the following parameters: the 10^{14} Hz flux densities of the cut-off and power law components (S_1 and S_2 respectively in units of mJy), the two spectral indices (α_1 and α_2), the cut-off frequency (ν_c in units of 10^{14} Hz) and the polarization parameter (Π). This latter is defined such that at low frequencies the polarization is $\Pi(1+\alpha_1)/(\frac{5}{3}+\alpha_1)$. All the frequencies are measured in the observer's frame. Fig. 11 shows the best fit to the 1986 August 5 data. In contrast to

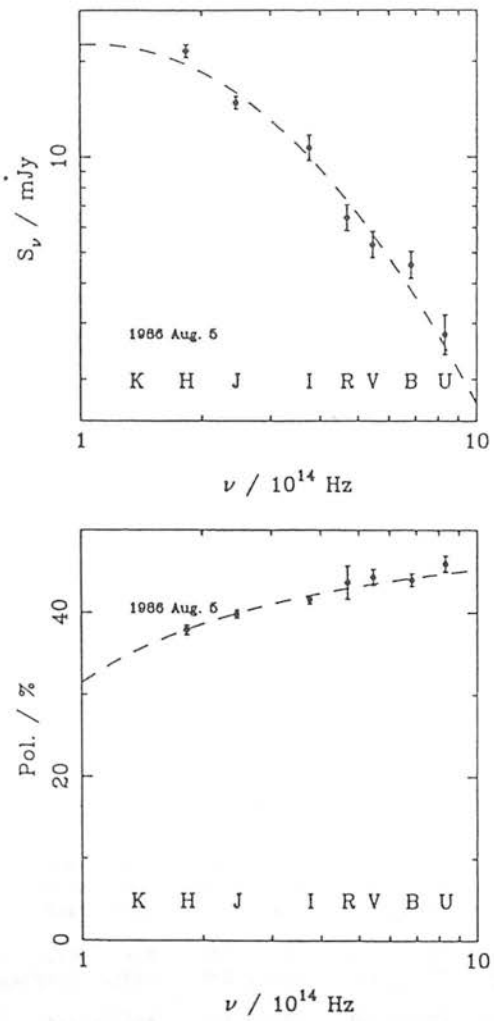


Figure 10. α-parameterization fits to the flux and polarization data of 1986 August 5.

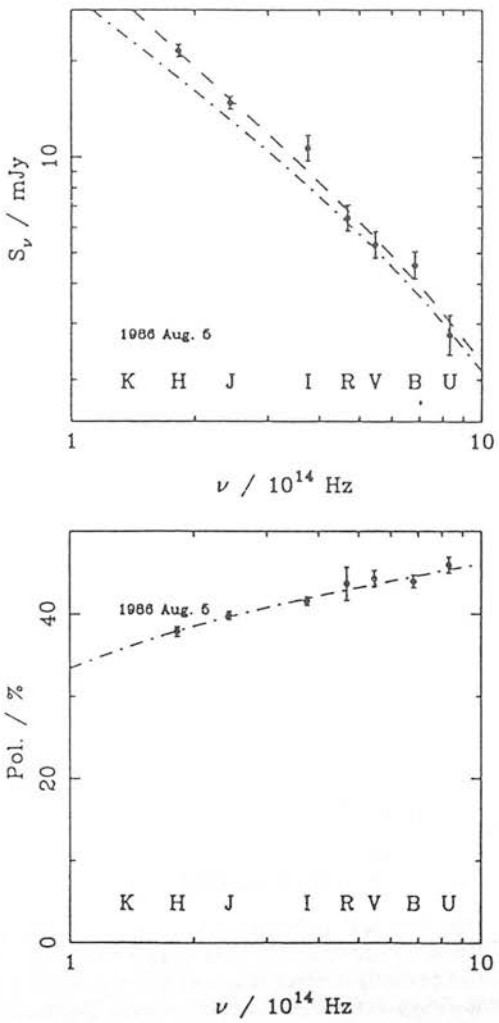


Figure 11. Cut-off fits to the flux and polarization data of 1986 August 5.

the α-parameterization, the cut-off explanation appears to be more complicated as it requires an extra unpolarized component, and an extra parameter is needed to fit the data. However, it has a number of advantages over the α-parameterization. In the cut-off picture the FDP results from the intrinsic properties of the synchrotron process and from the relative spectral indices of the two components. The spectral index parameter (α_1) appears to be consistent with relativistic strong shocks and synchrotron losses. The unpolarized component appears to have a higher spectral index which is inconsistent with current estimates of those obtainable from shock acceleration, but which is well within the range of spectral indices that can be predicted by the SSC models. This leads to a tentative identification of this component with the integrated emission of the quiescent jet, while the cut-off would arise from a shock in this flow. Hence, the cut-off model contains more information about the physical nature of the source than the α-parameterization. Those observations which displayed FDθ could be interpreted similarly to the explanation given for the α-parameterization. Alternatively, if the second component had a small polarization at a slightly different position angle to that of the cut-off component, then the observed FDθ could result without requiring intrinsic FDθ in the cut-off polarization behaviour. This is, of course, the explanation used by Holmes *et al.* (1984b) to describe their data. This would introduce extra parameters, reducing the degree of freedom to the fits which, in themselves, would make more likely a satisfactory fit to the limited number of points available.

Even assuming normal statistics, neither the α-parameterization nor the cut-off model can be rejected on the basis of the χ^2 value at the 1 per cent level. An important question is the constancy of these parameters as the source varies. Can the observed variations be characterized as simple changes in a small subset of parameters in either of these pictures? In the α-parameterization the fitted flux parameters appear to change markedly from 1986 August 2 to 1986 August 6. However, it was stated above that the spectral index did not change significantly during the observed period. If the parameters in Table 7 are used to calculate the spectral index at I (which is the best constrained part of the fit), then the calculated spectral indices remain essentially constant over the period of the 1986 August observations. The variability arises from changes in log S_{14} and II. In the cut-off model the situation is more complicated. The variations in ν_c, α_1 and α_2 in 1986 August are within the uncertainties in these parameters. The observed variations, in this picture, result from a complicated mix of variations in log S_1, log S_2 and II.

4.3 Application to other data

The α-parameterization was fitted to all polarized observations with four or more flux points, excluding the observations of 1641 + 399, for which an alternative explanation of dilution by unpolarized flux was advanced by Mead *et al.* (1988). There were 84 sets of observations satisfying this criterion, of which 58 were successfully fitted at the 5 per cent level. No allowance was made for the presence of FDθ. The fitted values of the B-field ordering parameter are displayed in Fig. 12. This confirms that in many cases the radiation can be modelled in terms of a simple inhomogeneous source. The rejected fits are summarized in Table 8. As six of the rejected fits are not significant at the 1 per cent level, it is probably reasonable to consider these consistent with the α-parameterization. The remaining cases in Table 8 are worthy of discussion. There was no *a priori* reason to expect that the α-parameterization would not have been able to fit the data of 0048 − 097 (1986 August 3), 0754 + 100 (1988 February 15, 16), or 0818 − 128 (1988 February 16, 17). The observations of 0109 + 224 and 1418 + 546 were characterized by sharp rises in the degree of polarization

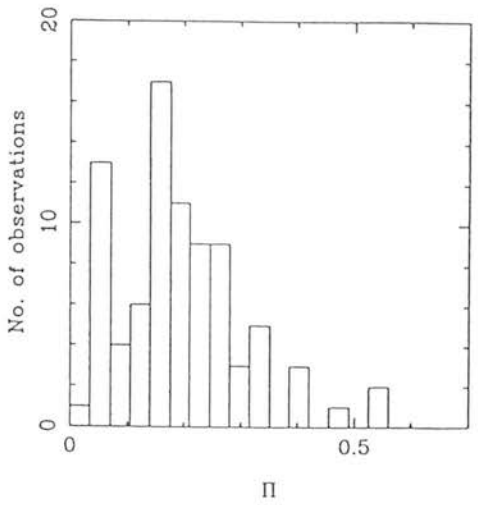


Figure 12. Distribution of the fitted values of II, and B-field ordering parameter.

Table 8. Summary of the fits to the polarization behaviour using the α-parameterization. Only those fits rejected at the 5 per cent level are shown.

Name	U.T. Date	χ^2	$P(\chi^2)$	Π
0048 − 097	1986 Aug. 3	27.40	0.0006	0.20
0048 − 097	1986 Aug. 4	15.98	0.0139	0.22
0109 + 224	1986 Aug. 7	45.77	2.6×10^{-7}	0.09
0118 − 272	1986 Jul. 27	38.27	6.7×10^{-6}	0.19
0138 − 097	1987 Sep. 20	13.93	0.0304	0.32
0235 + 164	1987 Sep. 20	27.27	0.0236	0.12
0754 + 100	1988 Feb. 15	65.8	2.0×10^{-8}	0.15
0754 + 100	1988 Feb. 16	26.8	0.0028	0.15
0818 − 128	1988 Feb. 15	18.7	0.0437	0.27
0818 − 128	1988 Feb. 16	25.8	0.0040	0.28
0818 − 128	1988 Feb. 17	5.0×10^5	0.0	0.28
1101 + 384	1988 Feb. 15	67.8	8.3×10^{-8}	0.05
1147 + 245	1988 Feb. 18	43.3	4.4×10^{-4}	0.04
1418 + 546	1987 Jul. 30	62.9	1.0×10^{-8}	0.07
1418 + 546	1988 Feb. 16	31.9	0.0014	0.04
1418 + 546	1988 Feb. 17	54.7	2.0×10^{-5}	0.06
1514 − 241	1987 Jul. 27	24.44	0.0177	0.06
2155 − 304	1986 Aug. 1	30.96	2.6×10^{-5}	0.05
2200 + 420	1986 Aug. 3	57.43	1.5×10^{-9}	0.16
2200 + 420	1987 Jul. 27	124.5	7.7×10^{-21}	0.11
2200 + 420	1987 Jul. 28	91.85	2.2×10^{-13}	0.14
2200 + 420	1987 Jul. 30	41.40	1.8×10^{-6}	0.16
2200 + 420	1987 Sep. 19	54.84	4.8×10^{-9}	0.11
2254 + 074	1986 Aug. 4	75.67	3.5×10^{-12}	0.14
2254 + 074	1987 Jul. 28	28.87	0.0003	0.15
2254 + 074	1987 Jul. 30	16.94	0.0307	0.15

with frequency (a factor of ~6 for the 0109+224 data). Extrapolating from the observed frequencies the polarization would have fallen to zero below 10^{14} Hz in both cases. The observations of 0109+224 were preceded by a number of three-frequency measurements which exhibited FDP with both $dp/d\nu > 0$ and $dp/d\nu < 0$. It is possible that a multi-component model can explain these data but without observations at lower frequencies such an explanation cannot be tested. The observation of 0118−272 has a peculiar spectrum due to an anomalous H flux point (see MEA). 1101+384, 1514−241 and 2200+420 are located in nearby galaxies so the fluxes contain an unquantifiable amount of starlight. The observations of 2155−304 and 1147+245 showed FDθ as well as FDP, and so the α-parameterization would not be strictly applicable. A likely explanation for the observed FDθ in BL Lac is a multi-component picture such as that of Holmes *et al.* (1984b) and Brindle *et al.* (1986). However, the fitting of such a model requires accurate photometry of the blazar component not achieved by these data. The observation of 2254+074 (1986 August 4) showed a frequency-independent degree of polarization from H to B and a significantly higher U polarization – behaviour inconsistent with the α-parameterization. The observations of 1987 July also showed a tendency for FDP to be more marked at the higher optical frequencies.

The cut-off model was fitted to a number of objects, selected as showing FDP with $dp/d\nu > 0$, with no evidence for either FDθ or convex spectra. Also excluded were those objects which the α-parameterization failed to explain because of the reasons given above, although the data of 0048−097 (1986 August 3), 0754+100 (1988 February 18) and 0818−128 (1988 February) were included as there was no *a posteriori* justification for the failure of the α-parameterization. The results are tabulated in Table 9. Eight of the attempted fits could be rejected at the 5 per cent level. The observations of 0048−097 (1986 August 3) were acceptable at the more correct 1 per cent level. The fits to the data of 0109+224 (1986 August 7) and 0138−097 (1987 September 20) were achieved with a high value of ν_c and thus the fits were essentially a result of FDP due to the effect of the unpolarized component on the apparent behaviour.

4.4 Discussion

Two pictures have been presented in the previous sections which explain the salient features of the observations. The α-parameterization is a simple description of the expected properties of synchrotron radiation from an inhomogeneous source. The form of the flux distribution is entirely empirical. All the information contained in the flux and polarization has been removed to this parameterization of the observed spectra. No further insights into the physics of the emission region can be obtained beyond the conclusion that the radiation is consistent with emission from an inhomogeneous source and that, to first order, equation (4) applies. SSC pictures provide a parameterization whereby the spectrum expected from an inhomogeneous relativistic jet can be evaluated (e.g. Madau, Ghisellini & Persic 1987) but rely on observations over decades of frequency to constrain their parameters.

The cut-off picture can explain many of the examples of FDP including some that the α-parameterization cannot.

Table 9. A summary of the fits to the FDP of blazars, using the cut-off model outlined in Section 4.2 for fits to the data of 1253−055.

Name	U.T. Date	χ^2	$P(\chi^2)$	$\log S_1$	α_1	Π	ν_c	$\log S_2$	α_2
0048 − 097	1986 Aug. 3	16.7	0.01	0.48	1.06	0.95	11.5	1.24	1.48
0048 − 097	1986 Aug. 4	4.79	0.31	0.39	0.87	0.85	119.	1.00	1.12
0048 − 097	1986 Aug. 6	8.47	0.21	0.85	0.50	0.25	14.1	1.13	3.24
0109 + 224	1986 Aug. 7	7.19	0.30	−0.38	0.50	1.00	6.8×10^6	1.16	1.33
0118 − 272	1986 Aug. 5	7.88	0.25	0.50	0.93	1.00	12.9	1.13	1.18
0118 − 272	1987 Jul. 30	4.44	0.82	0.99	1.01	0.27	26.9	0.86	1.38
0118 − 272	1987 Sep. 21	6.53	0.16	0.99	1.04	0.28	21.9	0.76	1.51
0138 − 097	1987 Jul. 28	1.71	0.94	0.68	0.98	0.27	32.4	0.89	3.67
0138 − 097	1987 Sep. 19	10.6	0.23	0.50	0.95	0.52	15.3	0.58	1.56
0138 − 097	1987 Sep. 21	3.01	0.56	0.69	0.72	0.32	9.96	−0.62	2.01
0138 − 097	1987 Sep. 20	10.2	0.04	0.29	0.74	0.69	50.0	0.56	1.14
0235 + 164	1987 Sep. 20	23.2	0.003	0.74	0.69	0.13	2.42	0.92	3.28
0735 + 178	1988 Feb. 16	6.55	0.36	0.83	0.62	0.42	4.17	1.08	1.65
0754 + 100	1988 Feb. 18	7.84	0.45	0.91	0.62	0.24	3.89	1.08	1.33
0818 − 128	1988 Feb. 15	26.3	9.5×10^{-3}	0.51	0.50	0.30	3.52	0.17	2.22
0818 − 128	1988 Feb. 16	16.9	0.03	0.43	0.83	0.35	4.07	0.45	2.39
0818 − 128	1988 Feb. 17	17.8	0.007	0.21	1.53	0.78	16.8	0.75	1.93
0851 + 202	1988 Feb. 16	12.3	0.056	1.06	1.16	0.23	9.53	0.52	1.88
1418 + 546	1987 Jul. 30	13.3	0.10	−0.38	0.56	1.00	20.0	1.38	1.65
1749 + 096	1987 Jul. 27	29.2	5.7×10^{-5}	0.80	1.32	0.12	1.95	−4.66	−3.18
2155 − 304	1987 Jul. 27	9.40	0.40	1.72	0.50	0.21	33.3	1.56	0.86
2254 + 074	1987 Aug. 3	41.8	2.0×10^{-7}	0.68	0.84	0.17	2.87	0.46	3.25

This model is attractive because an upper energy cut-off in the electron energy distribution is an expected feature of particle acceleration at shocks. This cut-off occurs because above certain energies the acceleration time-scale is longer than the energy loss time-scale. Biermann & Strittmatter (1987) derive a cut-off frequency between 3×10^{14} Hz and 2×10^{15} Hz. The value of the cut-off frequency is critically dependent on the assumptions made in its derivation, e.g. a Kolmogorov spectrum ($\propto k^{-5/3}$) was used to model the spectrum of the turbulent magnetic energy density responsible for the scattering. Heavens (1984) showed that a k^{-1} spectrum can give rise to a cut-off at X-ray frequencies for shocks speeds greater than 3000 km s^{-1}.

It is worth comparing the expected values with the fitted results for ν_c given in Sections 4.2 and 4.3. For 1253−055 the Biermann & Strittmatter (1987) upper limit corresponds to an observer's frame frequency of 1.3×10^{15} Hz. The 1986 August fits are broadly consistent with being around this upper limit, while the 1987 July and 1988 February fits are at lower frequencies. Although interpretation of ν_c for the objects in Table 9 is harder, as not all the objects have measured redshifts, most of the fits have best-fit cut-off frequencies which are either close to Biermann & Strittmatter's (1987) upper limit or higher. This is consistent with the fact that the polarization 'hump', at frequencies above ν_c, is not observed. In fact, instances of $dp/d\nu<0$ are either transient phenomena (e.g. 0109+224 in 1986 August) or have explanations in terms of the line emission and 'blue bump'. No significant cases also showing $d^2\alpha/d^2\nu>0$ were seen. This is a problem for the interpretation of the observed FDP in terms of cut-offs unless the Biermann & Strittmatter (1987) model genuinely indicates that the physical processes in shock acceleration produce cut-offs over a relatively restricted range of frequency.

How can the cut-off picture be placed in a framework which explains all the observed features of blazar emission? It has been shown how the cut-off picture can explain FDP where it occurs without significant FDθ. However, whenever FDθ occurs, this model fails to explain the observations. The motivation for advancing the cut-off model was FDP, although many observations were characterized by power-law spectra, frequency-independent degrees of polarization and frequency-independent position angles of polarization. These observations can be placed in the same framework as the cut-off model in two ways. Either they could be shock components observed at frequencies much lower than ν_c or they could be the integrated emission of the quiescent jet which may be slightly polarized. If the 'best guess' model of the polarization behaviour is taken to be a cut-off component representing a shock with a steeper spectrum component representing the underlying jet, how would this relate to observations outside of the IR/optical spectral region? Rusk & Seaquist (1985) and Impey (1987) have shown that there is a tendency for preferred optical polarization position angles to be aligned along the *VLBI* structure. Such quiescent position angles could be speculatively assigned to a polarized jet component (Brindle *et al.* 1986). The other component would then be the shock or cut-off component, which may in principle have a different position angle. In one case the consistency of this can be checked by reference to the observations of 0851+202. In their two-component decomposition Holmes *et al.* (1984b) identified two distinct orientations. One appeared constant (at about 100°) and the other rotated during the observations of 1983 January. Roberts & Wardle (1987) report *VLBI* flux and polarization observations of this object at two epochs (1981.9 and 1982.9), which show highly polarized extensions from the moderately polarized core at a position angle around −100°. The polarization of the supposed constant component is close to being orientated along the jet axis, but is not quite close enough to be definitely linked with the quiescent components described above.

5 CONCLUDING REMARKS

This paper analyses the results of a programme of multi-frequency polarimetric observations of blazars. 157 observations of 44 blazars were obtained during four separate observing runs in 1986 August, 1987 July, 1987 September and 1988 February. The data consist of simultaneous observations using up to 8 different filters at infrared to optical frequencies. The major results will now be summarized.

(i) In many cases the spectrum could be represented as a power law over the full range of observed frequencies, but significant spectral curvature was common. The behaviour of the polarized flux density was similar to that of the total flux density whose properties could be assumed to be representative of the synchrotron component. The spectral behaviour was inconsistent with that expected from particle acceleration at relativistic shocks unless the observations were made in the region of a high-frequency cut-off. Inhomogeneous source models can also explain the observed behaviour.

(ii) Frequency dependence of the degree of polarization was seen in many cases, generally with $dp/d\nu>0$. The counter examples were often associated with the 'blue bumps' discussed in Section 2. Frequency dependence of the polarization position angle was a less common feature of blazar behaviour. No evidence was found to support claims that frequency dependence is associated with high levels of polarization.

(iii) Variability was a common feature of the observations. No simple characteristics have yet been found to describe all the variations.

(iv) Most of the observed frequency dependence of polarization in blazars is linked to the intrinsic curvature in the blazar flux spectrum. This suggests an inhomogeneous model for the emission region is required. An example is provided by a polarized component with a high-frequency cut-off and a second component with a steeper spectral index and no significant polarization, tentatively identified with shock accelerated electrons and a quiescent jet, respectively. This cannot simply explain the observed frequency dependence of position angle without recourse to separate polarized components.

Further work must provide a rigorous test of the cut-off model. The fits would be better constrained if flux information were available from higher and lower frequencies than those observed. This would make the empirical determination of the spectral index more secure at *K*, *H* and *J*, and *B* and *U*. This in turn would more severely test the observed polarization behaviour at these frequencies. Polarization information at lower frequencies where the hypotheti-

cal unpolarized (or low polarization) and steeper spectrum component will dominate would be particularly important. If, on the other hand, a theoretical prediction of the observed spectral shape of a blazar is not available, then the validity of the α-parameterization must be tested using an empirical fit to the observed flux data. In this case, the α-parameterization would also benefit from wider frequency coverage.

ACKNOWLEDGMENTS

We wish to thank J. Bailey and C. Brindle for assistance in making these observations and A. Heavens for useful discussions. K. Ballard and A. Mead wish to acknowledge the receipt of SERC studentships.

REFERENCES

Abramowicz, M. A., Calvani, M. & Nobili, L., 1980. *Astrophys. J.*, **242**, 772.

Allen, D. A., Ward, M. J. & Hyland, A. R., 1982. *Mon. Not. R. astr. Soc.*, **199**, 969.

Bailey, J. A., Hough, J. H. & Axon, D. J., 1983. *Mon. Not. R. astr. Soc.*, **203**, 339.

Begelman, M. C., Blandford, R. D. & Rees, M. J., 1984. *Rev. Mod. Phys.*, **56**, 255.

Biermann, P. L. & Strittmatter, P. A., 1987. *Astrophys. J.*, **322**, 643.

Björnsson, C.-I., 1985. *Mon. Not. R. astr. Soc.*, **216**, 241.

Björnsson, C.-I. & Blumenthal, G. R., 1982. *Astrophys. J.*, **259**, 805.

Blandford, R. D., 1979. In: *Particle Acceleration Mechanisms in Astrophysics, Proc. A.I.P. Conf.* No. 56, p. 335.

Bregman, J. N., Lebofsky, M. J., Aller, M. F., Rieke, G. H., Aller, H. D., Hodge, P. E., Glassgold, A. E. & Huggins, P. J., 1981. *Nature*, **293**, 714.

Brindle, C. *et al.*, 1985. *Mon. Not. R. astr. Soc.*, **214**, 619.

Brindle, C., Hough, J. H., Bailey, J. A., Axon, D. J. & Hyland, A. R., 1986. *Mon. Not. R. astr. Soc.*, **221**, 739.

Conover, W. J., 1980. *Practical Nonparametric Statistics*, 2nd edn, Wiley, New York.

Cotton, W. D., Wittels, J. J., Shapiro, I. I., Marcaide, J., Owen, F. N., Spangler, S. R., Rius, A., Angulo, C., Clark, T. A. & Knight, C. A., 1980. *Astrophys. J.*, **238**, L123.

Cruz-Gonzalez, I. & Huchra, J. P., 1984. *Astr. J.*, **89**, 441.

Gear, W. K., Robson, E. I., Ade, P. A. R., Griffin, M. J., Brown, L. M. J., Smith, M. G., Nolt, I. G., Radostitz, J. V., Veeder, G. & Lebofsky, L., 1985. *Astrophys. J.*, **291**, 511.

Gear, W. K., Robson, E. I. & Brown, L. M. J., 1986. *Nature*, **324**, 546.

Ghisellini, G., Maraschi, L. & Treves, A., 1985. *Astr. Astrophys.*, **146**, 204.

Heavens, A. F., 1984. *Mon. Not. R. astr. Soc.*, **207**, 1P.

Heavens, A. F., 1988. In: *Workshop on Hotspots in Extragalactic Radio Sources*, eds Meisenheimer, K. & Röser, H.-J., Springer-Verlag, Berlin.

Heavens, A. F. & Drury, L. O'C., 1988. *Mon. Not. R. astr. Soc.*, **235**, 997.

Holmes, P. A., 1985. *PhD. thesis*, University of Edinburgh.

Holmes, P. A., Brand, P. W. J. L., Impey, C. D. & Williams, P. M., 1984a. *Mon. Not. R. astr. Soc.*, **210**, 961.

Holmes, P. A. *et al.*, 1984b. *Mon. Not. R. astr. Soc.*, **211**, 497.

Impey, C. D., 1987. In: *Superluminal Radio Sources*, p. 233, eds Zensus, J. A. & Pearson, T. J., Cambridge University Press, Cambridge.

Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1982. *Mon. Not. R. astr. Soc.*, **200**, 19.

Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1984. *Mon. Not. R. astr. Soc.*, **209**, 245.

Kardashev, N. S., 1962. *Sov. Astr. A. J.*, **6**, 317.

Kikuchi, S. & Mikami, Y., 1987. *Publs astr. Soc. Japan*, 39, 237.

Kirk, J. G. & Heavens, A. F., 1989. *Mon. Not. R. astr. Soc.*, 239, 995.

Kirk, J. G. & Schneider, P., 1987. *Astrophys. J.*, **315**, 425.

Königl, A., 1981. *Astrophys. J.*, **243**, 700.

Landau, R. *et al.*, 1986. *Astrophys. J.*, **308**, 78.

Longair, M. S., 1981. *High Energy Astrophysics*, Cambridge University Press, Cambridge.

Madau, P., Ghisellini, G. & Persic, M., 1987. *Mon. Not. R. astr. Soc.*, **224**, 257.

Malkan, M. A. & Sargent, W. L. W., 1982. *Astrophys. J.*, **254**, 22.

Marscher, A. P., 1980. *Astrophys. J.*, **235**, 386.

Marscher, A. P. & Gear, W. K., 1985. *Astrophys. J.*, **298**, 114.

Mead, A. R. G., Brand, P. W. J. L., Hough, J. H. & Bailey, J. A., 1988. *Mon. Not. R. astr. Soc.*, **233**, 503.

Mead, A. R. G., Ballard, K. R., Brand, P. W. J. L., Hough, J. H., Bailey, J. A. & Brindle, C., 1990. *Astr. Astrophys. Suppl.*, in press.

Meisenheimer, K. & Heavens, A. F., 1986. *Nature*, **323**, 419.

Moore, R. L. *et al.*, 1982. *Astrophys. J.*, **260**, 415.

Moore, R. L. & Stockman, H. S., 1984. *Astrophys. J.*, **279**, 465.

Nordsieck, K. H., 1976. *Astrophys. J.*, **209**, 653.

Pacholczyk, A. G., 1970. *Radio Astrophysics*, Freeman, San Francisco.

Pacholczyk, A. G., 1977. *Radio Galaxies*, Pergamon, Oxford.

Puschell, J. J. & Stein, W. A., 1980. *Astrophys. J.*, **237**, 331.

Rees, M. J., 1984. *Ann. Rev. Astr. Astrophys.*, **22**, 471.

Rieke, G. H., Lebofsky, M. J. & Kinman, T. D., 1979. *Astrophys. J.*, **232**, L151.

Roberts, D. H. & Wardle, J. F. C., 1987. In: *Superluminal Radio Sources*, p. 193, eds Zensus, J. A. & Pearson, T. J., Cambridge University Press, Cambridge.

Rusk, R. & Seaquist, E. R., 1985. *Astr. J.*, **90**, 30.

Shields, G., 1978. *Nature*, **272**, 706.

Smith, P. S., Balonek, T. J., Elston, R. & Heckert, P. A., 1987. *Astrophys. J. Suppl.*, **64**, 459.

ASTRONOMY & ASTROPHYSICS
SUPPLEMENT SERIES

APRIL 1990, PAGE 183

Astron. Astrophys. Suppl. Ser. **83**, 183-204 (1990)

5

Optical and infrared polarimetry and photometry of blazars

A.R.G. Mead([1]), K.R. Ballard([1]), P.W.J.L. Brand([1]), J.H. Hough([2]), C. Brindle([2]) and J.A. Bailey([3])

([1]) Dept of Astronomy, Edinburgh University, Edinburgh, Scotland EH9 3HJ, G.B.
([2]) Division of Physical Sciences, Hatfield Polytechnic, Hatfield, Herts, AL10 9AB, G.B.
([3]) Joint Astronomy Center, 665, Komohana St, Hilo, Hawaii 96720, U.S.A.

Received September 12, accepted November 14, 1989

Abstract. — We have carried out a polarization study of 44 blazars and candidate blazars at optical and near-infrared frequencies covering the period 1986 August to 1988 February. The data show flux and polarization properties changing on internight timescales, and commonly exhibit frequency dependence of spectral index, polarization and polarization position angle. The observations of the quasar 1253 – 055(3C279) showed a *U*-band polarization of $45.5 \pm 0.9\%$, the highest ever seen in a blazar. The interpretation of the data is the subject of a further paper (Ballard *et al.*, 1989).

Key words : BL Lacertae objects.

1. Introduction.

We have the results of a continuing programme of simultaneous measurements of the linear Stokes parameters in optical and near-infrared wavebands. This is the fourth paper in the series (see Impey *et al.*, 1982 ; Impey *et al.*, 1984 ; Holmes *et al.*, 1984a). The observations, described in section 2, are tabulated in table III. In all 44 objects were observed at the 3.8m United Kingdom Infrared Telescope (UKIRT) during four observing runs from 1986 July 31 to August 7, from 1987 July 27 to 30, from 1987 September 18 to 21 and from 1988 February 15 to 18. This yielded 135 sets of simultaneous photometric and polarimetric data, 23 sets of polarimetric data without photometric calibration and 23 sets of photometric data with no significant polarization at the 3σ level. The data which lack photometric calibration come from poor observing conditions during the 1987 September run. Consequently, for those objects where no polarization was detected (at 3σ), some of the upper limits are quite high. These data constitute the largest such set of observations yet obtained.

2. Observational techniques.

Two instruments were used to obtain the data presented in this paper. The Mark I Hatfield Polarimeter (as described by Brindle *et al.*(1986), with the addition of an extra optical channel) was used for the observations of 1986. The Mark II instrument, described below, was used for the observations of 1987 and 1988. Both operate at the $f/35$ Cassegrain focus of UKIRT. All the observations were taken with an 8 arcsec aperture.

2.1 THE MARK II HATFIELD POLARIMETER.— The Mark II Hatfield Polarimeter was constructed at the Hatfield Polytechnic and was first used in 1987 July at UKIRT. Its optical layout is shown in figure 1. It is a twin-beam instrument with two beams separated by 50.8 mm., which corresponds to an angular separation on the sky of 81.3 arcsec at the $f/35$ focus of UKIRT. The advantage of this set-up over the Mark I polarimeter is that the optical and infrared measurements are now performed on separate beams so that the wave-plates and other optics have been optimised for the appropriate wavelength ranges.

The infrared beam passes through an IR-achromatic wave-plate ($1 - 2.5\mu$m) and then an IR/optical dichroic. The reflected IR beam passes through a wire-grid analyser and into the observatory cryostat (UKT9). The transmitted optical beam is sent to the UKIRT TV camera for acquisition and guiding. An efficiency correction must be applied to the infrared polarization measurements, due to inefficiencies which arise from the frequency response of the detector and the wire-grid analyser. The appropriate efficiency to use for correcting the measurements is the measured polarization of an unpolarized star with the Glan prism inserted into the IR beam. The wire grid indicated on the calibration slide in figure 1 is only used if calibration at longer wavelengths than *K* is required, where the Glan prism can no longer be used. No such observations are

* *Send offprint requests to* : K.R. Ballard.

presented in this paper.

The optical beam is measured in a similar to the Mark I device. A super-achromatic wave plate is followed by a Foster prism which produces two orthogonally polarized beams. Dichroics are then used to split the beams into U and R channels, and the B, V and I beams channels with filters used in each beam to form the separate passbands. The polarimeter can thus measure the U, B, V, R, I wavebands and one IR waveband simultaneously. The filters quite closely reproduce the Cousins *UBVRI* system (Bessell, 1979) and any differences have been ignored. The effective wavelengths of the filters are listed in table I.

The polarimeter is operated in a similar manner to the Mark I device. A cycle consists of four rotations of the wave-plate. At the start of the cycle the object is centred on the optical beam. At the end of the first rotation the telescope is nodded to bring the object into the IR beam and then nodded back after the third rotation. While not centred on the object, the beams are measuring the sky. As for the Mark I device, on-line software updates the polarization and flux measurements and their combined errors — obtained from photon statistics — at the end of each cycle.

The polarimeter returns relative fluxes for each waveband, which are then calibrated by observations of standard stars. The infrared magnitudes were calibrated by observations of stars from the UKIRT standards list and the optical standards were obtained from the catalogue of Landolt (1983). Airmass corrections were performed using observationally determined extinctions. The fluxes have all been corrected for interstellar extinction with the values of $E(B-V)$, taken from Burstein and Heiles (1982), given in table III. The extinction curve is taken from Rieke and Lebofsky (1985).

The degree of polarization observed is subject to a statistical bias which must be corrected. This is inherent in the measurement of a vector quantity in the presence of noise. Wardle and Kronberg (1974) give the expression for the modal value, which is used to relate the true polarization p to the measured polarization p'.

$$p = p'\sqrt{1 - \left(\frac{\sigma_p}{p'}\right)^2} \qquad (1)$$

The best estimate of the error on p is the error on the measured polarization σ_p. The measured position angle is the best estimate of the true position angle, but the error is underestimated in the case of poor signal-to-noise (Wardle and Kronberg, 1974).

The glan is used to measure the relative position angles in each band. The zero of position angle on the sky is determined by measurements of polarized standard stars taken from Serkowski (1974). These position angles are only quoted to one degree accuracy and no errors are given. Consequently for many of the measurements given, the position angle error is dominated by the unknown error in these angles.

3. Results.

3.1 INDIVIDUAL OBJECTS.— The shape of the spectrum will b characterised by the use of the (local) spectral index;

$$\alpha(\nu) = -\frac{\mathrm{d}\log S_\nu(\nu)}{\mathrm{d}\log \nu}. \qquad (2$$

The condition of a frequency-dependent spectral inde which results in curvature in the $\log S_\nu(\nu) - \log \nu$ plane i referred to as spectral curvature.

To test spectral curvature, a power-law was fitted to al the UKIRT photometric data. All the Mark I Hatfiel Polarimeter data consist of measurements at H, I and B a least. The Mark II data all have one infrared and at leas four optical measurements. For some high airmass (se $z > 1.5$) observations, the U data had to be rejected becaus atmospheric refraction makes the photometry unreliable. Consequently all the data can be fitted by a power-law with at least one degree of freedom. These fits were achieved by minimising the χ^2 statistic, the use of which i only valid if the errors concerned are distributed normally. Unfortunately, this is certainly *not* the case here. The errors on the photometry are not simply the random errors associated with photon counting, but include the systematic uncertainties in the photometric calibration. It is most unlikely that these latter errors are distributed normally. However its use here is justified as the aim is to obtain an idea of the frequency of spectral curvature within the blazar population rather than to reject a power-law fit for any one individual object.

All the data which had two or more polarization measurements were tested for frequency dependence of the degree of polarization (hereafter FDP). The weighted mean of the measured polarizations was calculated and then tested again all the measured polarizations by use of the χ^2 statistic with the loss of one degree of freedom. Similarly the data were also tested for frequency dependence of the position angle of polarization (hereafter FDθ). Polarimetric errors should have an approximately normal error distribution in the case of high signal-to-noise, and the position angle error distribution is also approximately normal for high signal-to-noise observations (Wardle and Kronberg, 1974).

We now give a brief description of the observations of each object.

0048 – 097, OB – 081

This object was confirmed as a blazar by Angel and Stockman (1980) (hereafter AS80). 0048 – 097 was observed extensively in the 1986 August, 1987 July and 1987 September runs. It consistently showed FDP, but not FDθ, in 1986 August. No such behaviour was observed in 1987 July or September. Typically the position angle changed from night to night by a few degrees.

PKS 0106 + 013

This highly polarized quasar, HPQ, (confirmed by Moore and Stockman, 1984) was observed on one night in each of the first three runs. No significant polarizations were recorded during the 1986 July/August and the 1987 September runs. The only significant polarization measured was the U polarization of 1987 July 28 ($p(U) = 14.37 \pm 2.39\%$)

GC 0109 + 224

An AS80 blazar, 0109 + 224 was observed during all three runs. The polarization in 1986 August was highly variable and displayed both FDP and FDθ. Both $dp/d\nu > 0$ and $dp/d\nu < 0$ were observed on successive nights. The polarization observed in 1987 July was $\sim 10\%$ but essentially constant. The 1987 September data showed more variation but only one case of FDP.

PKS 0118 – 272

This object was a radio source identification with a smooth IR/optical spectrum (Wilkes *et al.*, 1983). Impey and Tapia (1988) measured one significant (i.e. $> 3\%$) polarization. The observations presented here confirm that this object is a blazar. The polarization was high ($p \sim 17\%$) but constant through 1987 August. FDP was seen in 1987 July with $dp/d\nu > 0$.

0138 – 097

This was another radio source identification with a smooth IR/optical spectrum (Fricke *et al.*, 1983), which was confirmed as a blazar by Impey and Tapia (1988). The polarization in 1986 August was $\sim 6\%$ with marginal indications of $dp/d\nu < 0$. The 1987 July data showed higher polarization ($\sim 20\%$) without FDP, while the 1987 September polarizations were of similar amplitude but with $dp/d\nu > 0$.

PKS 0219 – 164

This object was identified as a blazar by Meisenheimer and Röser (1984). It was observed only once (1987 July 28), when it had (frequency averaged) polarization of $12.63 \pm 0.18\%$ at a position angle of $160.9 \pm 0.5°$. This position angle is consistent with the range predicted by the "oblique rotator" model of Meisenheimer and Röser; though of course this model cannot be confirmed without some data on variability.

0219 + 428, 3C 66A

An AS80 blazar which was observed in all three runs. The polarization behaviour in 1986 August was fairly constant with no marked frequency dependence, and this was also the case in 1987 July. Marginal evidence for FDP was seen in 1987 September. Inoue (personal communication) reports 10 GHz observations of this object on 1987 September 17 which show 2.8% polarization at 5°.

AO 0235 + 164

This is the object for which Impey *et al.* (1982) measured their record polarization of $p(V) = 43.9 \pm 1.4\%$. Two photometric observations of this object were made. The spectrum was very steep $\alpha(B) = 4.61$, but the polarizations were always of the order of 10% . Inoue (personal communication) reports 10 GHz observations on 1987 September 17 which show 1.5% polarization at 14°. This position angle is consistent with the IR/optical data.

0300 + 470, 4C 47.08

This AS80 blazar was observed only once on 1987 September 20 in poor photometric conditions. The data were consistent with $p = 8.91 \pm 0.64\%$ and $\theta = 9.42 \pm 1.92°$, independent of frequency.

1H 0323 + 022

This X-ray selected object was classified as a blazar by Feigelson *et al.* (1986) who, in particular, noted its extremely rapid X-ray variability. They observed it to be optically polarized (2 – 9%) in 1983 and 1984. No significant polarization was observed on 1986 August 1($p < 6.3\%; 3\sigma$ upper limit at I). Polarization was measured on 1987 September 21, with the average over all wavebands being $3.56 \pm 0.45\%$.

0336 – 019, CTA 26

This HPQ (Moore and Stockman, 1981), was observed only once on 1986 August 5, and no significant polarization was measured as the object was faint and the integration was accordingly terminated early.

0338 – 214

This is a smooth optical spectrum radio source identification (Wilkes *et al.*, 1983). The polarization was observed to be $\sim 10\%$ in both 1986 August and 1987 September with no frequency dependence.

1H 0414 + 009

This X-ray selected object is a candidate blazar. Ulmer *et al.* (1983) classified it as a BL Lac object and Impey and Tapia (1988) reported an optical polarization of $2.76 \pm 0.29\%$. On 1988 February 15 and February 16 the B polarization was $6.4 \pm 0.8\%$. There was no evidence for frequency dependent polarization on either night.

PKS 0422 + 004

This AS80 blazar was observed on 1988 February 16 and 18. There was no evidence for frequency dependent polarization, but the polarization decreased from 21.4 ± 0.3 to $12.4 \pm 0.2\%$ between these dates.

PKS 0735 + 178

This well known blazar was observed on 1987 September 19 in poor photometric conditions. The measured polarization was frequency independent ($p = 6.9 \pm 0.4\%$) but the position angles were marginally frequency dependent (average value 136°). On 1988 February 16 there was evidence for $dp/d\nu > 0$.

PKS 0736 + 017

This AS80 blazar was observed on 1988 February 15, 16 when no significant polarization was detected at the 3% level.

0754 + 100, OI 090.4

0754 + 100 is a well known AS80 blazar which was observed on 1988 February 15, 16, 17 and 18. The polarization showed $dp/d\nu > 0$, the degree of which decreased with time. On the last two nights the position angle was frequency independent at an average angle of 36.2°.

0818 – 128, OJ – 131

This AS80 blazar was observed on 1988 February 15, 16 and 17. During the course of the observations the degree of frequency dependence of the polarization increased. The position angle was frequency independent throughout at an average angle of 81°.

0851 + 202, OJ287

This well known object was the subject of a previous paper (Holmes *et al.*, 1984b), when it displayed notable behaviour, interpreted in terms of a two-component model for the emission. It was observed on 1988 February 15, 16 and 17 to have a polarization around 10% . $dp/d\nu > 0$ was observed on February 16 and 17. The data presented here for OJ287 are satisfactorily fitted by the two-component model.

0906 + 015

0906 + 015 is an AS80 blazar which was observed once in February 1988. At the 3% level no significant polarization was detected.

1101 + 384, Mkn 421

This is an AS80 object, located in a galaxy at redshift, $z = 0.0308$. It was observed once, on 1988 February 15. The fact that the polarization showed complex behaviour and the curvature in the spectrum is in agreement with the hypothesis that the galactic flux significantly contaminates the blazar component.

1147 + 245

This AS80 blazar displayed $dp/d\nu > 0$ and complex position angle behaviour on the single occasion on which it was observed (1988, February 18).

1156 + 295

On 1988 February 15, 16 and 17 this well studied AS80 blazar was observed to have frequency independent polarizations and position angles (frequency averaged) of 24% at 165°.

1253 – 055, 3C 279

This object is a well studied superluminal radio source, which has long been known to be a blazar. The observations of 1986 August show a polarization flare, with the degree of polarization increasing during the course of the run, and consistently showing $dp/d\nu > 0$. The *U* data of 1986 August 5 show a polarization of 45.92 ± 0.98% . This is the largest polarization ever seen in the IR/optical for blazar (c.f. 0235 + 164 ; Impey *et al.*, 1982). On 198 July 28 the polarization was still high, but not at the recor level. In 1988 February the polarization rose throughou the observations from 19.6 ± 0.8% to 30.7 ± 0.6% in the band.

1413 + 135, OQ 122

This is a radio source with a very steep IR/optical spectrun (Beichman *et al.*, 1981 ; Bregman *et al.*, 1981). The latte gives the only recorded significant polarization of 19 ± 3% at H. This object was observed in 1986 August, but was to faint for IR polarimetry to be feasible in the time available The spectrum was so steep that only an upper limit i available for the optical flux.

1418 + 546, OQ 530

This AS80 blazar was observed during all four runs. In 198 August the polarization showed FDP with $dp/d\nu > 0$, which was repeated in the measurements of 1987 July 30, where the polarization increased from 2.5% at H to 8.7% at *U*. This latter was accompanied by significant FDθ. The data of 1987 September were all obtained in poor photometric conditions. However the polarization data still showed FDP but no FDθ. During the 1988 February the position angle was constant at 144° on the 16th and 17th. On the final night significant FDθ was observed, from 132.0 ± 2.5° at H to 156.2 ± 3.9° at *U*.

1424 + 240

This is a smooth spectrum radio source identification which was confirmed as a blazar by Impey and Tapia (1988). Frequency independent polarizations (19.6%) and position angles (116°) were recorded on 1988 February 16, 17 and 18.

PKS 1510 – 089

This object was confirmed to be a blazar by Moore and Stockman (1981) and Smith *et al.* (1987). It is one of the most violently variable HPQ's with $\Delta m = 5.4$ (Moore and Stockman, 1981). It was observed only once on 1986 August 1. The 3σ upper limit to the I polarization was 6.3% .

1514 – 241, AP Lib

One of the original BL Lac objects, 1514 – 241, AP Lib was observed only twice (1986 August 1 and 1987 July 27). The polarization of this object is typically quite low compared to most other blazars (e.g. AS80). Our observations are consistent with this. FDP was seen on 1987 July 27.

1538 + 149, 4C 14.60

An AS80 blazar which was observed on 1986 August 1 when no significant polarization was observed. The 3σ upper limit was high (27%) as the object was faint.

1641 + 399, 3C 345

The behaviour of this object was the subject of a recen paper (Mead *et al.*, 1988).

1652 + 398, Mkn.501

This BL Lac object was observed only once on 1987 September 21. Both the degree and position angle of polarization were observed to be frequency dependent (at the 0.5% level of significance).

1717 + 178, OT 129

This AS80 blazar was observed twice in 1986 August with strong $dp/d\nu < 0$ on the second night. No significant polarization was measured on 1987 July 30 ($p < 21\%$) (upper limit at R).

1727 + 502, 1 Zw. 186

An AS80 blazar which was observed twice in 1986 August, when it showed variable FDP and FDθ. It was also observed on 1987 September 21 when $dp/d\nu > 0$ was observed (without FDθ).

1749 + 096, OT081

An AS80 blazar which was extensively observed in 1986 August. No FDP was seen but significant variations were seen in the polarization. FDθ was seen on 1986 July 31. On 1987 July 27, 1749 + 096 was faint but $dp/d\nu > 0$

1921 – 293, OV – 236

This blazar was first classified as such by Wills and Wills (1981). Confirmation of it being polarized was provided by Impey *et al.* (1982). The observations of 1986 August indicate appreciable photometric variability (a factor of two from 1986 August 6 to August 7). Appreciable FDP was seen on 1986 August 6.

2032 + 107 MC

This candidate blazar was first classified by Zotov and Tapia (1979), who have published the only polarization measurement. Antonucci *et al.* (1987) have confirmed its extragalactic nature. One observation of this object was made on 1987 July 27. Significant polarization was not detected. The 3σ upper limit was 1.32% at R.

PKS 2155 – 304

This is one of the brightest of the AS80 BL Lac objects, and was observed as a part of the 1986 August and 1987 July and September runs, but its polarization is typically quite low (3 – 7% ; AS80). The data of 1986 August showed polarization at about this level and exhibited variability in both amplitude and frequency dependence. The data of 1987 July 27 showed a higher polarization ($\sim$ 10%) and both FDP and FDθ. The data of 1987 September 21 were obtained in poor conditions, but the polarization showed FDP.

2200 + 420, BL Lac

BL Lac was observed more often than any other object in this programme. Photometric variability was small over the 1986 August and 1987 July and September runs. Instances of both FDP and FDθ were observed in all three periods. The position angle of the polarization was in the range 10 – 40° for all observations. This is essentially the same position angle seen by Brindle *et al.* (1985). Inoue (personal communication) reports a high 10 GHz polarization of 9.2% at 21° on 1987 September 17. This position angle is somewhat different from that seen in the IR/optical on 1987 September 19 ($\sim$ 40°).

PKS 2208 – 137

This is an AS80 blazar which was observed four times in the 1986 August run. On each occasion it was observed to have an abnormally flat IR/optical spectrum and no polarization. The lowest 3σ upper limit was 1.78% at B on 1986 August 4.

2223 – 052, 3C 446

This is violently variable HPQ. During the 1986 August observations the object was faint and no reasonable limits on the optical polarization were obtained. In 1987 July frequency independent polarization was seen varying from 12% to 8% . The observations of 1987 September were obtained in poor photometric conditions. The noisy data showed some evidence for FDP. The 10 GHz polarization data of the 1987 September 17 ($p \approx 4\%$ *and* $\theta \approx 2°$), showed a position oriented roughly at 90° from the IR/optical data (Inoue, personal communication).

2230 + 114, CTA 102

This object is an AS80 blazar and a well known superluminal radio source. It was observed twice in 1986 August and no polarization was detected, the upper limits at B being 3.6% and 5.7% .

2251 + 158, 3CR 454.3

Another superluminal radio source and AS80 blazar, this object was observed once in 1986 August, 1987 July and 1987 September. No polarization was observed in the first two runs, but frequency independent polarization was observed on 1987 September 20 of $p = 3.8 \pm 0.3\%$. This position angle was oriented somewhat differently from the 10 GHz data of Inoue (personal communication) who measured $p \approx 3.2\%$ and $\theta \approx 11°$, on 1987 September 1987.

2254 + 074, OY 091

This AS80 blazar was extensively observed in 1986 August, 1987 July and 1987 September. The polarization behaviour was characterised by strong FDP with little evidence for FDθ in all three runs. On all dates except for 1986 August 6, the FDP was such that $dp/d\nu > 0$.

3.2 DISCUSSION.— The fits to the spectra of the 158 photometric data sets are shown in table II. These fits are separated into polarized and unpolarized samples. This is because it is only when polarizations in excess of the 3% limit are seen, that it is certain the observed flux contains a blazar contribution. The fits were all tested at the 5% significance level. If a power-law was rejected at this level,

then the spectral flux distribution is characterised as having either a convex spectrum ($d\alpha/d\nu > 0$), a concave spectrum ($d\alpha/d\nu < 0$) or a "complex" spectrum. The use of this level of significance is very lenient given the likely non-normal errors involved (it corresponds to a 1.96σ point of the normal distribution).

Table II lists all the polarimetric observations of blazars and whether they were fitted (at the 5% level of significance) by a constant polarization or FDP (either with $dp/d\nu > 0$, $dp/d\nu < 0$ or with more complex behaviour). In order to assign the FDP categories, a power-law polarization behaviour was fitted to these data ($p(\nu)\alpha\nu^{\beta}$). This was an empirical fit whose form was chosen as the polarization degree is contrained to be positive definite and this form allows an unconstrained minimisation to be performed with respect to log p_{14} and β. The two FDP categories ($dp/d\nu > 0$ and $dp/d\nu < 0$) were assigned on the basis of the sign of the fitted value of β. Those objects not fitted (at the 5% level) by this functional form, are described as having "complex" polarization behaviour. This procedure will flag behaviour, such as that seen in $1641 + 399$ and $2200 + 420$, where the FDP is only significant over a narro range of the observed frequencies.

Table II also presents the results of testing for frequenc dependence of the position angle of polarization. Th different categories were assigned using the same metho as for FDP.FDθ is seen to be rarer than FDP. There i no evidence suggesting any preference for clockwise ove counter-clockwise variations with frequency.

We have presented the results of our observing programme and described the operation of the MkII Hatfield polarimeter. Detailed discussion of individual events and the implications of the statistical properties of the sample is deferred to a further paper (Ballard *et al.*, 1989).

Acknowledgements.

We would like to thank the PATT committee for the generous allocation of telescope time awarded to this project and to the staff of UKIRT for their assistance during the observing program. K. Ballard and A. Mead wish to acknowledge the receipt of SERC studentships.

References

ANGEL J.R.P., STOCKMAN H.S. (AS80) : 1980, *Ann, Rev. Astron. Astrophys.* **18**, 321.
ANTONUCCI R.R.J., HICKSON P., MILLER J.S., OLSZEWSKI E.W. : 1987, *Astron. J.* **93**, 785.
BALLARD K.R., MEAD A.R.G., BRAND P.W.J.L., HOUGH J.H. : 1989, *Mon. Not. R. Astron. Soc.* submitted.
BEICHMAN C.A., NEUGEBAUER G., SOIFER B.T., WOOTTEN H.A., ROELLIG T., HARVEY P.M. : 1981, *Nature* **293**, 711.
BESSEL M.S. : 1979, *Publ. Astron. Soc. Pacific*, **91**, 589.
BREGMAN J.N., LEBOFSKY M.J., ALLER M.F., RIEKE G.H., ALLER H.D., HODGE P.E., GLASSGOLD A.E., HUGGINS P.J. : 1981, *Nature* **293**, 714.
BRINDLE C., HOUGH J.H., BAILEY J.A., AXON D.J., SCHULZ H., KIRUCHI S., MCGRAW J.T., WISNIEWSKI W.J., FONTAINE G., NADESU D., CLAYTON G., ANDERSON E., JAMESON R.F., SMITH R., WALLIS R.E. : 1985, *Mon. Not. R. Astron. Soc.* **214**, 619.
BRINDLE C., HOUGH J.H., BAILEY J.A., AXON D.J., HYLAND A.R. : 1986, *Mon. Not. R. Astron. Soc.* **221**, 739.
BURSTEIN D., HEILES C. : 1982, *Astron. J.* **87**, 1165.
FEIGELSON E.D., BRADT H., MCCLINTOCK J., REMILLARD R., URRY C.M., TAPIA S., GELDZAHLER B., JOHNSTON K., ROMANISHIN W., WEHINGER P.A., WYCKOFF S., MADEJSKI G., SCHWARTZ D.A., THORSTENSEN J., SCHAEFER B.E. : 1986, *Astrophys. J.* **302**, 337.
FRICKE K.J., KOLLATSCHNY W., WITZEL A. : 1983, *Astron. Astrophys.* **117**, 60.
HOLMES P.A., BRAND P.W.J.L., IMPEY C.D., WILLIAMS P.M. : 1984a, *Mon. Not. R. Astron. Soc.* **210**, 961.
HOLMES P.A., BRAND P.W.J.L., IMPEY C.D., WILLIAMS P.M., SMITH P., ELSTON R., BALONEK T., ZEILIK M., BURNS J., HECKERT P., BARVAINIS R., KENNY J., SCHMIDT G., PUSCHELL J. : 1984b, *Mon. Not. R. Astron. Soc.* **211**, 497.
IMPEY C.D., BRAND P.W.J.L., WOLSTENCROFT R.D., WILLIAMS P.M. : 1982, *Mon. Not. R. Astron. Soc.* **200**, 19.
IMPEY C.D., BRAND P.W.J.L., TAPIA S. : 1982, *Mon. Not. R. Astron. Soc.* **198**, 1.
IMPEY C.D., BRAND P.W.J.L., WOLSTENCROFT R.D., WILLIAMS P.M. : 1984, *Mon. Not. R. Astron. Soc.* **209**, 245.
IMPEY C.D., TAPIA S. : 1988, *Astrophys. J.* **333**, 666.
LANDOLT A.U. : 1983, *Astron. J.* **88**, 439.
MEISENHEIMER K., ROSER H.-J. : 1984, *Astrophys. J.* **279**, L39.
MEAD A.R.G., BRAND P.W.J.L., HOUGH J.H., BAILEY J.A. : 1988, *Mon. Not. R. Astron. Soc.* **233**, 503.
MOORE R.L., STOCKMAN H.S. : 1981, *Astrophys. J.* **243**, 60.
MOORE R.L., STOCKMAN H.S. : 1984, *Astrophys. J.* **279**, 465.
RIEKE G.H., LEBOFSKY M.J. : 1985, *Astrophys. J.* **288**, 618.
SERKOWSKI K. : 1974, *Methods. Exp. Phys.* **12 A** (Academic Press, New York) p. 361.
SMITH P.S., BALONEK T.J., ELSTON R., HECKERT P.A. : 1987, *Astrophys. J. Suppl. Ser.* **64**, 459.
ULMER M.P., BROWN R.L., SCHWARTZ D.A., PATTERSON J., CRUDDACE R.G. : 1983, *Astrophys. J.* **270**, L1.
WARDLE J.C.F., KRONBERG P.P. : 1974, *Astrophys. J.* **194**, 249.
WILKES B.J., WRIGHT A.E., JAUNCEY D.L., PETERSON B.A. : 1983, *Proc. Astron. Soc. Austr.* **5**, 2.
WILLS D., WILLS B.J. : 1981, *Nature* **289**, 384.
ZOTOV N.V., TAPIA S. : 1979, *Astrophys. J.* **229**, L5.

TABLE I.— *Effective wavelengths and zero-magnitude flux densities.*

Filter	λ/μm	$\nu/10^{14}$ Hz	S_o/kJy
K	2.18	1.38	0.667
H	1.64	1.83	1.075
J	1.23	2.44	1.603
I	0.80	3.75	2.55
R	0.64	4.68	3.08
V	0.55	5.45	3.64
B	0.44	6.81	4.26
U	0.36	8.33	1.81

TABLE II.— *Summary of the observations.* α_0 $(>0, <0)$ *refers to the spectral index constant (increasing, decreasing) with frequency. This is also indicated for polarization and position angle. Those observations not satisfactorily fitted are labelled complex.*

Object	E(B-v)	Date	Flux	Polarization	Position Angle
$0048-097OB-081$	0.00	86 Jul 31	α_0	$\Delta p>0$	θ_0
		86 Aug 3	$\Delta\alpha<0$	$\Delta p>0$	θ_0
		86 Aug 4	α_0	$\Delta p>0$	θ_0
		86 Aug 6	α_0	$\Delta p>0$	θ_0
		86 Aug 7	α_0	$\Delta p>0$	θ_0
		87 Jul 27	$\Delta\alpha<0$	p_0	θ_0
		87 Jul 29	α_0	p_0	θ_0
		87 Jul 30	α_0	p_0	θ_0
		87 Sep 18		p_0	θ_0
		87 Sep 19	α_0	p_0	θ_0
		87 Sep 20		p_0	θ_0
PKS 0106+013	0.00	86 Aug 4	α_0	unpolarized	
		87 Jul 28	$\Delta\alpha<0$	p_0	θ_0
		87 Sep 21	α_0	unpolarized	
GC 0109+224	0.06	86 Aug 1	α_0	p_0	θ_0
		86 Aug 4	α_0	$\Delta p<0$	$\Delta\theta<0$
		86 Aug 5		p_0	$\Delta\theta<0$
		86 Aug 6		$\Delta p>0$	θ_0
		86 Aug 7	α_0	$\Delta p>0$	θ_0
		87 Jul 27	α_0	p_0	θ_0
		87 Jul 30	$\Delta\alpha>0$	p_0	θ_0
		87 Sep 19	α_0	p_0	θ_0
		87 Sep 20		p_0	θ_0
		87 Sep 21	α_0	$\Delta p>0$	complex
0118−272	0.00	86 Aug 5	α_0	p_0	θ_0
		86 Aug 6		p_0	complex
		86 Aug 7	α_0	p_0	θ_0
		87 Jul 27	$\Delta\alpha<0$	$\Delta p>0$	θ_0
		87 Jul 30	α_0	$\Delta p>0$	θ_0
		87 Sep 20		complex	θ_0
		87 Sep 21	α_0	$\Delta p>0$	θ_0
0138−097	0.00	86 Aug 5	α_0	$\Delta p<0$	θ_0
		86 Aug 6		p_0	θ_0
		86 Aug 7	α_0	p_0	$\Delta\theta>0$
		87 Jul 28	α_0	p_0	θ_0
		87 Jul 30	α_0	p_0	θ_0
		87 Sep 19	α_0	$\Delta p>0$	θ_0
		87 Sep 20	α_0	$\Delta p>0$	θ_0
		87 Sep 21	α_0	p_0	θ_0

TABLE II (*continued*).

Object	E(B-v)	Date	Flux	Polarization	Position Angle
0219 − 164	0.00	87 Jul 28	α_0	p_0	θ_0
0219 + 428 3C 66A	0.09	86 Aug 1		p_0	$\Delta\theta > 0$
		86 Aug 4		$\Delta p > 0$	θ_0
		86 Aug 6		p_0	θ_0
		86 Aug 7		p_0	θ_0
		87 Jul 27	α_0	p_0	θ_0
		87 Jul 30	α_0	p_0	θ_0
		87 Sep 18		complex	θ_0
		87 Sep 19		p_0	θ_0
		87 Sep 20	α_0	p_0	θ_0
		87 Sep 21		p_0	θ_0
AO 0235 + 164	0.15	87 Jul 28	$\Delta\alpha > 0$	p_0	θ_0
		87 Sep 20	$\Delta\alpha > 0$	complex	θ_0
0300 + 470 4C 47.08	0.15	87 Sep 20		p_0	θ_0
0323 + 022	0.06	86 Aug 1			
		87 Sep 21		p_0	θ_0
0336 − 019	0.06	86 Aug 5		unpolarized	
0338 − 214	0.00	86 Aug 6		p_0	θ_0
		87 Sep 19		p_0	θ_0
0414 + 009	0.12	88 Feb 15	$\Delta\alpha > 0$	p_0	θ_0
		88 Feb 16	$\Delta\alpha > 0$	p_0	θ_0
0422 + 004	0.12	88 Feb 16	$\Delta\alpha > 0$	p_0	complex
		88 Feb 18	$\Delta\alpha > 0$	p_0	complex
PKS 0735 + 178	0.03	87 Sep 19		p_0	$\Delta\theta < 0$
		88 Feb 15	complex	p_0	complex
		88 Feb 16	α_0	p_0	θ_0
		88 Feb 18	α_0	p_0	complex
0736 + 017	0.12	88 Feb 15	complex	unpolarized	
		88 Feb 16	complex	unpolarized	
0754 + 100	0.00	88 Feb 15	$\Delta\alpha > 0$	$\Delta p > 0$	$\Delta\theta > 0$
		88 Feb 16	$\Delta\alpha > 0$	$\Delta p > 0$	complex
		88 Feb 17	α_0	p_0	θ_0
		88 Feb 18	α_0	p_0	θ_0
0818 − 128 OJ-131	0.09	88 Feb 15	complex	$\Delta p > 0$	θ_0
		88 Feb 16	$\Delta\alpha > 0$	$\Delta p > 0$	θ_0
		88 Feb 17	$\Delta\alpha > 0$	$\Delta p > 0$	θ_0
0851 + 202 OJ287	0.00	88 Feb 15	$\Delta\alpha > 0$	p_0	θ_0
		88 Feb 16	α_0	$\Delta p > 0$	θ_0
		88 Feb 17	α_0	$\Delta p > 0$	$\Delta\theta > 0$
0906 + 015	0.00	88 Feb 15	α_0	unpolarized	
1101 + 384	0.00	88 Feb 15	$\Delta\alpha > 0$	complex	θ_0
1147 + 245	0.00	88 Feb 18	α_0	$\Delta p > 0$	$\Delta\theta < 0$
1156 + 295	0.00	88 Feb 15	α_0	p_0	θ_0
		88 Feb 16	α_0	p_0	θ_0
		88 Feb 17	α_0	p_0	θ_0
1253 − 055 3C 279	0.03	86 Aug 1	α_0	p_0	θ_0
		86 Aug 2	α_0	$\Delta p > 0$	$\Delta\theta < 0$
		86 Aug 4	α_0	$\Delta p > 0$	complex
		86 Aug 5	α_0	$\Delta p > 0$	$\Delta\theta < 0$
		86 Aug 6	α_0	$\Delta p > 0$	$\Delta\theta < 0$
		87 Jul 28	$\Delta\alpha > 0$	$\Delta p > 0$	θ_0
		88 Feb 16	α_0	p_0	complex
		88 Feb 17	α_0	p_0	complex
		88 Feb 18	α_0	$\Delta p > 0$	θ_0
1413 + 135	0.03	86 Aug 6	α_0	unpolarized	
1418 + 546 OQ 530	0.03	86 Aug 4		$\Delta p > 0$	θ_0
		86 Aug 5		$\Delta p > 0$	complex

TABLE II (*continued*).

Object	E(B-v)	Date	Flux	Polarization	Position Angle
		86 Aug 6		complex	θ_0
		87 Jul 30	α_0	$\Delta p > 0$	complex
		87 Sep 20		$\Delta p > 0$	θ_0
		87 Sep 21		$\Delta p > 0$	θ_0
		88 Feb 16	$\Delta\alpha > 0$	complex	θ_0
		88 Feb 17	$\Delta\alpha > 0$	complex	θ_0
		88 Feb 18	$\Delta\alpha > 0$	p_0	$\Delta\theta > 0$
1424 + 240	0.00	88 Feb 16	complex	p_0	θ_0
		88 Feb 17	$\Delta\alpha > 0$	p_0	θ_0
		88 Feb 18	α_0	unpolarized	
PKS 1510 − 089	0.06	86 Aug 1	α_0	p_0	
1514 − 241 AP Libra	0.15	86 Aug 1		p_0	complex
		87 Jul 27	$\Delta\alpha > 0$	complex	θ_0
1538 + 149 4C 14.60	0.00	86 Aug 1	α_0	unpolarized	
1641 + 399 3C 345	0.00	86 Aug 1	$\Delta\alpha < 0$	complex	θ_0
		86 Aug 2	$\Delta\alpha < 0$	complex	θ_0
		86 Aug 4	$\Delta\alpha < 0$	complex	θ_0
		86 Aug 5	$\Delta\alpha < 0$	complex	θ_0
		86 Aug 6	complex	complex	θ_0
		86 Aug 7		$\Delta p < 0$	θ_0
		87 Jul 28	complex	unpolarized	
		87 Jul 30	$\Delta\alpha < 0$	unpolarized	
		87 Sep 19	$\Delta\alpha < 0$	unpolarized	
1652 + 398 Mkn 501	0.00	87 Sep 21		complex	$\Delta\theta < 0$
1717 + 178 OT 129	0.06	86 Aug 6	α_0	p_0	θ_0
		86 Aug 7		$\Delta p < 0$	θ_0
		87 Jul 30	α_0	unpolarized	
1727 + 502 I Zw 186	0.00	86 Aug 6		p_0	$\Delta\theta < 0$
		86 Aug 7		$\Delta p > 0$	θ_0
		87 Sep 21		$\Delta p > 0$	θ_0
1749 + 096 OT 081	0.15	86 Jul 31		p_0	$\Delta\theta > 0$
		86 Aug 5		p_0	θ_0
		86 Aug 6		p_0	θ_0
		86 Aug 7		$\Delta p > 0$	θ_0
		87 Jul 27	α_0	$\Delta p > 0$	θ_0
1921 − 293 OV-236	0.12	86 Aug 3		p_0	$\Delta\theta > 0$
		86 Aug 6		$\Delta p < 0$	θ_0
		86 Aug 7		p_0	θ_0
		87 Jul 27	complex	unpolarized	
2032 + 107	0.12	87 Jul 27	complex	unpolarized	
PKS 2155 − 304	0.00	86 Aug 1	α_0	$\Delta p > 0$	complex
		86 Aug 5	α_0	p_0	θ_0
		86 Aug 6		p_0	$\Delta\theta > 0$
		86 Aug 7	α_0	p_0	θ_0
		87 Jul 27	α_0	$\Delta p > 0$	$\Delta\theta < 0$
		87 Sep 21		complex	θ_0
2200 + 420 BL Lacertae	0.15	86 Jul 31		$\Delta p > 0$	θ_0
		86 Aug 1		p_0	θ_0
		86 Aug 3	$\Delta\alpha > 0$	complex	$\Delta\theta > 0$
		86 Aug 4		complex	$\Delta\theta > 0$
		86 Aug 5		complex	$\Delta\theta > 0$
		86 Aug 6		complex	θ_0
		86 Aug 7		complex	$\Delta\theta > 0$
		87 Jul 27	$\Delta\alpha > 0$	$\Delta p < 0$	$\Delta\theta < 0$
		87 Jul 28	$\Delta\alpha > 0$	complex	$\Delta\theta < 0$
		87 Jul 30	$\Delta\alpha > 0$	$\Delta p < 0$	$\Delta\theta < 0$
		87 Sep 19	$\Delta\alpha > 0$	complex	complex

TABLE II (*continued*).

Object	E(B-v)	Date	Flux	Polarization	Position Angle
		87 Sep 20		complex	$\Delta\theta > 0$
		87 Sep 21		complex	$\Delta\theta > 0$
PKS 2208 – 137	0.00	86 Aug 3		unpolarized	
		86 Aug 4		unpolarized	
		86 Aug 6	α_0	unpolarized	
		86 Aug 7		unpolarized	
2223 – 052 3C 446	0.03	86 Aug 4	α_0	unpolarized	
		87 Jul 27	α_0	p_0	θ_0
		87 Jul 29	α_0	p_0	θ_0
		87 Jul 30	α_0	p_0	θ_0
		87 Sep 20		$\Delta p > 0$	θ_0
		87 Sep 21		p_0	θ_0
2230 + 114 4C 11.69	0.03	86 Aug 3	α_0	unpolarized	
		86 Aug 4	α_0	unpolarized	
2251 + 158	0.06	86 Aug 5	α_0	unpolarized	
		87 Jul 28	$\Delta\alpha > 0$	unpolarized	
		87 Sep 20		p_0	θ_0
2254 + 074 OY 091	0.06	86 Jul 31		$\Delta p > 0$	θ_0
		86 Aug 1		p_0	θ_0
		86 Aug 3	complex	$\Delta p > 0$	θ_0
		86 Aug 4	$\Delta\alpha > 0$	complex	θ_0
		86 Aug 5		$\Delta p > 0$	θ_0
		86 Aug 6		$\Delta p > 0$	θ_0
		86 Aug 7	α_0	p_0	complex
		87 Jul 28	$\Delta\alpha > 0$	complex	θ_0
		87 Jul 30	$\Delta\alpha > 0$	complex	θ_0
		87 Sep 18		p_0	complex
		87 Sep 21		$\Delta p > 0$	θ_0

TABLE III.— *UKIRT data : 1986 July 31 - August 7, 1987 July 27-30, 1987 September 18-21, 1988 February 15-18.*

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
0048-097 OB-081			E(B-V) = 0.00			
1986 Jul. 31						
H	6.78	(0.63)	11.50	(1.80)	112.29	(4.85)
I	3.69	(0.34)	12.05	(0.52)	118.09	(1.28)
B	1.70	(0.24)	13.91	(0.44)	117.15	(0.83)
1986 Aug. 3						
H	9.63	(0.45)	13.17	(0.78)	116.34	(1.70)
J	5.26	(0.25)	15.09	(1.00)	117.58	(1.68)
I	3.01	(0.23)	14.55	(0.61)	118.13	(1.33)
R	2.34	(0.22)	16.02	(0.52)	116.38	(1.07)
V	1.74	(0.17)	16.59	(0.64)	118.54	(1.34)
B	1.62	(0.16)	16.90	(0.32)	116.45	(0.60)
1986 Aug 4						
H	6.54	(0.18)	14.80	(0.79)	111.71	(1.26)
J	4.71	(0.22)	14.31	(0.74)	112.37	(1.73)
I	3.12	(0.15)	15.67	(0.60)	111.29	(1.13)
V	1.86	(0.15)	17.42	(0.52)	113.68	(1.28)
B	1.73	(0.14)	17.85	(0.35)	111.50	(0.73)
1986 Aug. 6						
H	6.48	(0.30)	12.08	(0.88)	103.04	(1.75)
J	4.84	(0.23)	16.24	(0.68)	102.62	(1.39)
I	3.69	(0.34)	17.98	(0.60)	104.89	(1.02)
R	2.40	(0.23)	18.76	(0.49)	105.90	(0.78)
V	1.95	(0.15)	18.75	(0.57)	107.14	(1.32)
B	1.62	(0.16)	19.39	(0.38)	106.02	(0.57)
1986 Aug. 7						
H	6.60	(0.31)	12.89	(1.15)	106.69	(2.85)
I	3.15	(0.15)	15.79	(0.81)	107.36	(1.31)
B	1.66	(0.10)	18.12	(0.51)	108.82	(0.85)
1987 Jul. 27						
J	4.98	(0.32)	15.53	(1.30)	107.73	(1.95)
I	2.30	(0.11)	17.25	(0.82)	104.74	(1.27)
R	1.68	(0.09)	16.50	(0.80)	104.89	(1.22)
V	1.48	(0.08)	16.12	(0.86)	105.73	(2.43)
B	1.30	(0.08)	16.76	(0.71)	102.56	(1.04)
U	0.95	(0.10)	14.56	(0.75)	102.90	(1.46)
1987 Jul. 29						
K	9.64	(0.89)	19.34	(1.83)	117.44	(2.80)
H	6.24	(0.58)	21.72	(1.13)	117.82	(1.71)
I	2.55	(0.24)	21.17	(1.02)	116.85	(1.45)
R	1.94	(0.18)	21.32	(0.70)	118.82	(1.38)
V	1.59	(0.15)	20.54	(1.41)	118.50	(2.39)
B	1.17	(0.12)	22.69	(0.66)	118.26	(1.14)
U	0.79	(0.08)	23.03	(0.81)	117.66	(1.30)
1987 Jul. 30						
K	8.02	(0.44)	12.08	(1.53)	115.22	(3.67)
H	5.80	(0.11)	11.72	(0.98)	115.47	(2.50)
I	2.33	(0.11)	12.92	(0.59)	115.59	(1.71)
R	1.63	(0.08)	13.55	(0.57)	117.00	(1.32)
V	1.38	(0.07)	13.65	(0.86)	118.35	(2.12)
B	1.05	(0.06)	13.56	(0.57)	113.95	(1.56)
U	0.68	(0.07)	14.48	(1.16)	111.16	(2.42)
1987 Sep. 18						
H			5.40	(0.84)	81.26	(4.09)
I			3.64	(0.76)	78.81	(5.89)
R			4.52	(0.50)	88.50	(3.03)
V			3.60	(1.26)	87.11	(6.75)
B			4.09	(0.41)	94.10	(3.04)
U			3.12	(0.73)	95.17	(9.28)
1987 Sep. 19						
H	8.15	(0.75)	4.52	(0.60)	94.86	(4.07)
J	5.31	(0.49)	4.50	(0.75)	90.39	(4.22)
I	2.80	(0.26)	5.35	(0.50)	97.69	(5.21)
R	2.34	(0.22)	5.93	(0.30)	90.39	(1.37)
V	1.91	(0.18)	6.18	(0.70)	94.85	(2.29)
B	1.55	(0.15)	5.76	(0.44)	92.14	(1.76)
U	0.95	(0.10)	5.71	(0.51)	92.43	(2.25)
1987 Sep. 20						
H			3.18	(0.60)	90.70	(4.31)
I			4.17	(0.54)	86.53	(3.93)
R			4.15	(0.41)	89.69	(2.50)
V			2.46	(0.90)	83.31	(9.45)
B			5.23	(0.50)	88.92	(2.50)
U			4.08	(0.72)	94.61	(5.36)
PKS 0106+013			E(B-V) = 0.00			
1986 Aug. 4						
H	0.43	(0.06)	33.15	(18.62)	127.96	(15.16)
I	0.23	(0.02)	0.00	(7.84)		
B	0.20	(0.02)	12.28	(5.26)	122.37	(8.54)
1987 Jul. 28						
H	0.65	(0.03)	0.00	(13.02)		
I	0.29	(0.02)	17.44	(5.83)	138.64	(8.52)
R	0.19	(0.01)	11.30	(5.55)	146.86	(15.43)
V	0.17	(0.02)	0.00	(8.75)		
B	0.16	(0.01)	7.12	(3.39)	146.48	(9.67)
U	0.11	(0.01)	14.37	(2.39)	130.09	(5.03)
1987 Sep. 21						
H	0.43	(0.10)	36.73	(13.65)	119.33	(9.98)
I	0.28	(0.04)	0.00	(40.86)		
R	0.19	(0.02)	5.64	(4.43)	95.11	(24.59)
V	0.21	(0.03)	0.00	(10.99)		
B	0.15	(0.01)	0.54	(4.57)	56.51	(29.35)
U			0.00	(7.61)		
GC 0109+224			E(B-V) = 0.06			
1986 Aug. 1						
H	7.09	(0.33)	9.06	(1.20)	40.95	(3.55)
I	3.10	(0.29)	10.05	(0.43)	35.64	(1.30)
B	2.14	(0.21)	10.88	(0.34)	34.36	(1.04)
1986 Aug. 4						
H	7.02	(0.26)	5.66	(0.51)	54.70	(4.63)
I	3.07	(0.15)	3.81	(0.50)	42.74	(5.81)
B	1.26	(0.13)	4.13	(0.37)	31.15	(3.22)
1986 Aug. 5						
H	6.96	(0.32)	4.80	(1.90)	61.81	(7.07)
I	3.07	(0.15)	2.70	(0.45)	48.36	(4.71)
B	1.13	(0.07)	1.64	(0.49)	29.07	(7.95)
1986 Aug. 6						
H	6.58	(0.31)	0.00	(1.17)		
I	3.18	(0.15)	0.96	(0.63)	24.63	(3.43)
B	1.35	(0.08)	3.74	(0.47)	13.16	(3.96)
1986 Aug. 7						
H	6.58	(0.31)	1.58	(0.87)	4.06	(8.23)
J	4.70	(0.22)	3.41	(1.06)	22.18	(8.68)
I	2.90	(0.14)	4.95	(0.64)	17.54	(3.90)
R	2.13	(0.16)	6.35	(0.58)	20.31	(2.50)

TABLE III (*continued*).

Filter	Flux density (mJy)	Polarization (%)	Position angle (deg)
V	1.61 (0.16)	7.53 (0.59)	21.90 (2.14)
B	1.29 (0.08)	8.63 (0.30)	17.20 (1.03)
1987 Jul. 27			
H	4.81 (0.22)	13.63 (1.10)	100.66 (2.25)
I	1.77 (0.08)	14.04 (0.90)	103.58 (1.57)
R	1.19 (0.06)	13.79 (0.60)	104.61 (1.74)
V	0.94 (0.05)	15.04 (1.67)	107.92 (2.04)
B	0.73 (0.04)	15.24 (0.60)	105.83 (1.48)
U	0.51 (0.05)	14.15 (0.99)	107.34 (2.19)
1987 Jul. 30			
H	4.86 (0.23)	9.37 (1.12)	109.42 (3.73)
I	1.92 (0.09)	8.91 (0.85)	108.91 (2.25)
R	1.29 (0.07)	9.51 (0.49)	109.15 (2.28)
V	1.01 (0.05)	9.08 (0.94)	104.22 (3.96)
B	0.72 (0.04)	9.12 (0.57)	110.09 (1.60)
U	0.43 (0.03)	8.35 (1.04)	113.16 (4.37)
1987 Sep. 19			
H	7.22 (0.67)	4.12 (0.70)	117.86 (5.10)
I	2.58 (0.24)	3.92 (0.69)	124.48 (3.95)
R	1.88 (0.18)	4.02 (0.40)	115.87 (2.84)
V	1.34 (0.13)	3.89 (1.22)	123.55 (8.53)
B	1.03 (0.10)	3.66 (0.54)	121.28 (4.32)
U	0.67 (0.07)	3.88 (0.81)	123.90 (5.81)
1987 Sep. 20			
H		13.16 (1.03)	107.82 (2.92)
I		11.64 (1.16)	105.58 (2.25)
R		15.01 (0.98)	107.98 (1.41)
V		15.90 (1.66)	110.97 (3.01)
B		14.37 (0.98)	112.02 (1.39)
U		17.26 (1.81)	109.32 (1.80)
1987 Sep. 21			
H	5.48 (0.51)	13.87 (0.74)	105.70 (1.51)
I	1.95 (0.18)	13.81 (0.57)	103.46 (1.54)
R	1.36 (0.13)	15.29 (0.50)	100.78 (0.76)
V	1.02 (0.10)	17.27 (1.37)	98.60 (2.35)
B	0.71 (0.07)	16.44 (0.64)	105.38 (1.14)
U		16.14 (0.82)	102.34 (1.79)
0118-272		E(B-V) = 0.00	
1986 Aug. 5			
H	8.15 (0.38)	15.37 (0.77)	147.79 (1.46)
J	5.98 (0.28)	16.26 (0.91)	148.01 (1.75)
I	4.00 (0.19)	17.82 (0.58)	147.24 (0.97)
R	2.81 (0.14)	17.87 (0.47)	149.97 (0.81)
V	2.19 (0.12)	17.33 (0.63)	149.19 (1.49)
B	1.86 (0.19)	17.21 (0.35)	149.90 (0.59)
1986 Aug. 6			
H	8.08 (0.37)	16.63 (1.02)	148.71 (1.85)
I	3.97 (0.19)	17.50 (0.53)	148.14 (0.89)
B	1.49 (0.09)	17.56 (0.76)	154.44 (1.33)
1986 Aug. 7			
H	8.38 (0.39)	15.96 (1.42)	149.48 (1.89)
I	3.90 (0.19)	16.49 (0.68)	148.94 (1.20)
B	2.06 (0.12)	17.74 (0.65)	151.60 (1.01)
1987 Jul. 27			
H	9.02 (0.42)	12.53 (0.71)	25.62 (1.73)
I	2.65 (0.13)	12.96 (1.15)	29.58 (2.29)
R	2.25 (0.11)	13.92 (0.98)	27.80 (1.83)

Filter	Flux density (mJy)	Polarization (%)	Position angle (deg)
V	2.09 (0.11)	14.18 (0.86)	24.41 (1.92)
B	1.89 (0.11)	15.90 (0.57)	26.59 (1.23)
U	1.37 (0.14)	16.02 (1.08)	28.05 (1.97)
1987 Jul. 30			
K	11.48 (0.53)	12.31 (1.15)	28.51 (2.85)
H	8.62 (0.40)	12.82 (0.86)	29.04 (2.20)
I	3.62 (0.17)	14.19 (0.53)	30.66 (1.19)
R	2.68 (0.14)	13.58 (0.50)	30.47 (1.36)
V	2.25 (0.12)	13.94 (0.64)	27.20 (1.85)
B	1.78 (0.11)	15.03 (0.39)	28.54 (1.08)
U	1.44 (0.15)	15.01 (0.39)	27.58 (1.48)
1987 Sep. 20			
H		14.68 (0.54)	150.70 (1.26)
I		16.77 (0.68)	150.14 (1.04)
R		16.21 (0.56)	150.42 (0.73)
V		18.69 (1.01)	151.92 (1.37)
B		18.92 (0.99)	149.22 (0.72)
U		15.23 (1.37)	150.95 (1.84)
1987 Sep. 21			
H	7.44 (0.69)	15.25 (0.66)	148.50 (1.39)
I	3.07 (0.29)	16.00 (0.72)	149.91 (1.30)
R	2.34 (0.22)	15.81 (0.47)	150.33 (0.79)
V	1.91 (0.18)	17.94 (0.96)	152.33 (1.42)
B	1.48 (0.15)	17.95 (0.53)	151.56 (0.64)
U		16.54 (0.88)	150.45 (1.57)
0138-097		E(B-V) = 0.00	
1986 Aug. 5			
H	2.96 (0.22)	10.06 (2.02)	66.88 (6.31)
I	1.18 (0.06)	6.00 (1.08)	63.95 (4.93)
B	0.45 (0.04)	4.24 (1.02)	72.79 (6.52)
1986 Aug. 6			
H	2.82 (0.16)	7.68 (2.25)	64.73 (7.67)
I	1.16 (0.06)	7.63 (3.62)	73.96 (3.67)
B	0.30 (0.02)	5.31 (1.21)	76.10 (6.29)
1986 Aug. 7			
H	3.13 (0.14)	10.35 (2.65)	61.06 (5.41)
I	1.14 (0.05)	3.67 (1.38)	70.41 (8.77)
B	0.46 (0.03)	4.45 (1.24)	92.49 (7.18)
1987 Jul. 28			
H	3.49 (0.16)	14.76 (3.03)	104.72 (3.90)
I	1.34 (0.06)	19.09 (1.33)	103.14 (3.28)
R	1.03 (0.05)	20.56 (1.24)	107.72 (1.83)
V	0.87 (0.05)	19.95 (1.65)	104.91 (2.60)
B	0.71 (0.04)	19.99 (1.01)	104.83 (1.33)
U	0.50 (0.05)	20.88 (1.41)	103.47 (1.92)
1987 Jul. 30			
H	3.07 (0.14)	21.05 (2.94)	104.76 (3.58)
I	1.35 (0.06)	22.86 (2.04)	99.90 (2.08)
R	1.01 (0.05)	18.38 (1.81)	106.11 (2.08)
V	0.85 (0.05)	20.71 (2.38)	105.77 (2.19)
B	0.69 (0.05)	22.29 (1.45)	102.91 (2.26)
U	0.43 (0.04)	20.18 (1.73)	106.91 (2.83)
1987 Sep. 19			
K	4.21 (0.39)	21.61 (2.98)	74.36 (2.34)
H	3.40 (0.31)	20.60 (1.62)	73.96 (2.26)
I	1.34 (0.12)	24.80 (1.08)	71.11 (1.29)
R	1.02 (0.10)	25.86 (0.67)	72.81 (0.77)
V	0.83 (0.08)	26.04 (1.22)	74.30 (1.39)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
B	0.62	(0.06)	29.25	(0.92)	71.62	(0.83)
U	0.38	(0.04)	26.81	(1.35)	74.89	(1.49)
1987 Sep. 20						
H	2.96	(0.27)	22.23	(1.34)	70.40	(1.66)
I	1.61	(0.15)	21.45	(0.99)	74.33	(1.34)
R	1.23	(0.12)	24.51	(0.66)	73.32	(0.81)
V	1.00	(0.10)	27.33	(1.59)	73.54	(1.66)
B	0.81	(0.08)	27.07	(0.87)	73.53	(0.89)
U			27.05	(1.48)	74.69	(1.58)
1987 Sep. 21						
H	2.96	(0.27)	23.95	(1.44)	72.70	(1.72)
I	1.47	(0.14)	23.53	(1.41)	73.40	(1.80)
R	1.12	(0.11)	24.75	(0.96)	72.79	(1.12)
V	0.91	(0.09)	23.08	(1.82)	71.63	(2.31)
B	0.71	(0.07)	26.63	(1.07)	73.13	(1.13)
U			24.95	(1.47)	76.01	(1.70)
0219-164			E(B-V) = 0.00			
1987 Jul. 28						
K	13.31	(0.61)	14.83	(2.46)	160.96	(4.55)
H	11.15	(0.52)	13.67	(1.84)	170.47	(3.04)
J	7.67	(0.71)	12.85	(1.06)	159.53	(1.54)
I	5.48	(0.26)	12.45	(0.42)	159.96	(1.21)
R	4.14	(0.21)	12.65	(0.34)	161.33	(1.07)
V	3.64	(0.20)	12.37	(0.50)	160.92	(1.52)
B	3.00	(0.18)	12.70	(0.33)	161.16	(1.14)
U	2.10	(0.13)	12.83	(0.68)	160.19	(1.29)
0219+428 3C 66A			E(B-V) = 0.09			
1986 Aug. 1						
H	12.13	(0.56)	8.29	(0.71)	41.07	(2.60)
I	5.66	(0.53)	9.59	(0.24)	43.29	(0.69)
B	1.75	(0.17)	10.05	(0.31)	47.83	(0.78)
1986 Aug 4						
H	9.91	(0.37)	9.57	(0.60)	33.03	(1.75)
I	4.29	(0.21)	10.86	(0.38)	35.17	(0.84)
B	1.51	(0.12)	12.32	(0.53)	35.26	(1.13)
1986 Aug. 6						
H	9.91	(0.46)	9.41	(1.01)	30.80	(3.16)
I	4.80	(0.23)	9.24	(0.43)	34.25	(1.18)
B	1.95	(0.12)	10.85	(0.64)	35.60	(1.45)
1986 Aug. 7						
H	10.47	(0.49)	11.08	(0.74)	30.28	(1.90)
I	4.54	(0.22)	10.19	(0.34)	32.79	(1.05)
B	1.80	(0.11)	10.51	(0.33)	34.41	(1.07)
1987 Jul. 27						
H	11.17	(0.52)	12.53	(0.61)	164.20	(1.65)
I	3.92	(0.19)	11.20	(1.02)	162.63	(1.93)
R	2.47	(0.15)	12.52	(0.72)	162.87	(1.98)
V	2.34	(0.13)	14.25	(0.69)	161.93	(1.70)
B	1.81	(0.14)	13.16	(0.62)	163.65	(1.63)
U	1.34	(0.14)	14.47	(1.02)	165.89	(2.02)
1987 Jul. 30						
H	9.55	(0.44)	13.30	(0.99)	159.89	(2.39)
I	4.58	(0.22)	11.72	(1.34)	158.10	(2.97)
R	3.29	(0.17)	12.54	(0.88)	159.94	(2.13)
V	2.64	(0.14)	13.93	(0.76)	164.56	(1.63)
B	2.02	(0.12)	14.39	(0.50)	162.82	(1.20)
U	1.34	(0.14)	12.15	(1.53)	158.73	(2.19)

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
1987 Sep. 18						
H			14.80	(0.49)	22.16	(0.98)
I			13.38	(0.78)	22.86	(1.71)
R			13.49	(0.43)	24.04	(0.89)
V			15.43	(0.73)	24.19	(0.37)
B			14.59	(0.73)	23.28	(1.17)
U			12.15	(1.07)	22.06	(2.16)
1987 Sep. 19						
H			14.11	(0.53)	22.86	(1.06)
I			14.19	(0.80)	24.47	(1.69)
R			13.73	(0.38)	24.06	(0.76)
V			15.14	(0.74)	26.87	(1.48)
B			15.29	(0.53)	24.75	(0.95)
U			13.98	(0.61)	24.79	(1.18)
1987 Sep. 20						
H	11.27	(1.04)	14.11	(0.43)	23.50	(0.95)
I	4.71	(0.44)	15.46	(0.55)	26.52	(1.05)
R	3.68	(0.35)	14.30	(0.37)	25.95	(0.76)
V	3.09	(0.30)	15.94	(0.69)	25.80	(1.35)
B	2.30	(0.23)	14.38	(0.46)	26.50	(0.96)
U			14.25	(0.64)	28.47	(1.39)
1987 Sep. 21						
H			12.50	(0.55)	24.90	(1.25)
I			12.82	(0.69)	27.53	(1.54)
R			13.31	(0.37)	27.40	(0.84)
V			14.26	(0.60)	28.07	(1.21)
B			13.43	(0.35)	28.92	(0.73)
U			12.10	(0.68)	30.04	(1.64)
AO 0235+164			E(B-V) = 0.15			
1987 Jul. 28						
K	11.79	(0.54)	12.05	(1.04)	42.55	(2.92)
H	6.97	(0.32)	11.77	(1.02)	45.41	(3.25)
J	3.26	(0.30)	14.73	(4.42)	33.01	(9.88)
I	0.96	(0.05)	15.52	(1.59)	49.44	(3.04)
R	0.41	(0.02)	10.18	(1.58)	56.00	(3.80)
V	0.22	(0.01)	14.52	(3.67)	50.30	(9.31)
B	0.09	(0.01)	12.08	(3.73)	43.43	(8.36)
U	0.03	(0.00)	33.02	(10.17)	47.91	(8.59)
1987 Sep. 20						
K	5.80	(0.54)	1.66	(1.90)	47.71	(18.00)
H	3.97	(0.37)	5.78	(0.95)	52.70	(4.58)
J	2.47	(0.23)	8.59	(1.26)	56.59	(4.42)
I	0.99	(0.09)	11.04	(0.83)	55.91	(1.59)
R	0.56	(0.05)	10.10	(0.70)	55.26	(1.84)
V	0.31	(0.03)	10.00	(2.24)	57.95	(5.46)
B	0.14	(0.01)	11.65	(1.34)	60.67	(3.26)
U			8.28	(3.42)	68.16	(9.88)
0300+470 4C 47.08			E(B-V) = 0.15			
1987 Sep. 20						
H			9.55	(1.16)	9.00	(3.39)
I			9.44	(1.52)	1.76	(4.48)
R			8.69	(1.02)	13.96	(2.97)
V			5.05	(4.91)		
B			6.58	(2.12)	8.84	(8.92)
U			7.53	(5.37)	4.07	(11.70)
0323+022			E(B-V) = 0.06			
1986 Aug. 1						
H	2.16	(0.14)	0.00	(7.42)		

TABLE III (*continued*).

Filter	Flux density (mJy)	Polarization (%)	Position angle (deg)
I	1.03 (0.10)	0.00 (2.09)	
B	0.31 (0.03)	0.00 (2.25)	
1987 Sep. 21			
H		4.65 (1.63)	8.60 (6.61)
I		3.84 (1.02)	3.53 (8.40)
R		3.78 (0.70)	14.46 (5.13)
V		0.00 (1.84)	
B		3.49 (0.90)	
U		5.27 (0.86)	11.59 (6.65)
0336-019		E(B-V) = 0.06	
1986 Aug. 5			
H	0.76 (0.13)		
I	0.32 (0.02)		
B	0.10 (0.01)		
0338-214		E(B-V) = 0.00	
1986 Aug. 6			
H	4.01 (0.22)	11.39 (1.92)	31.92 (4.61)
I	1.47 (0.07)	10.45 (1.40)	29.31 (3.75)
B	0.29 (0.03)	10.60 (1.47)	37.32 (7.71)
1987 Sep. 19			
H		11.48 (1.60)	70.46 (3.69)
I		11.07 (1.56)	62.59 (3.88)
R		10.62 (0.95)	65.86 (2.00)
V		15.58 (2.90)	65.05 (6.33)
B		12.37 (1.13)	64.02 (3.21)
U		8.15 (3.39)	60.82 (11.61)
0414+009		E(B-V) = 0.12	
15th. Feb 1988			
H	1.63 (0.15)	2.29 (2.07)	158.56 (24.65)
J	1.19 (0.04)	4.14 (1.92)	174.45 (10.58)
I	0.68 (0.02)	4.24 (1.64)	159.87 (9.55)
R	0.46 (0.01)	3.87 (0.66)	153.22 (4.61)
V	0.32 (0.01)	5.08 (1.43)	170.57 (7.70)
B	0.23 (0.01)	6.41 (0.79)	157.69 (3.60)
U	0.15 (0.01)	4.94 (1.08)	153.76 (6.11)
16th. Feb 1988			
H	1.56 (0.12)	6.20 (2.01)	146.42 (1.38)
I	0.70 (0.03)	7.82 (2.25)	154.03 (8.43)
R	0.45 (0.02)	5.72 (1.08)	155.71 (5.44)
V	0.33 (0.01)	4.73 (1.89)	151.40 (9.96)
B	0.23 (0.01)	3.93 (0.97)	157.80 (7.44)
U	0.15 (0.01)	7.38 (1.44)	159.53 (6.16)
0422+004		E(B-V) = 0.12	
16th. Feb 1988			
H	7.47 (0.28)	20.27 (0.64)	155.80 (1.04)
I	2.60 (0.10)	20.79 (1.12)	148.91 (1.38)
R	1.71 (0.06)	21.84 (0.65)	151.77 (0.81)
V	1.22 (0.05)	23.30 (1.07)	153.20 (1.50)
B	0.78 (0.04)	21.47 (0.75)	151.76 (0.91)
U	0.47 (0.03)	22.14 (0.97)	150.93 (1.47)
18th. Feb 1988			
K	10.56 (1.07)	12.32 (0.73)	147.60 (1.09)
H	7.75 (1.15)	11.25 (0.87)	148.00 (1.38)
I	2.93 (0.19)	12.55 (0.67)	141.90 (1.51)
R	1.74 (0.07)	12.25 (0.34)	140.42 (0.69)
V	1.25 (0.08)	12.93 (0.69)	140.24 (1.34)
B	0.79 (0.06)	12.48 (0.48)	136.95 (0.96)
U	0.44 (0.03)	13.69 (0.83)	141.36 (1.44)
PKS 0735+178		E(B-V) = 0.03	
1987 Sep. 19			
H		6.64 (0.68)	144.76 (2.88)
I		8.08 (0.88)	132.34 (3.03)
R		6.77 (0.61)	137.22 (2.85)
V		7.44 (1.52)	131.02 (5.49)
B		6.59 (0.93)	130.39 (3.66)
U		5.68 (1.59)	129.65 (8.10)
15th. Feb 1988			
H	7.96 (0.22)	16.24 (0.54)	131.37 (1.06)
I	2.40 (0.07)	15.48 (1.02)	125.56 (1.64)
R	1.84 (0.05)	17.17 (0.47)	130.41 (0.65)
V	1.33 (0.04)	18.05 (1.00)	127.83 (1.71)
B	0.96 (0.04)	16.80 (0.56)	130.68 (1.15)
U	0.62 (0.03)	17.57 (1.03)	136.73 (1.90)
16th. Feb 1988			
H	8.41 (0.31)	15.23 (0.61)	130.73 (0.94)
I	2.73 (0.11)	17.66 (0.82)	132.51 (1.44)
R	2.00 (0.07)	17.90 (0.51)	129.97 (0.79)
V	1.46 (0.06)	17.19 (1.06)	130.60 (1.32)
B	1.05 (0.05)	17.49 (0.73)	130.98 (0.99)
U	0.67 (0.04)	17.94 (1.26)	132.36 (1.66)
18th. Feb 1988			
H	8.65 (1.36)	20.50 (0.89)	127.25 (0.78)
I	3.14 (0.44)	21.90 (0.98)	126.70 (1.14)
R	2.19 (0.31)	21.15 (0.68)	126.60 (0.73)
V	1.66 (0.27)	22.50 (1.00)	128.80 (1.11)
B	1.21 (0.20)	22.50 (0.82)	124.00 (0.89)
U	0.80 (0.14)	21.80 (0.93)	125.90 (1.11)
0736+017		E(B-V) = 0.12	
15th. Feb 1988			
J	2.20 (0.06)	1.53 (0.93)	105.58 (17.81)
I	1.66 (0.05)	2.98 (1.24)	172.76 (9.01)
R	0.96 (0.03)	0.77 (0.67)	7.56 (18.61)
V	0.75 (0.03)	0.49 (1.00)	152.87 (64.32)
B	0.60 (0.03)	0.22 (0.64)	73.87 (83.42)
U	0.46 (0.03)	1.28 (0.78)	107.22 (19.18)
16th. Feb 1988			
J	2.16 (0.10)	4.43 (2.04)	171.13 (11.22)
I	1.50 (0.05)	0.53 (1.39)	46.94 (83.32)
R	0.82 (0.03)	1.53 (0.93)	91.61 (16.10)
V	0.67 (0.03)	2.89 (1.92)	9.21 (18.86)
B	0.52 (0.02)	0.77 (0.81)	15.05 (28.29)
U	0.40 (0.02)	0.96 (1.25)	56.86 (40.83)
0754+100		E(B-V) = 0.00	
15th. Feb 1988			
K	14.33 (0.40)	8.58 (0.49)	40.50 (1.37)
H	10.17 (0.29)	10.56 (0.42)	41.31 (0.81)
I	4.08 (0.09)	11.22 (0.51)	43.20 (1.00)
R	2.89 (0.08)	12.33 (0.24)	44.15 (0.58)
V	2.15 (0.07)	12.22 (0.43)	46.12 (0.92)
B	1.62 (0.07)	13.39 (0.29)	46.39 (0.66)
U	1.07 (0.06)	14.39 (0.48)	48.43 (0.87)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
16th. Feb 1988						
K	13.81	(0.38)	10.34	(0.38)	34.40	(1.28)
H	9.63	(0.27)	10.96	(0.42)	35.20	(1.23)
I	3.62	(0.08)	12.39	(0.48)	30.20	(1.23)
R	2.27	(0.08)	12.59	(0.32)	34.74	(0.76)
V	1.86	(0.06)	13.22	(0.57)	35.45	(1.24)
B	1.36	(0.06)	13.10	(0.42)	35.69	(0.81)
U	0.91	(0.05)	13.25	(0.62)	38.54	(1.30)
17th. Feb 1988						
H	9.71	(0.36)	11.41	(0.51)	34.80	(1.67)
I	3.69	(0.24)	11.09	(0.59)	34.60	(1.95)
R	2.36	(0.12)	11.80	(0.39)	34.70	(1.23)
V	1.88	(0.10)	10.07	(0.58)	37.20	(2.27)
B	1.40	(0.07)	11.52	(0.54)	36.30	(1.27)
U	0.93	(0.06)	12.09	(0.77)	39.60	(2.27)
18th. Feb 1988						
K	14.73	(1.22)	8.67	(0.69)	36.20	(1.41)
H	10.75	(1.59)	8.17	(0.67)	39.00	(1.25)
I	3.75	(0.12)	9.28	(0.67)	32.33	(2.06)
R	2.73	(0.14)	8.65	(0.32)	36.67	(0.93)
V	2.08	(0.13)	8.43	(0.62)	36.40	(1.34)
B	1.55	(0.12)	6.95	(0.40)	36.35	(1.49)
U	1.03	(0.08)	7.56	(0.63)	38.28	(2.36)
0818-128 OJ-131			E(B-V) = 0.09			
15th. Feb 1988						
K	3.02	(0.14)	17.13	(1.60)	82.16	(2.81)
H	2.19	(0.10)	18.66	(1.72)	85.68	(2.15)
I	0.84	(0.03)	24.73	(1.66)	79.09	(1.92)
R	0.47	(0.02)	22.45	(0.72)	80.44	(1.01)
V	0.34	(0.01)	25.33	(1.81)	82.30	(1.85)
B	0.23	(0.01)	24.09	(0.86)	79.49	(1.04)
U	0.13	(0.01)	24.92	(1.66)	81.51	(1.99)
16th. Feb 1988						
K	3.16	(0.15)	15.48	(1.82)	79.61	(2.72)
J	1.32	(0.07)	21.87	(1.88)	83.54	(2.54)
I	0.68	(0.03)	21.92	(2.07)	88.24	(2.70)
R	0.38	(0.01)	22.22	(1.02)	81.56	(1.23)
V	0.28	(0.01)	25.68	(2.32)	81.85	(2.44)
B	0.19	(0.01)	26.75	(1.45)	80.49	(1.56)
U	0.11	(0.01)	27.67	(2.29)	81.59	(2.27)
17th. Feb 1988						
J	3.47	(0.16)	17.41	(1.62)	78.40	(2.33)
R	0.97	(0.04)	21.66	(2.43)	83.30	(3.20)
V	0.66	(0.04)	19.54	(1.21)	79.86	(1.85)
B	0.48	(0.03)	23.19	(2.78)	85.80	(3.03)
U	0.13	(0.01)	23.65	(1.55)	81.00	(1.83)
0851+202 OJ287			E(B-V) = 0.00			
15th. Feb 1988						
H	6.60	(0.19)	17.33	(0.59)	61.99	(0.96)
I	2.50	(0.08)	18.36	(1.22)	57.81	(1.91)
R	1.65	(0.06)	17.20	(0.59)	60.17	(0.99)
V	1.26	(0.05)	17.80	(1.13)	60.25	(1.53)
B	0.96	(0.04)	18.49	(0.65)	59.57	(0.88)
U	0.63	(0.04)	15.90	(1.06)	63.12	(2.07)
16th. Feb 1988						
H	6.54	(0.30)	15.62	(0.64)	60.40	(1.12)
I	2.50	(0.12)	15.20	(1.02)	66.07	(1.90)
R	1.65	(0.08)	15.55	(0.63)	61.96	(1.22)
V	1.31	(0.06)	17.38	(0.99)	63.58	(1.53)
B	1.00	(0.06)	18.55	(0.72)	62.56	(1.16)
U	0.70	(0.04)	17.65	(1.02)	64.89	(1.54)
17th. Feb 1988						
K	10.10	(0.28)	18.81	(0.50)	64.20	(0.77)
H	6.85	(0.19)	18.90	(0.58)	64.80	(0.88)
I	2.60	(0.08)	18.06	(0.66)	63.88	(1.08)
R	1.82	(0.09)	19.72	(0.33)	65.78	(0.48)
V	1.37	(0.05)	18.43	(0.74)	67.90	(1.17)
B	1.05	(0.05)	20.14	(0.44)	66.77	(0.65)
U	0.71	(0.04)	20.83	(0.76)	69.26	(0.99)
0906+015			E(B-V) = 0.00			
15th. Feb 1988						
H	0.48	(0.42)	51.43	(51.89)	142.20	(16.30)
I	0.37	(0.04)	48.40	(13.40)	49.50	(10.30)
R	0.36	(0.02)	2.08	(3.92)	6.20	(56.00)
V	0.35	(0.02)	8.33	(5.13)	73.90	(14.90)
B	0.31	(0.02)	2.16	(2.90)	35.50	(47.30)
U	0.24	(0.02)	10.10	(5.38)	135.10	(10.20)
1101+384			E(B-V) = 0.00			
15th. Feb 1988						
K	50.13	(0.94)	3.64	(0.30)	43.91	(2.17)
H	46.07	(1.29)	3.22	(0.14)	40.85	(1.19)
J	36.72	(0.74)	3.50	(0.20)	44.60	(1.09)
I	22.83	(0.90)	3.73	(0.13)	42.21	(0.90)
R	18.22	(0.63)	3.86	(0.08)	42.12	(0.55)
V	13.97	(0.55)	3.82	(0.11)	40.55	(0.82)
B	11.31	(0.53)	4.68	(0.10)	41.83	(0.58)
U	7.55	(0.43)	4.64	(0.13)	42.26	(0.83)
1147+245			E(B-V) = 0.00			
18th. Feb 1988						
H	8.62	(0.16)	1.52	(0.40)		
J	6.04	(0.12)	1.82	(0.54)	177.50	(8.05)
I	3.33	(0.10)	2.71	(0.67)	168.71	(6.70)
R	2.52	(0.11)	2.50	(0.26)	162.87	(3.05)
V	1.91	(0.10)	2.99	(0.46)	160.49	(4.41)
B	1.52	(0.09)	4.29	(0.32)	150.46	(2.08)
U	1.02	(0.07)	4.34	(0.52)	153.51	(3.62)
1156+295			E(B-V) = 0.00			
15th. Feb 1988						
H	2.35	(0.13)	23.04	(1.63)	168.83	(1.86)
I	0.97	(0.05)	24.07	(2.38)	162.06	(2.82)
R	0.77	(0.03)	24.37	(0.90)	164.87	(1.11)
V	0.62	(0.02)	24.76	(1.66)	166.81	(1.85)
B	0.53	(0.02)	22.83	(0.80)	164.20	(1.00)
U	0.35	(0.02)	24.25	(1.33)	163.78	(1.55)
16th. Feb 1988						
H	2.49	(0.21)	24.15	(2.53)	164.95	(1.55)
I	1.01	(0.05)	26.42	(2.87)	168.59	(3.25)
R	0.77	(0.03)	23.85	(1.02)	164.46	(1.16)
V	0.63	(0.03)	19.35	(2.46)	161.88	(3.68)
B	0.54	(0.03)	21.88	(1.24)	164.78	(1.64)
U	0.35	(0.02)	25.06	(2.17)	168.22	(2.41)
17th. Feb 1988						
H	2.37	(0.11)	28.06	(1.58)	166.80	(1.55)
I	0.88	(0.04)	27.27	(2.08)	161.20	(2.32)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
R	0.71	(0.02)	24.99	(0.99)	164.90	(1.02)
V	0.57	(0.02)	22.54	(2.12)	162.20	(2.44)
B	0.48	(0.02)	23.08	(0.82)	164.60	(1.07)
U	0.31	(0.02)	24.87	(1.78)	168.10	(2.02)
1253-055 3C 279			E(B-V) = 0.03			
1986 Aug. 1						
H	25.17	(1.17)	28.19	(0.55)	120.52	(0.72)
I	11.71	(1.09)	29.17	(0.76)	120.62	(0.37)
B	6.03	(0.60)	30.03	(0.96)	119.88	(0.76)
1986 Aug. 2						
H	25.64	(1.42)	28.71	(0.59)	121.63	(0.60)
J	17.32	(1.28)	28.18	(1.06)	120.15	(0.39)
I	12.84	(1.20)	29.88	(0.36)	118.44	(0.47)
V	6.67	(0.64)	30.46	(0.79)	117.47	(0.55)
B	5.50	(0.55)	30.91	(0.28)	116.94	(0.28)
1986 Aug. 4						
H	23.82	(0.67)	33.22	(0.26)	125.35	(0.25)
J	17.64	(1.31)	34.66	(0.33)	126.97	(0.28)
I	12.15	(0.58)	36.57	(0.30)	125.01	(0.26)
R	8.49	(0.65)	37.70	(0.29)	127.32	(0.19)
V	7.38	(0.71)	38.17	(0.44)	126.12	(1.05)
B	5.55	(0.41)	39.24	(0.18)	126.09	(0.18)
U	3.64	(0.37)	41.30	(1.09)	127.66	(0.51)
1986 Aug. 5						
H	21.52	(1.00)	37.87	(0.60)	133.34	(0.40)
J	14.81	(0.69)	39.80	(0.54)	132.65	(0.28)
I	10.68	(1.00)	41.58	(0.55)	131.92	(0.30)
R	6.44	(0.61)	43.67	(2.02)	130.75	(0.52)
V	5.30	(0.51)	44.28	(0.97)	131.99	(1.05)
B	4.57	(0.46)	43.95	(0.78)	131.63	(0.31)
U	2.76	(0.40)	45.92	(0.98)	131.77	(0.62)
1986 Aug. 6						
H	17.10	(0.64)	38.13	(0.60)	136.70	(0.47)
J	12.09	(0.57)	38.53	(0.77)	136.30	(0.45)
I	8.10	(0.61)	40.51	(0.89)	136.29	(0.74)
R	6.15	(0.47)	43.07	(0.48)	135.52	(0.31)
V	5.75	(0.56)	43.31	(0.64)	136.14	(0.80)
B	3.70	(0.37)	44.36	(0.64)	134.73	(0.32)
U	2.68	(0.27)	45.47	(0.89)	135.51	(0.48)
1987 Jul. 28						
K	21.88	(2.02)	26.69	(0.77)	108.71	(1.15)
H	13.96	(1.29)	27.53	(0.68)	109.21	(1.12)
J	7.77	(0.72)	31.01	(0.84)	108.87	(1.17)
I	4.45	(0.42)	32.43	(0.77)	110.57	(0.95)
R	2.89	(0.27)	34.10	(0.61)	110.90	(0.94)
V	1.92	(0.19)	34.78	(1.14)	106.41	(1.14)
B	1.26	(0.13)	35.56	(1.43)	109.41	(1.11)
U	0.63	(0.06)	35.43	(2.29)	107.74	(1.50)
16th. Feb 1988						
H	29.07	(1.35)	21.36	(0.70)	114.44	(0.79)
I	12.55	(0.94)	20.94	(0.70)	110.07	(0.97)
R	9.22	(0.63)	20.08	(0.59)	111.38	(0.83)
V	7.00	(0.49)	19.58	(0.64)	109.58	(0.91)
B	5.17	(0.39)	19.47	(0.66)	109.65	(0.65)
U	3.29	(0.26)	19.65	(0.81)	112.30	(1.15)
17th. Feb 1988						
H	28.80	(0.55)	26.10	(0.20)	113.60	(0.24)
I	9.78	(0.30)	26.30	(0.42)	111.90	(0.46)

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
R	6.49	(0.23)	26.30	(0.36)	112.10	(0.39)
V	5.17	(0.20)	26.40	(0.53)	112.40	(0.58)
B	3.74	(0.17)	25.70	(0.42)	110.20	(0.38)
U	2.30	(0.12)	25.50	(0.67)	112.60	(0.69)
18th. Feb 1988						
H	32.17	(0.61)	30.12	(0.21)	113.80	(0.21)
I	10.53	(0.24)	31.30	(0.36)	113.70	(0.33)
R	7.12	(0.25)	31.30	(0.33)	113.30	(0.30)
V	5.56	(0.22)	31.50	(0.39)	114.10	(0.36)
B	3.96	(0.17)	30.70	(0.33)	113.30	(0.30)
U	2.57	(0.13)	30.70	(0.58)	114.90	(0.53)
1413+135			E(B-V) = 0.03			
1986 Aug. 6						
H	0.75	(0.11)				
I	0.06	(0.01)				
B	0.00	(0.01)				
1418+546 OQ 530			E(B-V) = 0.03			
1986 Aug. 4						
H	9.40	(0.26)	3.56	(0.64)	138.41	(3.62)
I	4.02	(0.19)	4.62	(0.31)	142.15	(2.03)
B	1.18	(0.10)	5.82	(0.46)	138.46	(2.28)
1986 Aug. 5						
H	9.48	(0.44)	4.55	(0.60)	140.42	(3.95)
I	3.23	(0.30)	4.83	(0.52)	149.43	(3.03)
B	2.89	(0.29)	6.41	(0.39)	138.75	(2.24)
1986 Aug. 6						
H	8.65	(0.40)	1.73	(0.78)	122.26	(11.51)
I	3.29	(0.16)	4.94	(0.64)	134.68	(3.63)
B	0.91	(0.05)	5.09	(0.70)	134.51	(3.88)
1987 Jul. 30						
H	8.49	(0.39)	2.50	(0.64)	54.32	(7.65)
J	5.79	(0.27)	4.12	(2.29)	41.04	(5.74)
I	2.94	(0.04)	3.69	(0.44)	52.71	(3.63)
R	2.00	(0.19)	5.17	(0.39)	64.37	(2.02)
V	1.46	(0.14)	5.71	(0.71)	64.43	(3.86)
B	1.05	(0.10)	7.31	(0.41)	65.58	(1.86)
U	0.69	(0.07)	8.68	(0.96)	60.93	(2.70)
1987 Sep. 20						
H			8.21	(0.99)	3.96	(2.42)
I			15.37	(2.82)	3.45	(4.35)
R			11.30	(2.23)	15.26	(5.02)
V			17.53	(4.76)	−0.66	(12.40)
B			0.00	(6.11)		
U			0.00	(35.16)		
1987 Sep. 21						
H			7.22	(1.46)	15.39	(7.78)
I			8.99	(1.48)	7.56	(5.26)
R			12.83	(1.12)	3.12	(2.95)
V			13.13	(3.13)	14.46	(4.09)
B			9.78	(2.52)	8.53	(6.39)
U			12.93	(7.57)	19.81	(15.08)
16th. Feb 1988						
K	19.96	(1.11)	2.49	(0.37)	134.85	(4.33)
H	16.88	(0.63)	2.90	(0.30)	142.21	(4.51)
J	12.27	(0.25)	3.57	(0.29)	143.80	(2.48)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
I	7.49	(0.36)	3.12	(0.31)	146.58	(3.09)
R	5.25	(0.27)	3.06	(0.18)	140.00	(1.87)
V	3.64	(0.20)	3.31	(0.39)	146.85	(3.50)
B	2.59	(0.15)	4.38	(0.23)	143.90	(1.55)
U	1.49	(0.10)	4.89	(0.50)	147.29	(2.56)
17th. Feb 1988						
K	17.38	(0.48)	4.09	(0.48)	142.10	(3.71)
H	12.92	(0.36)	4.04	(0.32)	145.50	(1.85)
J	8.97	(0.18)	4.06	(0.29)	138.90	(3.07)
I	4.81	(0.11)	4.48	(0.27)	142.63	(1.88)
R	3.38	(0.09)	4.51	(0.18)	142.82	(1.09)
V	2.34	(0.08)	5.05	(0.35)	146.93	(1.91)
B	1.70	(0.07)	6.28	(0.23)	144.21	(1.11)
U	0.96	(0.05)	5.88	(0.43)	147.41	(1.86)
18th. Feb 1988						
H	14.43	(0.41)	3.38	(0.25)	132.01	(2.63)
I	5.68	(0.13)	3.56	(0.51)	140.80	(4.60)
R	4.25	(0.12)	3.11	(0.25)	146.20	(2.27)
V	2.97	(0.10)	2.96	(0.55)	142.80	(5.86)
B	2.34	(0.10)	3.20	(0.38)	150.00	(3.49)
U	1.49	(0.08)	4.05	(0.64)	156.20	(3.86)
1424+240			E(B-V) = 0.00			
16th. Feb 1988						
H	9.28	(0.26)	4.86	(0.36)	119.82	(2.12)
I	5.74	(0.28)	4.98	(0.66)	121.54	(3.78)
R	4.21	(0.12)	4.70	(0.36)	113.04	(2.18)
V	3.81	(0.21)	3.46	(0.50)	116.21	(4.07)
B	3.17	(0.15)	4.63	(0.30)	113.05	(1.95)
U	1.81	(0.09)	4.32	(0.48)	115.97	(3.12)
17th. Feb 1988						
H	9.19	(0.26)	4.92	(0.40)	123.82	(2.24)
I	4.95	(0.11)	4.24	(0.53)	114.60	(3.52)
R	3.84	(0.13)	4.48	(0.31)	119.20	(1.86)
V	3.26	(0.13)	4.36	(0.38)	114.20	(2.33)
B	2.76	(0.13)	4.45	(0.24)	117.80	(1.88)
U	1.91	(0.11)	4.35	(0.58)	116.60	(3.38)
18th. Feb 1988						
H	9.36	(0.35)	5.01	(0.44)	122.58	(2.21)
I	4.95	(0.11)	5.14	(0.55)	120.30	(3.00)
R	4.06	(0.11)	4.93	(0.30)	118.30	(1.65)
V	3.38	(0.11)	4.99	(0.49)	116.30	(2.87)
B	2.87	(0.12)	4.56	(0.30)	114.20	(1.90)
U	2.12	(0.10)	4.42	(0.58)	113.20	(3.85)
PKS 1510-089			E(B-V) = 0.06			
1986 Aug. 1						
H	2.26	(0.13)	9.07	(5.34)	177.65	(7.87)
I	1.23	(0.11)	1.65	(2.09)		
B	0.71	(0.03)	0.00	(2.32)		
1514-241 AP Libra			E(B-V) = 0.15			
1986 Aug. 1						
H	15.96	(0.74)	1.70	(0.44)	177.65	(7.87)
I	6.87	(0.64)	1.86	(0.28)	3.27	(3.94)
B	1.16	(0.12)	2.38	(0.37)	170.79	(4.07)
1987 Jul. 27						
K	23.09	(0.07)	5.11	(0.72)	19.56	(4.28)
H	19.01	(1.76)	3.70	(0.62)	16.72	(3.85)
J	14.21	(1.32)	5.42	(0.44)	14.28	(2.05)
I	6.87	(0.64)	4.76	(0.30)	12.37	(1.61)
R	4.07	(0.39)	5.12	(0.24)	15.40	(1.35)
V	2.86	(0.28)	4.83	(0.34)	18.72	(2.25)
B	1.39	(0.14)	6.13	(0.23)	17.41	(1.22)
U	0.86	(0.04)	6.92	(0.65)	16.50	(2.56)
1538+149 4C 14.60			E(B-V) = 0.00			
1986 Aug. 1						
H	0.90	(0.14)	0.00	(16.62)		
I	0.34	(0.03)	10.99	(9.54)		
B	0.13	(0.01)	4.97	(7.54)		
1641+399 3C 345			E(B-V) = 0.00			
1986 Aug. 1						
H	4.40	(0.20)	16.37	(1.24)	51.41	(1.82)
J	2.43	(0.20)	14.52	(2.64)	43.60	(5.34)
I	1.45	(0.11)	13.85	(0.73)	54.84	(1.60)
R			12.80	(0.98)	55.96	(2.17)
V	0.91	(0.09)	11.32	(0.79)	52.63	(2.02)
B	0.98	(0.10)	7.85	(0.30)	52.76	(1.04)
1986 Aug. 2						
H	4.16	(0.23)	17.23	(0.95)	52.15	(1.35)
J	2.56	(0.14)	17.24	(1.02)	55.10	(1.39)
I	1.49	(0.11)	15.83	(0.82)	53.83	(1.46)
R	1.06	(0.08)	14.96	(0.66)	51.63	(1.40)
V	0.92	(0.07)	12.57	(0.79)	55.55	(1.72)
B	0.90	(0.05)	8.70	(0.29)	52.48	(0.90)
1986 Aug 4						
H	4.12	(0.19)	17.87	(0.78)	55.65	(0.99)
J	2.59	(0.12)	16.27	(0.79)	53.01	(1.35)
I	1.43	(0.07)	15.01	(0.68)	55.18	(1.34)
R	1.07	(0.04)	12.87	(0.51)	56.74	(1.13)
V	0.97	(0.09)	11.83	(0.46)	53.85	(1.36)
B	0.85	(0.07)	8.61	(0.57)	56.64	(1.67)
U	0.60	(0.06)	7.92	(0.67)	58.14	(1.13)
1986 Aug. 5						
K	8.02	(0.74)	17.93	(1.57)	63.05	(2.51)
H	4.16	(0.19)	19.56	(1.91)	55.29	(2.64)
I	1.39	(0.13)	16.09	(1.14)	59.64	(2.03)
B	0.89	(0.09)	7.73	(0.68)	57.20	(2.57)
1986 Aug. 6						
H	3.98	(0.18)	20.27	(1.30)	57.17	(1.96)
J	2.52	(0.12)	20.03	(1.23)	62.31	(1.75)
I	1.44	(0.07)	14.98	(0.72)	57.00	(1.39)
V	0.87	(0.03)	11.14	(0.81)	60.17	(2.30)
B	0.90	(0.04)	8.65	(0.38)	57.24	(1.22)
1986 Aug. 7						
H	4.05	(0.19)	18.59	(1.55)	60.72	(2.06)
I	1.45	(0.07)	14.39	(0.81)	58.34	(1.61)
B	0.91	(0.05)	9.60	(0.71)	58.48	(2.48)
1987 Jul. 28						
K	2.66	(0.12)	1.52	(2.19)	179.39	(17.65)
H	1.92	(0.09)	0.00	(1.93)		
J	1.15	(0.05)	2.32	(3.51)	72.37	(23.98)
I	0.58	(0.03)	2.32	(1.28)	50.69	(13.47)
R	0.47	(0.02)	3.41	(0.95)	123.12	(7.39)
V	0.43	(0.02)	3.47	(1.24)	119.39	(9.18)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
B	0.44	(0.03)	1.76	(0.46)	131.81	(7.22)
U	0.25	(0.02)	2.62	(0.81)	127.29	(8.41)
1987 Jul. 30						
H	1.69	(0.08)	2.74	(1.78)	30.62	(13.86)
I	0.58	(0.05)	0.86	(1.97)	131.75	(26.86)
R	0.47	(0.04)	1.61	(1.29)	127.34	(16.72)
V	0.44	(0.04)	0.00	(3.39)		
B	0.43	(0.04)	1.38	(0.77)	123.69	(12.96)
U	0.24	(0.02)	1.99	(1.09)	73.13	(13.91)
1987 Sep. 19						
H	1.55	(0.14)	4.02	(3.28)	169.86	(18.80)
I	0.49	(0.05)	11.15	(3.89)	82.71	(8.74)
R	0.37	(0.04)	0.00	(1.62)		
V	0.36	(0.04)	0.00	(1.47)		
B	0.39	(0.04)	0.00	(1.41)		
U	0.24	(0.03)	1.49	(2.13)	138.80	(26.06)
1652+398 Mkn 501			E(B-V) = 0.00			
1987 Sep. 21						
K			1.34	(0.26)	138.09	(4.81)
H			1.24	(0.19)	125.70	(4.83)
J			1.57	(0.20)	124.98	(3.85)
I			1.56	(0.16)	126.94	(2.52)
R			1.95	(0.11)	122.09	(1.55)
V			2.50	(0.18)	124.08	(2.06)
B			3.53	(0.14)	124.06	(1.18)
U			3.76	(0.22)	119.20	(1.62)
1717+178 OT 129			E(B-V) = 0.06			
1986 Aug. 6						
H	1.34	(0.06)	15.55	(3.97)	39.16	(8.07)
I	0.39	(0.02)	17.82	(3.62)	45.33	(5.08)
B	0.12	(0.01)	20.92	(3.67)	45.37	(4.54)
1986 Aug. 7						
H	1.13	(0.05)	18.59	(1.55)	60.72	(2.06)
I	0.37	(0.02)	14.39	(0.81)	58.34	(1.61)
B	0.10	(0.01)	9.60	(0.71)	58.48	(2.48)
1987 Jul. 30						
H	0.55	(0.10)	16.85	(11.90)	103.02	(13.86)
I	0.16	(0.02)	0.00	(11.88)		
R	0.10	(0.01)	6.33	(7.03)	150.64	(19.41)
V	0.08	(0.01)	10.37	(11.63)	34.78	(20.79)
B	0.04	(0.01)	7.11	(7.81)	134.50	(21.14)
U	0.02	(0.00)	0.00	(16.81)		
1727+502 I Zw 186			E(B-V) = 0.00			
1986 Aug. 6						
H	4.16	(0.19)	2.36	(1.55)	142.17	(15.26)
I	1.80	(0.09)	2.51	(0.82)	91.15	(7.55)
B	0.23	(0.01)	4.16	(0.88)	78.49	(6.95)
1986 Aug. 7						
H	4.20	(0.19)	0.43	(0.98)	56.35	(31.99)
I	2.08	(0.10)	2.44	(0.54)	91.88	(6.24)
B	0.65	(0.04)	4.67	(0.59)	92.73	(3.94)
1987 Sep. 21						
H			1.55	(1.62)	118.80	(20.70)
I			0.00	(0.84)		
R			0.92	(0.41)	95.85	(12.45)
V			1.10	(1.92)	126.27	(25.93)
B			3.14	(1.02)	96.91	(6.26)
U			5.96	(1.30)	110.38	(6.50)
1749+096 OT 081			E(B-V) = 0.15			
1986 Jul. 31						
H	4.35	(0.40)	16.70	(1.49)	153.42	(2.43)
I	1.44	(0.13)	16.53	(0.95)	165.69	(1.31)
B	0.32	(0.03)	18.41	(0.88)	169.93	(1.39)
1986 Aug. 5						
H	4.64	(0.22)	5.66	(1.85)	16.41	(8.04)
I	1.31	(0.06)	6.64	(1.13)	5.80	(4.53)
B	0.71	(0.04)	7.42	(1.18)	6.29	(4.29)
1986 Aug. 6						
H	5.00	(0.23)	10.92	(1.35)	27.54	(6.42)
I	1.48	(0.07)	9.62	(0.72)	35.81	(1.89)
B	0.38	(0.04)	9.76	(0.85)	34.36	(2.87)
1986 Aug. 7						
H	4.16	(0.19)	7.80	(1.15)		
I	1.27	(0.06)	8.16	(0.90)	3.99	(2.90)
B	0.30	(0.02)	10.87	(0.89)	1.85	(2.69)
1987 Jul. 27						
J	1.34	(0.09)	4.06	(2.23)	18.99	(16.94)
I	0.41	(0.02)	9.96	(3.87)	35.57	(12.25)
R	0.22	(0.01)	9.71	(4.01)	13.91	(10.32)
V	0.14	(0.01)	18.96	(6.76)		
B	0.07	(0.00)	23.85	(3.78)	0.71	(5.88)
U	0.04	(0.00)	13.87	(7.20)		
1921-293 OV-236			E(B-V) = 0.12			
1986 Aug. 3						
H	1.33	(0.12)	9.78	(3.71)	104.12	(12.02)
I	0.57	(0.05)	6.34	(4.13)	100.10	(17.85)
B	0.11	(0.02)	16.89	(4.53)	160.42	(7.29)
6th. Aug 1986						
H	5.08	(0.24)	13.94	(1.91)	126.06	(4.92)
I	1.54	(0.07)	6.98	(1.82)	117.72	(6.59)
B	0.43	(0.04)	5.76	(1.85)	109.18	(8.79)
1986 Aug. 7						
H	2.52	(0.12)	7.10	(2.03)	119.56	(7.91)
I	0.54	(0.03)	8.13	(1.65)	135.93	(6.78)
B	0.08	(0.01)	5.97	(3.04)	111.08	(13.03)
1987 Jul. 27						
J	1.11	(0.06)	6.01	(3.86)	85.18	(15.94)
I	0.79	(0.05)	3.91	(2.92)	130.94	(17.44)
R	0.69	(0.07)	5.60	(2.26)	105.55	(10.49)
V	0.19	(0.02)	0.00	(5.44)		
B	0.15	(0.01)	0.00	(3.32)		
2032+107			E(B-V) = 0.12			
1987 Jul. 27						
J	8.61	(0.40)	0.56	(0.47)		
I	4.36	(0.21)	0.86	(0.40)		
R	1.98	(0.19)	1.08	(0.44)		
V	1.03	(0.06)	1.38	(1.08)		

TABLE III (*continued*).

Filter	Flux density (mJy)	Polarization (%)	Position angle (deg)
B	0.28 (0.02)	0.00 (1.15)	
U	0.04 (0.00)	0.00 (7.18)	
PKS 2155-304		E(B-V) = 0.00	
1986 Aug. 1			
H	27.76 (1.29)	2.03 (0.30)	148.06 (3.24)
J	21.72 (1.02)	3.11 (0.36)	149.47 (4.93)
I	21.02 (1.58)	3.10 (0.12)	141.54 (1.18)
R	14.74 (1.40)	3.62 (0.08)	137.15 (0.76)
B	14.11 (1.41)	4.03 (0.07)	138.80 (0.53)
1986 Aug. 5			
H	30.44 (1.41)	1.24 (0.35)	
I	20.63 (0.99)	0.72 (0.19)	3.60 (8.10)
B	13.23 (0.79)	0.80 (0.19)	13.33 (7.97)
1986 Aug. 6			
H	28.02 (1.30)	0.73 (0.21)	
I	26.70 (2.49)	0.45 (0.16)	21.52 (9.23)
B	14.11 (1.41)	0.69 (0.21)	75.08 (7.70)
1986 Aug. 7			
H	29.88 (1.38)	1.85 (0.19)	147.57 (1.89)
J	26.12 (1.22)	1.67 (0.17)	149.75 (3.11)
I	20.82 (1.00)	1.94 (0.16)	149.14 (2.42)
R	18.90 (1.45)	2.03 (0.18)	145.32 (2.46)
V	15.60 (1.51)	2.06 (0.18)	146.78 (2.63)
B	13.47 (1.12)	2.11 (0.12)	142.71 (1.36)
1987 Jul. 27			
K	67.94 (3.14)	9.16 (0.24)	175.89 (1.11)
H	60.73 (2.81)	9.40 (0.70)	176.16 (1.51)
J	55.58 (3.61)	10.14 (0.97)	172.93 (1.01)
I	36.52 (2.74)	10.29 (0.23)	171.83 (1.07)
R	27.07 (2.56)	10.46 (0.23)	171.20 (1.07)
V	26.13 (2.06)	10.69 (0.21)	170.48 (1.03)
B	24.29 (1.81)	10.91 (0.15)	169.39 (0.94)
U	18.10 (1.85)	10.93 (0.16)	169.14 (0.96)
1987 Sep. 21			
K		7.14 (0.42)	170.19 (1.82)
H		7.60 (0.21)	170.00 (0.75)
I		8.33 (0.17)	170.41 (0.62)
R		8.52 (0.12)	169.67 (0.41)
V		9.05 (0.18)	169.96 (0.60)
B		8.61 (0.10)	169.91 (0.44)
U		8.35 (0.16)	171.62 (0.58)
2200+420 BL Lacertae		E(B-V) = 0.15	
1986 Jul. 31			
H	19.01 (1.76)	11.83 (0.48)	30.74 (1.17)
I	5.21 (0.49)	12.02 (0.50)	33.50 (0.76)
B	0.73 (0.07)	14.03 (0.65)	35.33 (1.14)
1986 Aug. 1			
H	21.04 (0.97)	12.52 (0.73)	26.56 (1.37)
I	3.96 (0.37)	11.36 (0.43)	27.93 (1.03)
B	0.61 (0.06)	12.57 (0.53)	30.78 (1.24)
1986 Aug. 3			
H	19.54 (0.91)	12.36 (0.24)	20.87 (0.81)
J	11.93 (0.56)	12.49 (0.24)	21.63 (0.57)
I	4.46 (0.33)	12.93 (0.34)	22.46 (0.72)
R	2.04 (0.19)	12.80 (0.42)	25.87 (1.06)
V	1.37 (0.13)	14.24 (0.45)	24.56 (1.32)
B	0.55 (0.06)	16.70 (0.39)	26.09 (0.71)
1986 Aug. 4			
H	18.32 (0.51)	12.24 (0.44)	17.50 (1.09)
I	4.54 (0.22)	13.34 (0.40)	19.62 (0.85)
B	0.61 (0.05)	17.17 (0.71)	22.74 (1.19)
1986 Aug. 5			
H	19.01 (0.88)	14.33 (0.36)	15.36 (0.58)
I	4.42 (0.21)	14.00 (0.43)	18.43 (0.94)
B	0.54 (0.03)	17.52 (0.74)	23.13 (1.22)
1986 Aug. 6			
H	15.96 (0.74)	14.92 (0.72)	19.39 (1.29)
I	5.21 (0.49)	14.15 (0.37)	22.31 (0.85)
B	0.63 (0.06)	17.24 (0.71)	23.17 (1.22)
1986 Aug. 7			
H	18.83 (0.87)	13.95 (0.38)	22.59 (0.83)
I	4.63 (0.22)	13.81 (0.40)	22.51 (0.89)
B	0.60 (0.04)	16.88 (0.68)	25.91 (1.12)
1987 Jul. 27			
K	28.53 (1.32)	10.42 (0.26)	16.50 (1.11)
H	20.84 (1.35)	9.60 (0.35)	16.44 (1.44)
J	13.32 (0.62)	8.52 (0.40)	16.12 (1.74)
I	5.12 (0.25)	7.83 (0.35)	15.57 (1.41)
R	3.12 (0.16)	8.40 (0.23)	14.93 (1.12)
V	1.60 (0.11)	8.50 (0.56)	10.96 (1.96)
B	0.67 (0.07)	8.30 (0.44)	10.56 (1.57)
U	0.30 (0.03)	8.32 (1.05)	11.59 (2.99)
1987 Jul. 28			
K	32.76 (1.51)	13.23 (0.42)	22.96 (1.24)
H	23.49 (1.09)	11.93 (0.19)	19.73 (0.91)
J	14.88 (1.10)	12.12 (0.41)	17.12 (1.31)
I	5.36 (0.31)	10.26 (0.40)	16.80 (1.16)
R	2.55 (0.13)	10.56 (0.41)	14.99 (1.12)
V	1.55 (0.08)	9.77 (0.70)	9.98 (1.82)
B	0.69 (0.04)	12.27 (0.59)	9.47 (1.46)
U	0.28 (0.02)	11.42 (1.20)	5.74 (2.79)
1987 Jul. 30			
H	23.28 (1.08)	13.85 (0.32)	18.81 (1.25)
I	5.56 (0.27)	12.57 (0.62)	16.96 (1.29)
R	2.84 (0.14)	12.01 (0.55)	12.54 (1.16)
V	1.60 (0.09)	10.36 (1.10)	15.46 (2.48)
B	0.74 (0.04)	13.25 (0.60)	10.06 (1.34)
U	0.29 (0.02)	13.04 (1.54)	4.75 (1.96)
1987 Sep. 19			
H	13.15 (1.21)	8.35 (0.61)	31.50 (2.36)
I	3.00 (0.28)	8.11 (0.59)	40.95 (1.84)
R	1.70 (0.16)	8.37 (0.35)	38.41 (1.38)
V	0.95 (0.09)	8.35 (1.24)	45.04 (3.82)
B	0.38 (0.04)	14.04 (0.69)	37.46 (1.42)
U	0.18 (0.02)	10.53 (1.60)	37.37 (4.74)
1987 Sep. 20			
H		6.96 (0.29)	25.96 (1.21)
I		8.43 (0.55)	32.97 (1.87)
R		9.07 (0.39)	31.05 (1.24)
V		10.52 (0.94)	38.75 (3.07)
B		13.09 (0.87)	38.30 (2.03)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
U			19.57	(2.06)	31.48	(3.00)
1987 Sep. 21						
K			8.56	(0.51)	20.69	(1.59)
H			8.44	(0.40)	24.40	(1.26)
I			9.34	(0.50)	31.11	(1.52)
R			9.88	(0.36)	33.51	(0.97)
V			12.19	(1.44)	32.25	(3.41)
B			13.90	(0.82)	33.18	(1.71)
U			18.59	(2.31)	35.95	(3.53)
PKS 2208-137			E(B-V) = 0.00			
1986 Aug. 3						
H	1.87	(0.09)	4.11	(4.16)	152.96	(15.06)
I	0.89	(0.07)	1.40	(1.55)	51.88	(20.48)
B	1.07	(0.11)	0.00	(0.74)		
1986 Aug. 4						
H	1.82	(0.08)	0.00	(3.20)		
I	0.97	(0.05)	1.95	(1.00)	114.07	(15.56)
B	1.10	(0.09)	0.39	(0.56)	66.94	(23.78)
1986 Aug. 6						
H	1.70	(0.13)	9.32	(4.21)	157.41	(12.16)
I	1.19	(0.11)	0.00	(2.28)	109.04	(30.02)
B	1.17	(0.12)	0.99	(0.74)	77.88	(13.63)
1986 Aug. 7						
H	1.35	(0.09)	2.62	(5.23)	74.70	(31.35)
I	0.87	(0.04)	1.31	(1.67)	45.38	(20.74)
B	1.00	(0.06)	0.31	(0.75)	132.34	(21.31)
2223-052 3C 446			E(B-V) = 0.03			
1986 Aug. 4						
H	0.58	(0.08)	3.65	(20.99)	116.30	(28.41)
I	0.27	(0.02)	2.80	(7.93)	105.49	(23.25)
B	0.15	(0.01)	0.75	(5.65)	142.91	(28.84)
1987 Jul. 27						
J	1.38	(0.09)	12.99	(3.11)	76.27	(6.18)
I	0.65	(0.03)	11.64	(1.87)	70.41	(4.34)
R	0.47	(0.02)	13.21	(1.24)	76.03	(2.95)
V	0.33	(0.02)	17.42	(2.35)	75.44	(4.28)
B	0.22	(0.01)	11.59	(1.20)	68.41	(2.77)
U	0.16	(0.02)	9.75	(1.84)	71.52	(6.29)
1987 Jul. 29						
H	2.06	(0.19)	9.29	(1.48)	88.83	(5.65)
I	0.61	(0.06)	5.87	(2.05)	87.07	(9.47)
R	0.44	(0.04)	7.90	(1.35)	82.59	(4.89)
V	0.33	(0.03)	7.35	(2.75)	73.49	(7.78)
B	0.24	(0.02)	9.57	(1.15)	73.00	(4.41)
U	0.17	(0.02)	8.34	(2.17)	78.73	(6.83)
1987 Jul. 30						
H	1.98	(0.11)	9.04	(1.83)	77.89	(5.61)
I	0.64	(0.03)	7.23	(2.44)	55.10	(8.95)
R	0.45	(0.02)	7.02	(1.88)	71.52	(7.08)
V	0.31	(0.02)	7.80	(3.33)	84.66	(11.35)
B	0.22	(0.01)	9.77	(1.88)	68.42	(5.38)
U	0.15	(0.01)	6.08	(2.51)	60.98	(10.71)
1987 Sep. 20						
H			6.55	(0.91)	92.16	(3.88)
I			7.98	(1.12)	91.95	(3.90)
R			9.03	(0.59)	94.03	(1.77)
V			12.35	(2.00)	97.50	(4.37)
B			10.30	(0.86)	92.58	(2.30)
U			8.89	(1.24)	99.50	(4.12)
21st. Sep 1987						
H			9.29	(1.09)	97.70	(3.28)
I			11.90	(1.21)	98.72	(2.58)
R			9.76	(0.59)	101.84	(2.04)
V			11.64	(2.03)	106.59	(3.92)
B			10.86	(0.83)	100.67	(2.06)
U			8.92	(1.19)	103.30	(4.37)
2230+114 4C 11.69			E(B-V) = 0.03			
1986 Aug. 3						
H	1.18	(0.10)	8.29	(4.76)	154.21	(14.60)
I	0.64	(0.05)	3.77	(1.97)	52.35	(12.08)
B	0.42	(0.04)	0.00	(1.20)		
1986 Aug. 4						
H	1.02	(0.05)	7.70	(7.49)	136.84	(23.37)
I	0.64	(0.03)	3.86	(2.13)	109.45	(16.37)
B	0.39	(0.03)	0.00	(1.89)	100.39	(34.85)
2251+158			E(B-V) = 0.06			
1986 Aug. 5						
H	1.97	(0.16)	6.40	(2.91)	152.90	(14.99)
I	1.12	(0.05)	1.11	(1.31)	27.84	(19.01)
B	0.65	(0.04)	0.00	(0.68)		
1987 Jul. 28						
H	2.41	(0.16)	4.62	(2.06)	10.28	(9.32)
I	1.16	(0.06)	0.55	(1.18)	61.84	(22.04)
R	0.98	(0.05)	1.38	(0.87)	24.86	(14.17)
V	0.85	(0.05)	0.00	(1.21)		
B	0.60	(0.04)	0.00	(0.63)		
U	0.36	(0.03)	1.17	(0.97)	155.83	(24.63)
1987 Sep. 20						
H			7.35	(1.22)	161.96	(4.66)
I			4.18	(0.96)	160.62	(6.32)
R			3.92	(0.45)	162.27	(3.31)
V			3.72	(1.10)	159.21	(7.36)
B			3.49	(0.47)	161.95	(4.52)
U			2.37	(0.79)	163.02	(9.09)
2254+074 OY 091			E(B-V) = 0.06			
1986 Jul. 31						
H	2.08	(0.19)	7.06	(4.13)	46.26	(12.00)
I	0.93	(0.09)	10.09	(1.30)	57.36	(3.64)
B	0.20	(0.03)	18.91	(2.53)	57.23	(3.77)
1986 Aug. 1						
H	1.78	(0.12)	7.90	(5.62)	45.31	(12.85)
I	0.85	(0.08)	10.77	(1.48)	52.72	(4.24)
B	0.16	(0.02)	14.75	(1.57)	43.95	(3.99)
1986 Aug. 3						
H	2.50	(0.12)	10.54	(1.37)	52.66	(4.66)
J	1.75	(0.08)	12.91	(1.55)	41.11	(5.37)
I	0.84	(0.06)	11.80	(1.08)	47.70	(2.50)
R	0.52	(0.05)	12.28	(0.93)	43.97	(2.91)
V	0.23	(0.02)	15.40	(1.93)	45.59	(4.41)
B	0.18	(0.02)	18.71	(1.12)	42.44	(1.48)

TABLE III (*continued*).

Filter	Flux density (mJy)		Polarization (%)		Position angle (deg)	
1986 Aug. 4						
H	2.30	(0.06)	10.31	(1.15)	44.45	(3.18)
J	1.72	(0.08)	9.03	(1.30)	43.38	(3.54)
I	0.85	(0.04)	8.99	(0.90)	48.91	(2.77)
R	0.52	(0.03)	10.27	(0.63)	45.69	(1.84)
V	0.36	(0.03)	10.92	(0.87)	47.60	(2.92)
B	0.18	(0.02)	19.17	(0.95)	44.17	(1.40)
U	0.12	(0.01)	18.36	(4.04)	47.46	(5.98)
1986 Aug. 5						
H	2.39	(0.11)	7.69	(2.40)	55.64	(6.75)
I	0.90	(0.04)	8.44	(1.31)	40.39	(3.05)
B	0.20	(0.01)	15.80	(1.33)	48.81	(2.32)
1986 Aug. 6						
H	2.48	(0.11)	4.22	(2.19)	59.06	(14.98)
I	1.18	(0.11)	7.38	(0.98)	57.42	(3.80)
B	0.21	(0.01)	11.63	(1.21)	55.39	(3.22)
1986 Aug. 7						
H	2.01	(0.11)	7.91	(4.38)	60.36	(11.00)
I	0.44	(0.02)	10.74	(3.18)	4.26	(7.83)
B	0.11	(0.01)	2.30	(3.42)	54.50	(26.31)
1987 Jul. 28						
H	2.46	(0.11)	11.55	(1.19)	136.59	(3.17)
I	1.04	(0.05)	9.07	(0.94)	138.02	(3.07)
R	0.67	(0.03)	11.87	(0.83)	142.79	(2.18)
V	0.45	(0.02)	13.76	(1.51)	148.83	(2.17)
B	0.24	(0.01)	17.40	(1.16)	141.65	(2.19)
U	0.12	(0.01)	14.28	(1.82)	146.91	(4.89)
1987 Jul. 30						
H	2.74	(0.13)	13.62	(1.48)	139.19	(2.72)
I	1.06	(0.05)	12.06	(1.26)	143.84	(3.05)
R	0.67	(0.03)	9.89	(1.09)	142.96	(2.96)
V	0.47	(0.03)	12.09	(1.72)	141.55	(4.19)
B	0.24	(0.01)	15.76	(1.30)	148.76	(2.21)
U	0.13	(0.01)	17.52	(2.94)	139.69	(4.49)
1987 Sep. 18						
H			3.34	(3.89)	45.57	(22.10)
I			9.65	(1.45)	47.85	(4.14)
R			11.81	(0.75)	38.76	(1.82)
V			14.12	(2.43)	58.38	(5.14)
B			14.26	(1.13)	39.71	(2.57)
U			13.74	(2.16)	56.63	(4.34)
1987 Sep. 21						
H			7.01	(1.83)	50.27	(6.09)
I			8.10	(1.55)	42.91	(5.73)
R			9.76	(0.85)	42.30	(2.62)
V			14.37	(3.06)	30.83	(5.90)
B			14.63	(1.31)	38.52	(2.67)
U			13.47	(2.73)	40.60	(6.88)

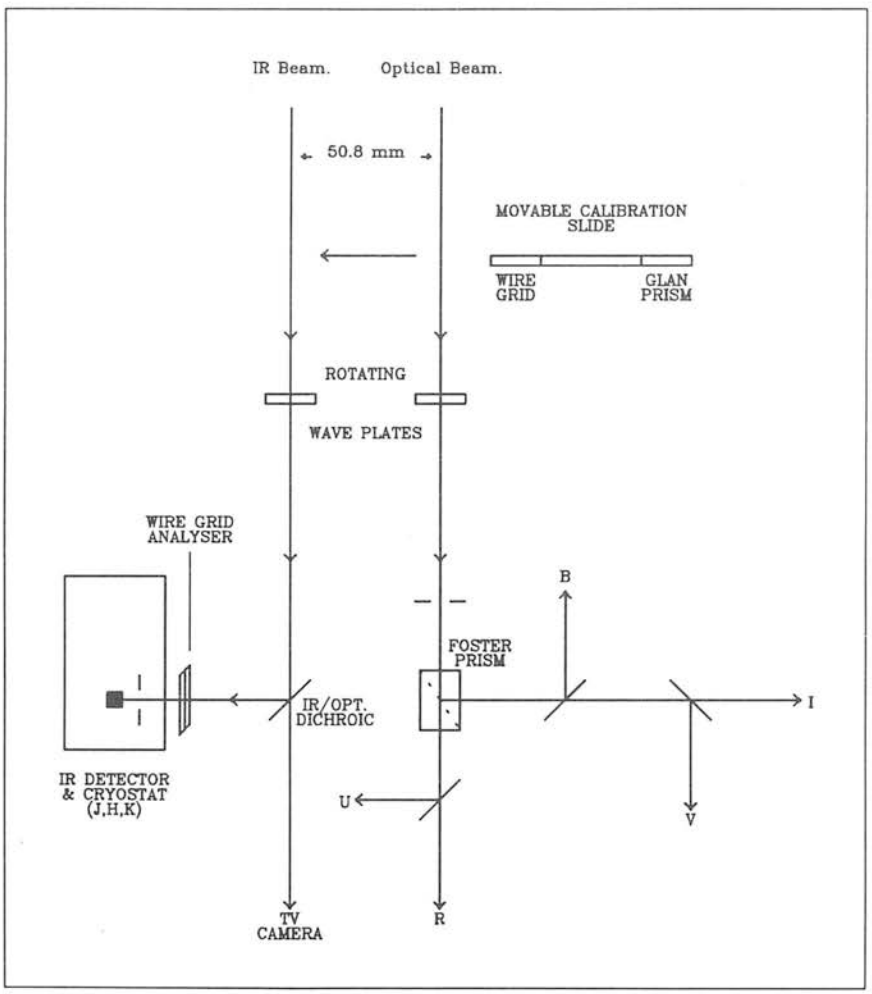


FIGURE 1.— The Mark II Hatfield Polarimeter.

Mon. Not. R. astr. Soc. (1982) 198, 1–9

A polarization burst in the BL Lac object AO 0235 + 164

C. D. Impey and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

S. Tapia *Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA*

Received 1981 February 9; in original form 1980 October 9

Summary. Simultaneous infrared and optical polarimetry and photometry have been obtained for AO 0235 + 164 covering a five night period. The object underwent a polarization burst during which the 2.2 μm polarization rose from 17.5 to 28.7 per cent and fell again to 14.9 per cent. At its peak the degree of optical polarization was 43.9 per cent, the highest linear polarization observed in a BL Lac object. The data show the degree of polarization to increase towards shorter wavelengths, and the effect is inconsistent with either dilution by a galactic component or simple one-component synchrotron models. The large changes in polarization are not accompanied by large changes in flux, a result which is difficult to explain using conventional models of these objects. Other implications of the luminosity, polarization and variability are discussed.

1 Introduction

AO 0235 + 164 is the most spectacular example of the class of extragalactic objects called BL Lac objects. BL Lac objects are characterized by a strong non-thermal continuum which rises into the infrared, relatively weak emission lines, rapid variability and a large and variable degree of polarization. In their properties they resemble the most extreme quasi-stellar objects.

The radio source AO 0235 + 164 was shown to be a radio and optical variable by Spinrad & Smith (1975). There are absorption line redshift systems at $z = 0.524$ and $z = 0.852$ (Rieke *et al.* 1976; Burbidge *et al.* 1976), and 21 cm absorption has been detected in the lower redshift system (Roberts *et al.* 1976). Although there are emission lines, they are too weak to establish a redshift. The source underwent a violent outburst in 1975, brightening by more than 5 mag with a simultaneous radio burst (Ledden, Allen & Dent 1976). It rose to the same flux level again in 1979 (Pica *et al.* 1980) with evidence for a second correlated radio/optical burst (Balonek & Dent 1980). The energetics of these variations have led to speculation on the nature of the emission in AO 0235 + 164 (Wolfe, Brown & Roberts 1976; Blandford & Rees 1978).

Large and variable linear polarization has been observed in AO 0235 + 164. During the

1975 burst, the optical polarization at peak brightness was 25 per cent with a maximum radio polarization of 4 per cent (Rieke *et al.* 1976; Ledden *et al.* 1976). Over a period of two months, the position angle of the polarization at two radio frequencies showed a remarkable linear rotation of 130°, suggesting a rotating magnetic field structure in the emitting region (Ledden & Aller 1979). The observations presented here are the first simultaneous optical/infrared polarimetry and photometry of AO 0235 + 164.

2 Observations

The observations span the five night period 1979 December 18–22, and are part of a continuing programme of infrared and optical polarimetry of BL Lac objects. Infrared polarimetry and photometry were carried out at the 3.8-m United Kingdom Infrared Telescope (UKIRT) on Mauna Kea using a photovoltaic InSb detector and rotating HR Polaroid. The optical polarimetry and photometry were carried out using Minipol (Frecker & Serkowski 1976) at the 1.54-m telescope of the University of Arizona on Mt Lemmon.

The data cover the $B(0.44\,\mu m)$, $V(0.55\,\mu m)$, $I(0.85\,\mu m)$, $J(1.25\,\mu m)$ $H(1.65\,\mu m)$ and $K(2.2\,\mu m)$ wavebands, and the spectral transmission of all these filters has been calibrated. Nearby unpolarized stars (Serkowski 1974) were observed to determine the instrumental polarization and the results have been corrected accordingly. In addition, the infrared data has been corrected for the polarizing efficiency of the HR Polaroid, and the BN object was measured at H and K as a polarized standard. Fluxes in absolute units were obtained by observing standard stars. The infrared data has been corrected for the difference between the spectral flux distribution of the BL Lac object and the calibrators. Normalized Stokes parameters and their errors were calculated from a least-squares analysis of the data, and the degree of polarization ($p = (Q^2 + U^2)^{1/2}/I$) and position angle ($\theta = {}^1/_2 \tan^{-1} Q/U$) were calculated from these quantities. The bias in p due to noise has been taken into account (Wardle & Kronberg 1974).

3 Results

The results of the infrared and visual observations are given in Table 1. Column 1 contains the date of observation and column 2 the starting and finishing times of the measurements. The remaining columns contain the data from 0.44 to 2.2 μm, consisting of the degree of polarization, position angle, flux and appropriate errors.

There are several interesting features of the data. Assuming the cosmological nature of the redshift systems, the optical/infrared luminosity of the BL Lac object in 1979 December was at least $2 \times 10^{48}\,erg\,s^{-1}$ which is one of the highest luminosities observed in an extragalactic object. Second, there is the remarkable polarization burst over the entire optical/infrared wavelength range. At 2.2 μm, the degree of polarization rose from 17.5 to 28.7 per cent and then fell by a factor of 2 over 24 hr. It is significant that the polarization rose by the same proportion at all wavelengths between December 18 and 20. At the peak of the burst, the optical polarization measured in three wavebands was 40 per cent or greater, which is the highest degree of polarization confirmed in any BL Lac object. Third, there is a wavelength dependence of polarization in the data, with polarization increasing towards shorter wavelengths. As shown in Fig. 1, the shape of the dependence did not change substantially during the burst. It is also highly significant that many of the polarization changes were accompanied by very little variation in total flux. The dramatic rise in polarized flux between December 18 and 20 was not matched by the total flux. At 0.55 μm (V), the degree of polarization rose from 20 to 44 per cent while the flux rose by only 9 per cent. On the other hand, at

Table 1.

Date 1979 December	U.T.		*B* (0.44 μm)	*V* (0.55 μm)	*I* (0.85 μm)	*J* (1.25 μm)	*H* (1.65 μm)	*K* (2.20 μm)
18	ir: 6:30–9:20	*P* (per cent)	25.4 ± 1.6	20.1 ± 1.8	25.5 ± 1.1	–	22.8 ± 1.8	17.5 ± 2.9
	opt: 7:30–8:50	*θ* (°)	150.6 ± 1.8	150.8 ± 2.6	139.0 ± 1.2	–	129.0 ± 2.3	137.0 ± 4.8
		F (mJy)	0.46 ± 0.02	0.78 ± 0.02	1.98 ± 0.04	–	6.7 ± 0.2	14.6 ± 0.3
19	ir: 6:30–8:35	*P* (per cent)	–	–	–	–	22.3 ± 1.6	18.7 ± 3.1
		θ (°)	–	–	–	–	114.0 ± 2.1	120.0 ± 4.8
		F (mJy)	–	–	–	–	8.6 ± 0.3	12.2 ± 0.3
20	ir: 7:00–9:00	*P* (per cent)	42.5 ± 0.8	43.9 ± 1.4	39.9 ± 0.5	36.2 ± 8.6	34.6 ± 2.9	28.7 ± 3.1
	opt: 6:25–7:20	*θ* (°)	132.8 ± 0.5	131.8 ± 0.9	131.0 ± 0.4	137.0 ± 6.8	128.0 ± 2.4	125.0 ± 3.1
		F (mJy)	0.43 ± 0.02	0.85 ± 0.02	2.3 ± 0.05	4.0 ± 0.2	7.6 ± 0.3	13.4 ± 0.4
21	ir: 7:00–7:20	*P* (per cent)	–	–	–	–	–	26.7 ± 3.6
		θ (°)	–	–	–	–	–	110.0 ± 3.8
		F (mJy)	–	–	–	–	–	13.3 ± 0.2
22	ir: 6:30–8:25	*P* (per cent)	–	–	–	22.1 ± 2.0	19.4 ± 1.2	14.9 ± 0.9
		θ (°)	–	–	–	119.0 ± 2.6	109.0 ± 1.8	108.0 ± 1.7
		F (mJy)	–	–	–	5.6 ± 0.3	10.2 ± 0.4	14.3 ± 0.2

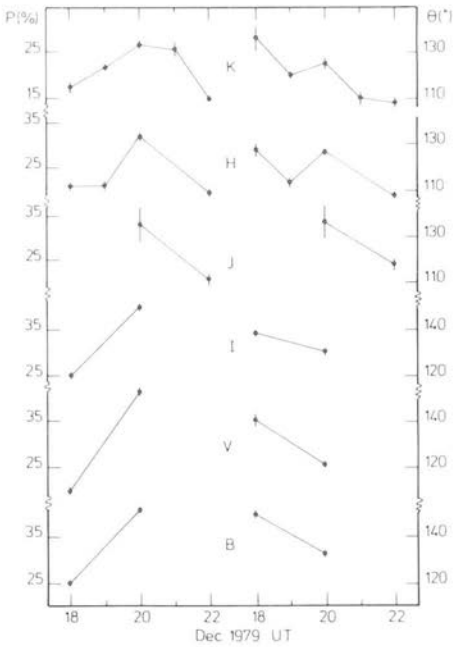


Figure 1. Degree of polarization and polarization position angle plotted for five nights of observation. Effective wavelengths for the spectral bands are given in the text.

0.44 μm (B) the degree of polarization rose from 25 to 43 per cent while the flux *fell* by 8 per cent. During the tail of the burst, the 2.2 μm (K) polarization fell from 29 to 15 per cent with a flux increase of 8 per cent. Polarization and flux appear to be uncorrelated, a result which must feature in any theoretical model of AO 0235 + 164. Over the period of the polarization burst the spectral shape of the emission changed. On December 18 the spectral flux was well fitted by a power law ($\alpha = 1.99 \pm 0.05$, where $\alpha = -d(\ln I)/d(\ln \nu)$) over the wavelength range 0.44–2.2 μm. However, by the fifth night the spectrum was curved with a flattening towards the infrared. In terms of $H-K$ colour, the change was from $H-K = 1.21 \pm 0.04$ on the first night, to $H-K = 1.03 \pm 0.04$ on the third night to $H-K = 0.75 \pm 0.04$ on the fifth night (Fig. 2). The spectral flux calibration used for all the data takes the flux

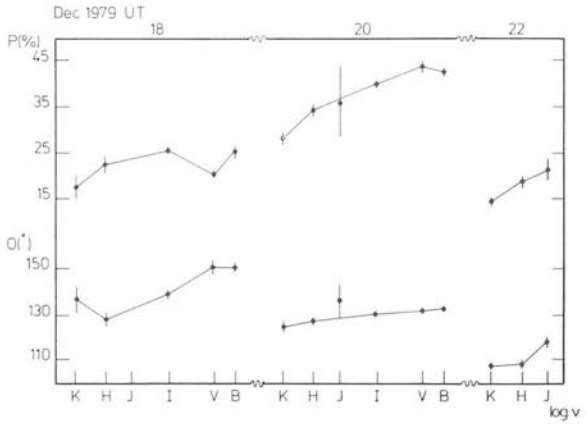


Figure 2. Degree of polarization and polarization position angle plotted against wavelength for the first, third and fifth nights of observation.

for a 0 mag star to be 1520 mJy at J (1.25 μm), 980 mJy at H (1.65 μm) and 620 mJy at K (2.2 μm). These figures are based on the calibration of α Lyr by Oke & Schild (1970) and the model atmosphere calculation by Schild, Peterson & Oke (1971). The data shows position angle rotations with both time and wavelength. On December 18 there is a 14° rotation between 0.44 and 2.2 μm, and the 2.2 μm measurements show a 30° rotation over five nights. The temporal rotation is seen at all six wavelengths between December 18 and 20.

The observations show sizeable changes in polarization and flux between consecutive nights; however, continuous observations in one colour do not indicate any significant variations on a time-scale of 1 hr. The data at all wavelengths on a given night was collected within a few hours, so the lack of complete simultaneity will not be important. Much of the published polarimetry of BL Lac objects at optical to radio wavelengths is not simultaneous and propagation effects cannot be ruled out. We note that the precision of the infrared polarimetry is better than that already published.

4 Discussion

4.1 THERMAL RADIATION

AO 0235+164 has no obvious galactic component but we consider the effect of the non-thermal source being embedded in an elliptical galaxy. This two-component model has been used to explain some of the properties of low redshift BL Lac objects, such as wavelength dependence of optical polarization (Maza, Martin & Angel 1978). It is easily demonstrated that the thermal flux from an elliptical galaxy around AO 0235+164 would be very small compared to the non-thermal component. The lower limit to the cosmological distance to the BL Lac object is given by the absorption redshift $z = 0.852$. Using $H_0 = 75\ \mathrm{km\ s^{-1}\ Mpc^{-1}}$ and $q_0 = 0$, this corresponds to a distance of 3.4 Gpc. To determine the visual luminosity, we can extrapolate the power law in Fig. 3 down to 0.33 μm which gives a rest frame luminosity per unit frequency of $L(0.55\,\mu\mathrm{m}) = 8 \times 10^{32}\ \mathrm{erg\ s^{-1}\ Hz^{-1}}$. This is one of the highest visual luminosities observed in a BL Lac object (Miller, French & Hawley 1978) and is more than 100 times the luminosity of a first-ranked giant elliptical (Frogel *et al.* 1978). In the infrared, the contamination is even less because of the steeply rising flux of the BL Lac object. Therefore contamination by a thermal component will have less than a 1 per cent effect on any of the observed parameters and cannot be responsible for the wavelength dependent polarization.

The other thermal mechanism which could generate wavelength dependent polarization in

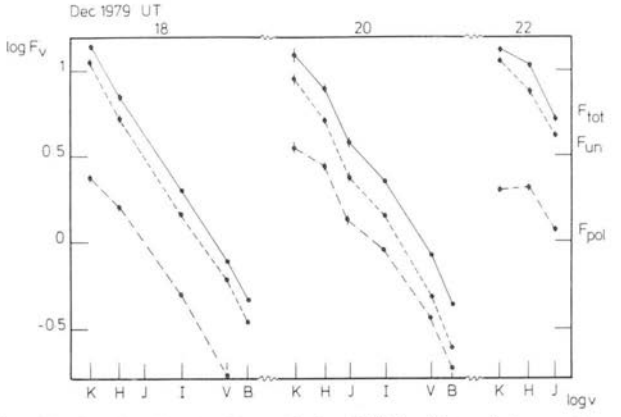


Figure 3. Spectral flux distribution for the first, third and fifth nights of observation. F_{tot}, F_{un} and F_{pol} are the total, unpolarized and polarized fluxes respectively.

the correct sense is scattering by dust or free electrons. To produce a high degree of polarization the source must be optically thin and absorbing clouds can be ruled out. The alternative is an asymmetric scattering cloud, but the rapid variability makes the model untenable. In their summary paper, Neugebauer *et al.* (1979) were able to rule out thermal radiation from dust as an important emission mechanism in high luminosity quasars.

4.2 POWER SOURCE

The size of the emitting volume in AO 0235+164 given by the variability time-scale is 2.6×10^{15} cm. The integrated optical/infrared luminosity (0.44–2.2 μm) is 2×10^{48} erg s^{-1}. This luminosity is at the top of the range for BL Lac objects and quasars, and is not far from the highest recorded luminosity for AO 0235+164 (Rieke *et al.* 1976), assuming isotropic emission. The current favourite for the central 'engine' in luminous extragalactic objects is a supermassive rotating magnetoid (black hole or spinar) fuelled by accretion (Rees 1978). An upper limit to the mass of a black hole radiating with a variability time-scale of 1 day, and the appropriate Eddington limit are: $M_H \leq 9 \times 10^9 M_\odot$, $L_{Edd} \leq 10^{48}$ erg s^{-1}. AO 0235+164 exceeded the Eddington limit by a factor of 2 during the burst, and the factor increases if emission comes from regions well away from the Schwartzschild radius. The energy density is even more severe at longer infrared wavelengths ($\lambda > 2\,\mu$m), because of the steeply rising flux distribution. However, it has been shown that certain accretion geometries can generate supercritical luminosities (Abramowicz, Calvani & Nobili 1980). Alternatively, the energy budget can be reduced if there is a bulk relativistic motion of the emitting region towards the observer.

4.3 NON-THERMAL RADIATION

The canonical model for describing the continuum emission in BL Lac objects is incoherent synchrotron radiation from a power law distribution of relativistic electrons. Significant Compton-scattered flux is ruled out by the low X-ray flux limit of 2×10^{-11} erg s^{-1} cm^{-2} from 2–11 keV (Dennison *et al.* 1978). If synchrotron losses dominate, the energy in the magnetic field must be greater than the energy in the photon field, implying that $B > 5700$ G. The high energy density means that the synchrotron cooling time is much less than the source crossing time defined by the variability time-scale. If the source is to remain optically thin, the electrons must be re-accelerated hundreds of times. Either the electrons are accelerated impulsively in a fraction of the emitting volume or continuously throughout the emitting volume.

The maximum degree of polarization from an isotropic synchrotron source depends only on the spectral index; $P_{max} = (3\alpha + 3)/(3\alpha + 5)$ for a uniform magnetic field. To produce greater than 40 per cent linear polarization, the source must not only be optically thin at optical and infrared frequences, but the magnetic field must be highly ordered also. Plasma propagation effects, such as Faraday rotation ($\propto \nu^{-2}$) can be ignored. Such a large degree of alignment in the integrated flux from an extragalactic object is remarkable. If the aligned field is localized within the source, the energy density is raised even higher.

The wavelength-dependent polarization in AO 0235+164 is not due to dilution by an underlying galaxy. $p(\lambda)$ which is intrinsic to the non-thermal source has also been observed in OJ 287 (Kikuchi *et al.* 1976), 0735+178 (Nordsieck 1976), BL Lac (Puschell & Stein 1980) and 3C 345 (Knacke, Capps & Johns 1979). In the first three cases the polarization decreased with increasing wavelength, while in the last it increased with increasing wavelength. The occurrence of $p(\lambda)$ can be transient, because BL Lac has also been observed to

have wavelength-independent polarization from 0.44 to 3.5 μm (Knacke, Capps & Johns 1976; Rudnick *et al.* 1978). Nordsieck (1976) has modelled synchrotron radiation in a partially aligned magnetic field, and shown that BL Lac objects with curved spectra are expected to show wavelength-dependent polarization. $dp/d\lambda$ at a given wavelength is a function of the local spectral index. The polarization of AO 0235 + 164 does not agree with the predictions of Nordsieck's model. On both the nights with coverage at six wavelengths, the spectral flux is well-fitted by a power law, yet there is substantial wavelength dependent polarization. On the third night of observation the polarization was ~ 15 per cent higher at 0.55 μm than at 2.2 μm. The conventional synchrotron model has difficulty explaining $p(\lambda)$, unless the radiation at different wavelengths comes from different parts of the source with varying field geometries. However, the close continuity of flux, polarization and position angle from optical to infrared wavelengths argues against such an interpretation.

No systematic rotation of position angle has been observed in AO 0235 + 164, apart from the dramatic 130° rotation at radio frequencies (Ledden & Allen 1979). The position angle of most BL Lac objects varies stochastically and widely with time (Maza *et al.* 1978), but rotations of position angle with wavelength have been observed in PKS 0735 + 178 and OI 090.4 (Rieke *et al.* 1977; Puschell & Stein 1980). In both cases the rotation was about 30° between 0.44 and 2.2 μm. Our data show rotations with both time and wavelength. Faraday rotation is ruled out because it implies $\langle n_e B \rangle \sim 10^{-5}$ G cm^{-3} along the line of sight, and since the rotation is proportional to ν^{-2}, effects at 0.44 μm should be mirrored with amplitude 25 times greater at 2.2 μm. The data shows that the rotations at 0.44 and 2.2 μm have similar amplitude. Jones & O'Dell (1977) have suggested that if rotation with wavelength is not due to thermal electrons, it may be due to the relativistic electrons themselves. However, to preserve the high degree of polarization in AO 0235 + 164, the Faraday optical depth due to any electron population must be small. Changes in opacity due to synchrotron self-absorption do not apply at these frequencies. The most straightforward mechanism for producing the rotation with time is rotation of the magnetic field within the emitting region.

The data for AO 0235 + 164 show large changes in the degree of polarization without corresponding changes in the total flux. Similar intriguing behaviour was noted in a burst of B2 1308 + 326 by Puschell *et al.* (1979) and it has also been observed in both the optical and radio emission of OJ 287 (Wardle 1978; Visvanathan 1973). Miller (1978) discussed similar results for 3C 371 and pointed out the significance of the effect in BL Lac objects. The changes in polarized and total flux are not correlated in AO 0235 + 164, and so certain models can be ruled out. Blandford & Rees (1978) suggested that a BL Lac object might be synthesized from sub-units, each with nearly uniform field and high synchrotron polarization. The number of sub-units must be very small to generate > 40 per cent polarization. However, none of the night-to-night changes can be accounted for by *any* combination of polarized sub-units. An alternative explanation which can also be excluded is a depolarizing screen of variable efficiency outside the source, for the same reason of Faraday optical depth already described. The most reasonable cause of the polarization changes is a physical reordering of the magnetic field.

AO 0235 + 164 is a BL Lac object with indirect evidence for relativistic motion. Ledden *et al.* (1976) observed a peak brightness temperature of $T_\beta > 10^{15}$ K at 8 GHz, which is far in excess of the maximum for an incoherent synchrotron emitter. The absence of interstellar scintillation at radio frequencies confirms the relativistic expansion, regardless of the emission mechanism (Scheuer 1976). Blandford & Rees (1978) have modelled the radio and optical burst of 1975 November by the emission behind a relativistic blast wave. In their model, BL Lac objects represent a preferred line of sight, down a jet where the emission is forward-focussed by the relativistic motion of the emitting region. Observational support for a form

of relativistic blast wave comes from the upper limit on the variable degree of Faraday rotation (Wardle 1977), which requires the number of ultrarelativistic electrons ($\gamma < 100$) to far outnumber lower energy ($\gamma < 10$) electrons. This picture of relativistic expansion has two important consequences on the data presented here. First, bulk motions reduce the energy requirements of the source and made super-critical luminosities unnecessary. The luminosity is reduced by a factor Γ^2, where Γ is the bulk Lorentz factor of the emitting region and the observer is in the emission core. The equipartition field is correspondingly reduced by Γ, and the synchrotron cooling time is increased by $\Gamma^{3/2}$ (Blandford & Rees 1978). Second, the dynamical effect of the strong shock can accelerate the electrons and amplify the magnetic field (Jones & Tobin 1977; Blandford & McKee 1977). In this case the polarization flare is a direct result of the interaction of the magnetized plasma in the jet with the surrounding medium (Marscher 1980).

5 Conclusion

AO 0235 + 164 is an exceptional object. It displays many of the most dramatic properties of BL Lac objects: large luminosity, high degree of polarization and rapid variability. The extreme conditions provide a good confrontation of observations with theory, and the observations presented have put important new constraints on the available models. Polarization studies are very important, because the polarization and flux are not correlated and photometry alone would not have revealed the large changes taking place in AO 0235 + 164. Simultaneous photometry and polarimetry will be important in monitoring BL Lac objects and understanding the source of their activity.

Acknowledgments

We thank the Panel for the Allocation of Telescope Time, and T. J. Lee and the UKIRT staff for excellent technical support and many helpful suggestions. Much credit is due to R. J. Beetles, W. A. Cormack and the staff of the ROE workshop for the design and construction of the Edinburgh Polarimeter. CDI acknowledges the receipt of an SRC Studentship. Optical observations with the Minipol polarimeter are supported by NSF Grant AST-79-23593.

References

Abramowicz, M. A., Calvani, M. & Nobili, L., 1980. *Astrophys. J.*, **242**, 772.
Balonek, T. J. & Dent, W. A., 1980. University of Massachusetts preprint.
Blandford, R. D. & McKee, C. F., 1977. *Mon. Not. R. astr. Soc.*, **180**, 343.
Blandford, R. D. & Rees, M. J., 1978. Pittsburgh Conference on BL Lac Objects, p. 328, ed. Wolfe, A. M., University of Pittsburgh Press.
Burbidge, E. M., Caldwell, R. D., Smith, H. E., Liebert, J. & Spinrad, H., 1976. *Astrophys. J.*, **205**, L117.
Dennison, B., Delvaille, J. P., Epstein, A. & Schnopper, H. W., 1978. *Nature*, **276**, 375.
Frecker, J. E. & Serkowski, K., 1976. *Appl. Opt.*, **15**, 605.
Frogel, J. A., Persson, S. E., Aaronson, M. & Matthews, K., 1978. *Astrophys. J.*, **220**, 75.
Jones, T. W. & O'Dell, S. L., 1977. *Astrophys. J.*, **214**, 522.
Jones, T. W. & Tobin, W., 1977. *Astrophys. J.*, **215**, 474.
Kikuchi, S., Mikami, Y., Konno, M. & Inoue, M., 1976. *Publs astr. Soc. Japan*, **28**, 117.
Knacke, R. F., Capps, R. W. & Johns, M., 1976. *Astrophys. J.*, **210**, L69.
Knacke, R. F., Capps, R. W. & Johns, M., 1979. *Nature*, **280**, 215.
Ledden, J. E., Allen, H. D. & Dent, W. A., 1976. *Nature*, **260**, 752.
Ledden, J. E. & Allen, H. D., 1979. *Astrophys. J.*, **229**, L1.
Marscher, A. P., 1980. *Astrophys. J.*, **235**, 386.
Maza, J., Martin, P. G. & Angel, J. R. P., 1978. *Astrophys. J.*, **224**, 368.

Miller, J. S., 1978. *Comm. Astrophys. Space Sci.*, **7**, 175.
Miller, J. S., French, H. B. & Hawley, S. A., 1978. *Astrophys. J.*, **219**, L85.
Neugebauer, G., Oke, J. B., Becklin, E. E. & Matthews, K., 1979. *Astrophys. J.*, **230**, 79.
Nordsieck, K. H., 1976. *Astrophys. J.*, **209**, 653.
Oke, J. B. & Schild, R. E., 1970. *Astrophys. J.*, **161**, 1015.
Pica, A. J., Pollock, J. T., Smith, A. G., Leacock, R. J., Edwards, P. L. & Scott, R. L., 1980. *Astr. J.*, **85**, 1442.
Puschell, J. J. & Stein, W. A., 1980. *Astrophys. J.*, **237**, 331.
Puschell, J. J., Stein, W. A., Jones, T. W., Warner, J. W., Owen, F., Rudnick, L., Allen, H. D. & Hodge, P., 1979. *Astrophys. J.*, **227**, L11.
Rees, M. J., 1978. *Phys. Scripta*, **17**, 201.
Rieke, G. H., Grasdalen, G. L., Kinman, T. D., Hintzen, P., Wills, B. J. & Wills, D., 1976. *Nature*, **260**, 754.
Rieke, G. H., Lebofsky, M. J., Kemp, J. C., Coyne, G. V. & Tapia, S., 1977. *Astrophys. J.*, **218**, L37.
Roberts, M. S., Brown, R. L., Brundage, W. D., Rots, A. H., Haynes, M. P. & Wolfe, A. M., 1976. *Astr. J.*, **81**, 293.
Rudnick, L., Owen, F., Jones, T. W., Puschell, J. J. & Stein, W. A., 1978. *Astrophys. J.*, **225**, L5.
Serkowski, K., 1974. *Planets, Stars & Nebulae Studied with Photopolarimetry*, p. 135, ed. Gehrels, T., University of Arizona Press.
Scheuer, P. A. G., 1976. *Mon. Not. R. astr. Soc.*, **117**, 1*P*.
Schild, R. E., Peterson, D. M. & Oke, J. B., 1971. *Astrophys. J.*, **166**, 95.
Spinrad, H. & Smith, H. E., 1975. *Astrophys. J.*, **201**, 275.
Visvanathan, N., 1973. *Astrophys. J.*, **179**, 1.
Wardle, J. F. C. & Kronberg, P. P., 1974. *Astrophys. J.*, **194**, 249.
Wardle, J. F. C., 1977. *Nature*, **269**, 563.
Wardle, J. F. C., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 117, ed. Wolfe, A. M., University of Pittsburgh Press.
Wolfe, A. M., Brown, R. L. & Roberts, M. S., 1976. *Phys. Rev. Lett.*, **37**, 179.

Mon. Not. R. astr. Soc. (1984) **211**, 497–506

7

A polarization flare in OJ 287

P. A. Holmes and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

C. D. Impey* *Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Hawaii 96822, USA*

P. M. Williams *UK Infrared Telescope, 900 Leilani Street, Hilo, Hawaii 96270, USA*

P. Smith, R. Elston, T. Balonek, M. Zeilik and J. Burns *Department of Physics and Astronomy, University of New Mexico, Albuquerque, NM 87131, USA*

P. Heckert *Department of Physics, California State College at San Bernardino, California 92407, USA*

R. Barvainis and J. Kenny *Department of Physics and Astronomy, University of Massuchusetts, Amherst, MA 01002, USA*

G. Schmidt *Steward Observatory, University of Arizona, Tucson, Arizona 85721, USA*

J. Puschell *Titan Systems, Inc., 8950 Villa La Jolla Drive, Suites 2232, La Jolla, California 92037, USA*

Accepted 1984 June 11. Received 1984 May 31; in original form 1984 March 5

Summary. Observations of the BL Lac object OJ 287 were made simultaneously at infrared *JHK* wavelengths and in optical and near infrared wavelengths at *UBVRI* in 1983 January. These observations, of total flux reaching a peak and then decaying, and of complex polarization changes, are explained in terms of a model of a jet emanating from an accretion disc.

1 Introduction

OJ 287 is one of the best studied and the most rapidly variable BL Lac objects known. The redshift of $z=0.306$ is deduced from a tentative identification of the O III (λ 500) emission line at 6538 Å (Miller, French & Hawley 1978). The lack of detection of any other features is consistent with the properties of blazars, and also with the lack of any detected nebulosity associated with OJ 287.

OJ 287 has shown large amplitude and rapid variability at all wavelengths (Visvanathan & Elliot 1973; Véron & Véron 1975; Valtaoja *et al.* 1982). The polarization has a range of 0.4–29 per

*Present address: California Institute of Technology, Division of Astronomy and Physics, Pasadena, California 91125, USA.

cent (this paper, Angel & Stockman 1980) in the optical region and 1.5–19.7 per cent (this paper, Impey *et al.* 1984) in the infrared. The polarization shows no preferred position angle. Periods of wavelength dependent polarization have been observed by Kikuchi *et al.* (1976) but there was no evidence for it in Impey *et al.* (1984).

The data presented here reveal large and rapid variations in flux and exotic polarization behaviour for which we propose a simple two component model.

2 Observations

Infrared observations were carried out on the 3.8 m UKIRT on Mauna Kea using a photovoltaic InSb detector and a rotating HR polaroid followed a Lyot depolarizer. Three-colour (*J*, *H*, *K*) photometry and polarimetry was collected on each of the 6 consecutive nights, and on 5 nights, *L* photometry was also collected. *JHK* photometry was performed twice on January 8, revealing a 7 per cent variation in flux in one hour. The f/35 chopping secondary was used with star/sky chopping frequency of 7–8 Hz. Instrumental sensitivity and polarization were calculated each night immediately prior to observing OJ 287. Broad band filters were used and their effective wavelengths were $J(1.25\,\mu m)$, $H(1.65\,\mu m)$, $K(2.2\,\mu m)$ and $L(3.4\,\mu m)$.

Calibration and reduction procedures for this project were described by Impey *et al.* (1982). Analyser efficiency and positions angle zero point were redetermined each time the polarimeter was mounted on the telescope. The polarimetry was calibrated with nearby unpolarized stars, and the instrumental polarization, of less than half a per cent, was repeatable over the run to better than one quarter per cent. Flux calibrations have taken into account the effect of different flux distributions of OJ 287 and the calibration stars in the broad pass brands. OJ 287 was acquired directly on the Scanco TV system. All observations were made with 8 arcsec aperture, and near dark of moon.

Optical observations were carried out on the 1.5-m UCSD/Minnesota telescope. Five colour (*UBVRI*) photometry and polarimetry was collected on 9 consecutive nights, from January 2nd to 10th inclusive. Only the 4 nights where simultaneous optical-infrared data were collected are discussed in this paper. The filters and effective wavelengths are $U(0.36\,\mu m)$, $B(0.44\,\mu m)$, $V(0.55\,\mu m)$, $R(0.64\,\mu m)$ and $I(0.79\,\mu m)$.

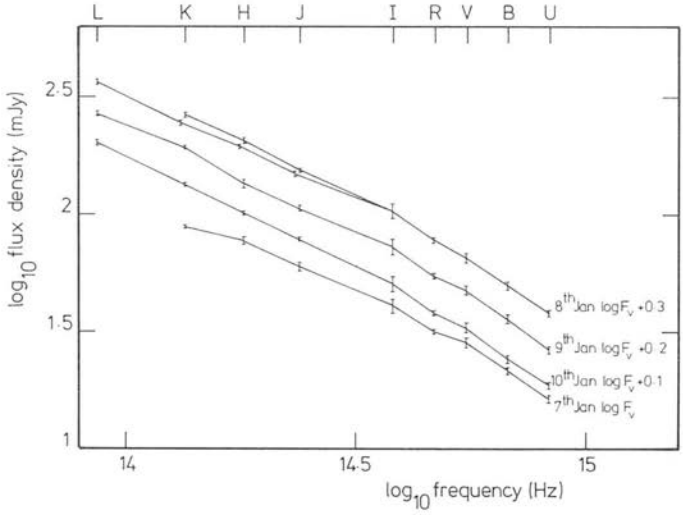


Figure 1. Flux spectrum. Log_{10} (flux density) in mJy plotted against log_{10} frequency.

3 Results

The infrared flux of most BL Lac objects is well approximated by a power law (e.g. Allen, Ward & Hyland 1982). However, our spectra of OJ 287 show significant curvature in the optical-infrared region (see Fig. 1) with local spectral index increasing with frequency. Typical values are $\tilde{\alpha}_{IR}=0.87$ and $\tilde{\alpha}_{opt}=1.25$ where $\alpha=-d \ln I/d \ln \nu$.

Similar curvature has also been seen in the optical-ultraviolet region in OJ 287 by Maraschi *et al.* (1983) and in the BL Lac object 1156+295 by Glassgold *et al.* (1983). The spectral index-redshift relation discussed by Holmes *et al.* (1984a) indicates that spectral curvature over longer frequency ranges may be general.

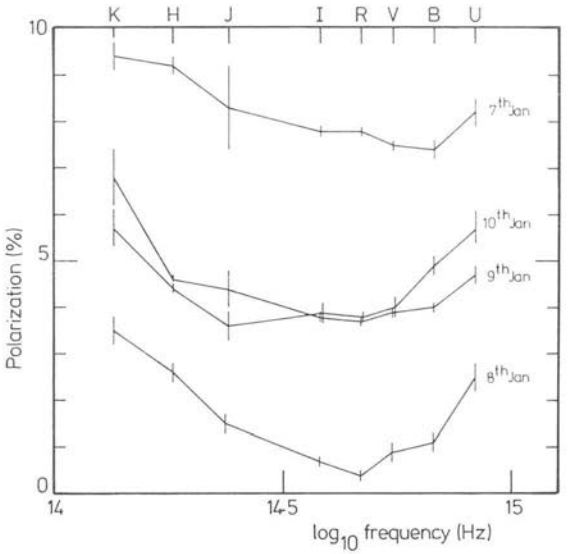


Figure 2. Spectral polarization. Polarization percentage plotted against $\log_{10}$ frequency.

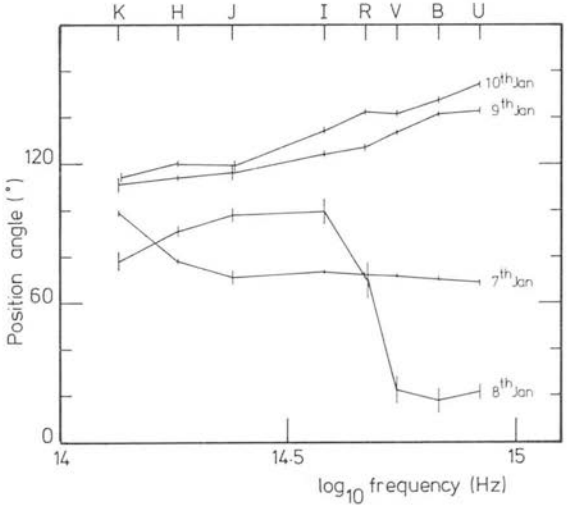


Figure 3. Spectral position angle. $\theta°$ plotted against $\log_{10}$ frequency.

The spectrum maintains its shape approximately from night to night, with the exception of the 2.2 μm flux on January 7. However, it is the unusual polarization data which give the best indication of the nature of the object. The degree of polarization, p, shows a strong wavelength dependence with p decreasing with frequency ($dp/d \log \nu < 0$) in the infrared region, reaching a minimum in the near infrared (I, R) and increasing with frequency thereafter (see Fig. 2). A wavelength dependence of the polarization position angle ($\theta°$) is also observed, with a large rotation ($\Delta\theta = 80°$), between 0.5–0.8 μm, occurring on January 8 (see Fig. 3). The flux, polarization degree and position angle data are tabulated in Table 1. Also on January 8, the polarized flux showed a pronounced dip at 0.64 μm.

We can immediately reject a single synchrotron component model, as this could not yield the observed wavelength dependence. We can also reject the model consisting of two components with spectral cut-offs. This model would produce an abrupt change in p and θ at the cut-off frequency of the first component, when the second component begins to dominate completely. Such behaviour is not observed. Also, such components would need to vary in unison in order to maintain the spectral shape and we would thus expect to see evolution of the cut-off frequency due to synchrotron energy losses on time-scales comparable with the variability time-scale ($\tau_{var} \sim 10$ hr), which is not observed. We may also reject the model consisting of two polarized non-variable components, which give rise to the spectral shape, and a third synchrotron component of spectral index ~ 1.5, which varies rapidly in flux. This describes well the internight variations in flux but can be rejected on the grounds that the variable component must have wavelength dependence of polarization which is theoretically prohibited in an optically thin incoherent synchrotron source.

Table 1. Flux and polarization measurements (with errors).

	w/b	F_ν (mJy)	$\sigma(F_\nu)$	p%	σ(p%)	$\theta°$	$\sigma(\theta°)$
7 Jan 83	U	16.3	0.5	8.2	0.3	68.8	1.0
	B	21.7	0.8	7.4	0.2	70.3	0.7
	V	28.5	1.2	7.5	0.1	71.4	0.5
	R	31.7	0.6	7.8	0.1	71.9	0.5
	I	41.0	3.0	7.8	0.1	73.4	0.5
	J	60.0	2.2	8.3	0.9	71	3
	H	77.1	2.8	9.2	0.2	78	1
	K	88.8	0.8	9.4	0.3	99	1
8 Jan 83	U	19.0	0.6	2.5	0.3	21.6	3.1
	B	25.0	0.9	1.1	0.2	17.8	5.2
	V	32.7	1.3	0.8	0.2	22.3	5.6
	R	38.9	0.8	0.4	0.1	69.5	7.6
	I	52.0	3.7	0.7	0.1	99.3	5.5
	J	77.6	1.4	1.5	0.2	98	3
		74.8	1.4				
	H	103.6	1.9	2.6	0.2	91	2
		98.0	1.8				
	K	133.2	3.6	3.5	0.3	78	4
		123.7	3.4				
	L	183.3	5.5	-	-	-	-
9 Jan 83	U	16.7	0.5	4.7	0.2	142.8	1.4
	B	22.7	0.9	4.0	0.1	141.3	1.1
	V	30.1	1.2	3.9	0.1	133.2	1.0
	R	34.5	0.7	3.7	0.1	126.8	0.9
	I	45.9	3.3	3.8	0.1	124.0	1.1
	J	67.0	1.2	4.4	0.4	116	3
	H	85.4	3.1	4.6	0.1	114	1
	K	121.4	1.2	6.8	0.6	111	3
	L	168.7	3.4	-	-	-	-
10 Jan 83	U	14.8	0.4	5.7	0.3	154.2	1.3
	B	19.3	0.7	4.9	0.2	147.4	1.0
	V	26.1	1.1	4.0	0.2	141.6	1.1
	R	30.2	0.6	3.8	0.1	142.1	1.0
	I	40.2	2.9	3.9	0.2	134.2	1.1
	J	62.2	0.6	3.6	0.3	119	2
	H	80.8	0.8	4.4	0.1	120	1
	K	106.8	1.9	5.7	0.4	114	2
	L	159.6	4.8	-	-	-	-

4 The two-component model

The model we propose seeks to explain the gross behaviour of the data in terms of two optically thin polarized components (p, θ independent of frequency). That such a superposition can produce p, θ curves similar to those observed can readily be seen. Consider two power-law components of spectral slope α_1, α_2 respectively.

If (I, p, θ), (I_1, p_1, θ_1) and (I_2, p_2, θ_2) are the intensity, polarization degree and position angle of the resultant radiation, and components 1 and 2 respectively, then we may write

$$p^2 = \frac{p_1^2 + p_2^2 (I_2/I_1)^2 + 2p_1 p_2 (I_2/I_1) \cos 2\xi}{(1 + I_2/I_1)^2} \tag{1}$$

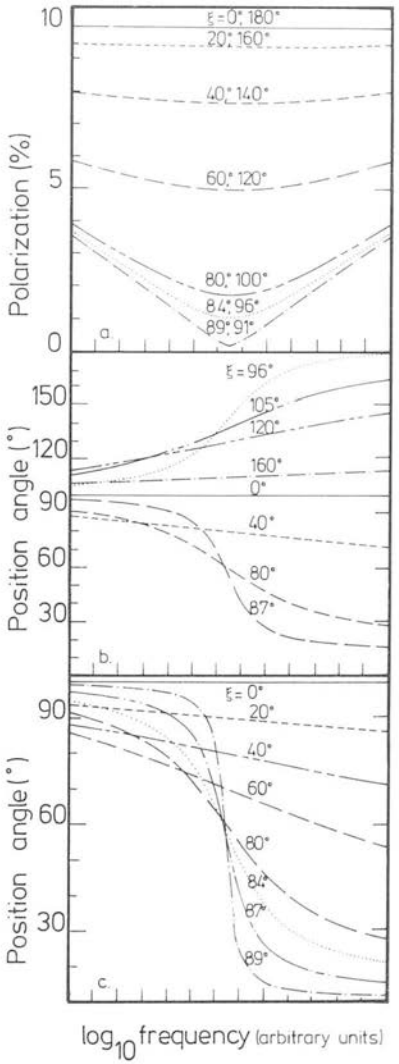


Figure 4. (a) polarizations and (b) and (c) position angles plotted against log frequency for several values of ξ. (b) shows rotations for $0 \leqslant \xi \leqslant 180$ while (c) shows, in more detail, rotations for $0 \leqslant \xi < 90$, emphasizing the rapid rotations with frequency for $\xi \leqslant 80°$. Parameter values used: $p_1 = p_2 = 10$ per cent, $\theta_1 = 100°$, $\nu_0 = 4.5$, $\alpha_2 - \alpha_1 = 1/2$.

where $\xi=\theta_1-\theta_2$ and

$$\tan 2\theta=\frac{p_1 \sin 2\theta_1+p_2\, I_2/I_1 \sin 2\theta_2}{p_1 \cos 2\theta_2+p_2\, I_2/I_1 \cos 2\theta_2}. \tag{2}$$

$I_2/I_1=(\nu/\nu_0)^{\alpha_1-\alpha_2}$ where ν_0 is the 'crossover frequency', i.e. the frequency at which the fluxes of each component are equal.

The resultant spectral polarizations and position angles are then sensitive functions of ξ. The $p(\nu)$ and $\theta(\nu)$ curves are well approximated by taking the values $p_1\simeq p_2\simeq 11$–20 per cent, $\theta_1\simeq 100°$, $(\alpha_1-\alpha_2)=1/2$ where p_1, p_2, θ_1, θ_2 are all independent of frequency, and only θ_2 and ν_0 vary with time. The minimum degree of polarization on each night occurs at the frequency for which the polarized fluxes of each component are equal.

The $p(\nu)$, $\theta(\nu)$ curves are relatively insensitive to the exact values of the parameters p_1, p_2, θ_1, θ_2, ν_0. The $p(\nu)$ curve is independent of θ_1 and θ_2, and the $\theta(\nu)$ curve is independent of the absolute values of p_1 and p_2, but is sensitive to the ratio p_1/p_2. Both curves are very sensitive to the value of ξ, especially for $\xi>80°$. This is well illustrated by the resultant curves shown in Figs 4(a, b, c).

However, such a double power-law component model is only a first approximation. Although it reproduces the p and θ curves remarkably well, the resulting spectrum must necessarily be concave ($d\alpha/d\ln\nu<0$), whereas our observations show convex spectra. The reason we would find concave (or negative) curvature can readily be seen from equation 3 and 4.

The resultant local spectral index for two superposed sources is given by (see Björnsson & Blumenthal 1982)

$$\alpha=\delta_1\alpha_1+\delta_2\alpha_2 \tag{3}$$

where $\delta_1=I_1/(I_1+I_2)$ and $\delta_2=1-\delta_1$.

To find the curvature, we differentiate (3) to give

$$\frac{d\alpha}{d\ln\nu}=\delta_1\frac{d\alpha_1}{d\ln\nu}+\delta_2\frac{d\alpha_2}{d\ln\omega}-\delta_1\delta_2(\alpha_2-\alpha_1)^2. \tag{4}$$

For two power laws we see that $d\alpha/d\ln\nu<0$. However, we need to reproduce a convex spectrum, with $\Delta\alpha\sim 0.4$ from infrared to optical wavelengths. This requires a relatively small degree of curvature in the spectra of the two sources. The reason for this curvature is not well understood. We will discuss possible causes and the incidence of spectral curvature in a future paper. Let us for the meantime assume an arbitrary form for the flux distribution by extending the power law relation to second order in $\ln\nu$. For successful modelling, we require only that the fluxes of the two components vary smoothly with respect to one another, and that, as already noted, the polarized fluxes are equal at some point within the observed spectral range.

We therefore assume the following forms:

$$\begin{aligned}\ln I_1&=\ln k_1-a\ln\nu-b\ln^2\nu\\ \ln I_2&=\ln k_2-a\ln(\nu+\nu')-b\ln^2(\nu+\nu')\end{aligned} \tag{5}$$

where (k_1-k_2) and ν' represent a relative shift in the spectra such that $\alpha_1(\nu)=\alpha_2(\nu+\nu')$. We would emphasize that these forms are arbitrary, their sole purpose being to reproduce the *observed* curvature, thus enabling us to calculate the expected polarization properties accurately. There is no loss of generality in adopting these forms in preference to any other.

We then allow the polarizations to be functions of the local spectral index, such that,

$$p_{1,2}=\pi_{1,2}(\alpha_{1,2}+1)/(\alpha_{1,2}+5/3) \tag{6}$$

where π is the 'alignment parameter' ($\pi=1$ for perfect alignment, $\pi=0$ for a tangled field) which results in

$$p^2=(p_1\,\delta_1)^2+(p_2\,\delta_2)^2-2(p_1\,\delta_1)(p_2\,\delta_2)\cos 2\xi \tag{7}$$

and

$$\tan 2\theta=\frac{(p_1\,\delta_1)\sin 2\theta_1+(p_2\,\delta_2)\sin 2\theta_2}{(p_1\,\delta_1)\cos 2\theta_1+(p_2\,\delta_2)\cos 2\theta_2}. \tag{8}$$

The spectra may be well approximated by taking $a=0.8$–0.95, $b=0.12$–0.15 and $\nu'=3$, and where ν, ν' are in units of 10^{14} Hz. The flux constants k_1, k_2 need not be specified, but are defined by the values of ν' and the cross-over frequency ν_0.

Parameter value which give a reasonable fit to the observational p, θ curves for each night, and also closely describe the spectral curvature are:

	7th	8th	9th	10th
π_1	26	17	20	22
π_2	21	18	20	22
θ_1	99	104	91	96
ξ	64	89	104	103
ν_0	3.9	4.8	3.9	3.1

These values are by no means unique, and the observational data are insensitive particularly to ν_0 over a small range. We have arbitrarily chosen $a=0.8$, $b=0.15$, $\nu_1=3$ for each night to generate the theoretical curves (Figs 5a,b). However a range of values of a and b is required to fit the spectral curvatures more accurately (as explained above). For example, the spectrum of January 10 is reproduced using $a=0.95$ and $b=0.15$, together with the polarization values tabulated above (see Fig. 6).

It is clear that the cross-over frequency does change from night to night, representing a brightening of one source relative to another. The large rotation over a small wavelength range, and the pronounced dip in the polarization curve, which occur on January 8 define the frequency at which the polarized fluxes of the two components are equal. It is possible that at the longest wavelengths, the θ curve indicates the presence of a third component. The most noticeable aspect of this model is that the polarizations of each component are very similar, and tend to vary in unison from night to night. This is true of the flux distribution also, for although the cross-over frequency varies from night, it does so only over a narrow range considering the relatively large internight variations in total flux. This strongly suggests that these two components are physically connected. The first component has a stable position angle ($\theta_1\simeq 100°$) while the second component shows a gradual rotation from east through north. This behaviour may be understood in terms of a physical rotation of the magnetic field.

5 Discussion

We have described the complex polarization behaviour of OJ 287 during the January outburst in terms of two 'simple' synchrotron emitters, i.e. regions where a (presumably partially tangled) magnetic field gives rise to a wavelength independent polarization and position angle of the emitted radiation. The emitted flux densities from the two postulated sources need not be power laws in frequency but their ratios must vary slowly and smoothly with frequency, as expressed by the index difference α_1–α_2, which we take to be 1/2. This value was chosen to represent the steepening of the flux spectrum in the presence of a constant injection of electrons and energy

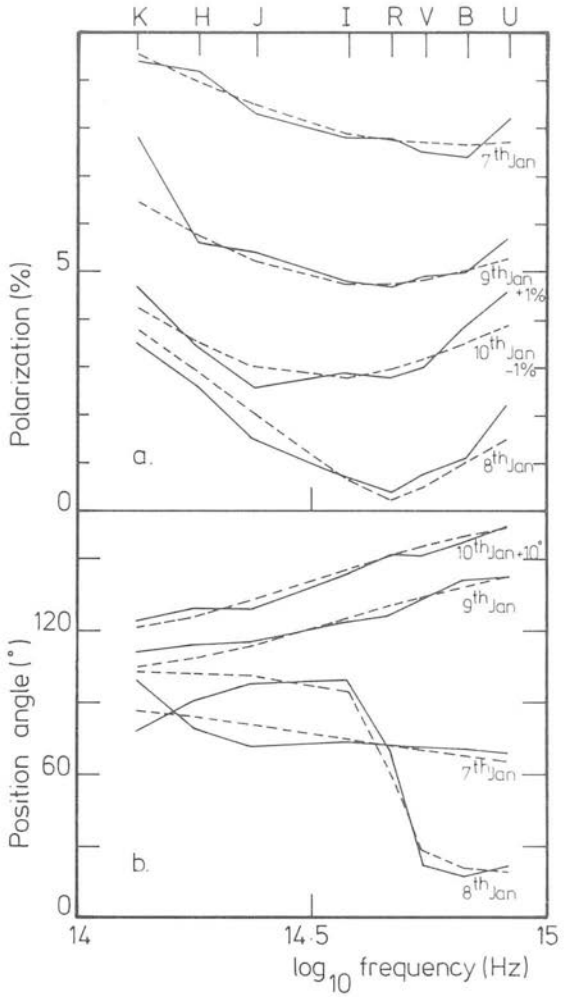


Figure 5. (a) and (b) Theoretical fits to the $p(\nu)$ and $\theta(\nu)$ curves.

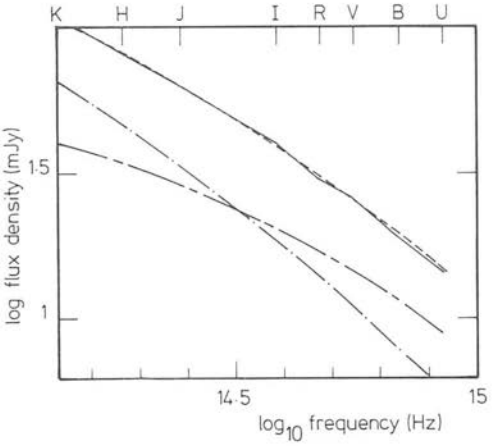


Figure 6. Log (flux density) plotted against log frequency. The solid line is the resultant of the two superposed spectra (wide dashes). The narrow dashed line is the observational data for January 10.

losses by synchrotron radiation (Kardashev 1962). Although the slight variation of ν_0 suggests that the flux densities of these two sources do not vary quite in unison, the differences are slight, certainly much less than if the sources were independent.

We stress that a *minimum* of two components are required to *dominate* the emission in the optical-infrared region only. This is further supported by infrared monitoring of BL Lacs (Holmes *et al.* 1984a,b). This by no means excludes the possibility that there are other components dominating at other wavelengths such as far-infrared and radio.

The flux variability of $\simeq 7$ per cent in one hour observed on January 8, shows that the infrared luminosity is near the Eddington limit for accretion onto a massive black hole (see Holmes *et al.* 1984a) which strongly suggest the presence of a jet (Rees 1978). The observed radiation can then be understood in terms of a cloud of emitting electrons being accelerated in the jet (e.g. Blandford & Königl 1979; Marscher 1980). We have seen, however, that the fluxes for each component vary approximately in unison while the curvature remains fairly constant. This leads inevitably to a picture of an *independent* source of electrons impinging upon the magnetic field regions.

One possibility is a magnetic field with a core-halo structure protruding from the inner edge of a geometrically thick accretion torus. The core region may have a stronger field and be better anchored to the wall of the torus, while the halo field would be weaker, and may therefore rotate under the influence of mechanical pressure from the jet.

Other possibilities exist. One explanation of the observed behaviour is that of two individual 'clouds' being accelerated in a jet of electrons separating with a Lorentz factor $\gamma \sim 2$ or 3. This would account for the apparent shift in frequency of $\nu' \simeq 3 \times 10^{14}$ Hz, which would be necessary if energy losses were not responsible for the spectral steepening of one component. The flux variation may then be due to a differing amount of intersection of the electron beam with the magnetic fields in the clouds.

Another scenario under consideration involves the structure of a spherical shock front produced when an electron jet encounters a denser medium, and the consequent magnetic field alignment due to compression. Several such possibilities are being investigated in attempt to explain the unusual observations, and these will be discussed in a future paper.

Also, we would point out that we have an explanation for the intriguing observed correlation between the wavelength dependence of polarization and its degree, in a sample of blazars (Bailey, Hough & Axon 1983; Holmes *et al.* 1984a). Since the production of strong wavelength dependence in the polarization requires a small number of separate emitting regions (i.e. two!) to dominate the luminosity of the source, and since in a multi-component model high observed polarization also requires a small number of components contributing, this could be the cause of the correlation. Large polarization wavelength-dependence is associated with high polarization. Investigation of these ideas is under way, and more data are being sought against which to test them.

It may be rarely that such observations are possible. OJ 287 was undergoing a flare of up to four times its normal near-infrared flux. It is clear, however, that with variability time-scales of less than half a day, we are looking either very close to the central power source (half a light day is the Schwarzschild radius for a $4 \times 10^9 M_\odot$ black hole, for instance) (Lynden-Bell 1978), or at a region of comparable size near a highly collimated jet whose diameter cannot be much greater than the accelerating region. This follows because the rate of energy release in blazars such as OJ 287 sets severe constraints on total energy supply, unless a massive black hole is working as a 'heat engine' at high efficiency (Lynden–Bell 1978).

Observationally, it will be important to examine OJ 287 again, to see if any vestige of the physics we have outlined above persists, or whether only in outburst do simple enough clues appear to tell us what is going on in a blazar.

Acknowledgments

We thank the staff of UKIRT for excellent support. We are very grateful to Mel Dyck for the loan of the Lyot depolarizer. PAH acknowledges receipt of a Robert Cormack Bequest Fellowship and financial assistance from SERC for part of this work.

References

Allen, D., Ward, M. J. & Hyland, A. R., 1982. *Mon. Not. R. astr. Soc.*, **199,** 969.
Angel, J. R. P. & Stockman, H. S., 1980. *Ann. Rev. Astr. Astrophys.*, **18,** 321.
Bailey, J., Hough, J. H. & Axon, D. J., 1983. *Mon. Not. R. astr. Soc.*, **203,** 339.
Björnsson, C.-I. & Blumenthal, G. R. 1982. *Astrophys. J.*, **259,** 805.
Blandford, R. D. & Königl, A., 1979. *Astrophys.*, **232,** 34.
Glassgold *et al.*, 1983. *Astrophys. J.*, **274,** 101.
Holmes, P. A., Brand, P. W. J. L., Impey, C. D. & Williams, P. M. 1984a. *Mon. Not. R. astr. Soc.*, **210,** 961.
Holmes, P. A., Brand, P. W. J. L., Impey, C. D. & Williams, P. M., 1984b. *Mon. Not. R. astr. Soc.*, submitted.
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1982. *Mon. Not. R. astr. Soc.*, **200,** 19.
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D. & Williams, P. M., 1984. *Mon. Not. R. astr. Soc.*, **209,** 245.
Kardashev, N. S. 1962. *Soviet Astronomy – AJ*, **6,** (3) 317.
Kikuchi, S., Mikami, Y., Konno, M. & Inoue, M., 1976. *Publs astr. Soc. Japan*, **28,** 117.
Lynden-Bell, D. 1978. *Physica Scripta*, **17,** 185.
Maraschi, L., Tanzi, E. G., Treves, A. & Falomo, R., 1983. *Astr. Astrophys.*, **127,** L17.
Marscher, A. P. 1980. *Astrophys. J.*, **239,** 296.
Miller, J. S., French, H. B. & Hawley, S. W., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 176, ed. Wolfe, A. M., University of Pittsburgh.
Rees, M. J. 1978. *Nature*, **275,** 516.
Valtoaja, E., Lehto, H., Teerikorpi, P., Haarala, S., Korhonen, T., Valtonen, M., Teräsranta, H., Salonen, E., Urpo, S., Tiuri, M. & Piirola, V., 1982. Turku Univ. Obs. Informo. No. 62. *Astrophys. J.*, submitted.
Véron, P. & Véron, M. P., 1975. *Astr. Astrophys.*, **39,** 281.
Visvanathan, N. & Elliot, J. L. 1973. *Astrophys. J.*, **179,** 721.

Mon. Not. R. astr. Soc. (1988) **233**, 503–509

Polarimetric observations of the quasar 3C 345

A. R. G. Mead and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

J. H. Hough *School of Natural Sciences, Hatfield Polytechnic, PO Box 109, Hatfield, Hertfordshire AL10 9AB*

J. A. Bailey *Anglo–Australian Observatory, PO Box 296, Epping, NSW 2121, Australia*

Accepted 1988 February 8. Received 1988 February 5; in original form 1987 November 23

Summary. Infrared and optical polarimetric observations of the quasar 3C 345 are presented, which show a decrease in the degree of polarization with frequency previously observed by Smith *et al.*, and interpreted by them as the result of dilution of the polarized flux by a non-polarized thermal component (the 'blue bump'). The new data are compared to previous fits of this model to 3C 345, and the validity of the model is discussed. The uncertainty in the estimates of the Balmer continuum and Fe^+ line emission is such, that attempts to determine whether the observed fall-off in polarization is the result of blackbody emission are impossible without detailed spectrophotometry.

1 Introduction

3C 345 (1641+399) is a well studied superluminal radio source and optically violently variable quasar. It has long been known to be a highly polarized quasar (Angel & Stockman 1980). Recently Smith *et al.* (1986) (hereafter referred to as SBHE) reported observations of the polarization behaviour of this object over a range of dates in 1983 and 1984. They found that the polarization always decreased with increasing frequency from the infrared into the optical region. They were able to model their observations by including an unpolarized blackbody in addition to the typical power-law continuum emission of a blazar. This blackbody dilutes the polarized flux of the power-law component and causes the decrease in polarization observed. Such blackbodies have been proposed by Malkan & Sargent (1982), who used them to fit 'bumps' observed in the optical–ultraviolet continua of quasars. They proposed that these blackbodies were the radiating accretion discs around the black holes, which are believed to be at the centre of all active galactic nuclei. Malkan (1983) improved upon this model of quasar continua by considering more realistic accretion disc spectra, which allowed for temperature gradients across the surface of the disc.

Malkan & Moore (1986) have also applied this model to observations of two blazars PKS 0736+017 and PKS 1510−089. SBHE then predicted that the blackbody component seen in the polarization data for 3C 345 should remain stable for at least a few years, since the size of the proposed accretion disc indicates the minimum variability time-scale.

In 1986 August, 3C 345 was observed as part of a continuing programme to study the optical–infrared polarization properties of blazars. The new data also show a decrease in the polarization from the infrared into the blue together with a flattening of the flux spectrum in the optical wavebands.

2 Observations

The observations were made at the 3.8-m UK Infrared Telescope (UKIRT) on Mauna Kea, Hawaii, between 1986 August 1 and August 7. The instrument used was the Hatfield Polarimeter which can make simultaneous observations using one infrared and two optical channels. The instrument is described more fully by Brindle (1986). Observations were made over a total of eight wavebands (*U, B, V, R, I, J, H, K*). The optical flux zero points were taken from Bessell

Table 1. Flux and polarization measurements for 3C 345. Errors are given in parentheses.

U.T. Date	Filter	Flux (mJy)	Polarization (%)	Position angle (°)
01.08.86	H	4.40 (0.20)	16.37 (1.24)	51.41 (1.82)
	J	2.43 (0.20)	14.52 (2.64)	43.60 (5.34)
	I	1.45 (0.10)	13.85 (0.73)	54.84 (1.60)
	V	0.91 (0.09)	11.32 (0.79)	55.96 (2.17)
	B	0.98 (0.10)	7.85 (0.30)	52.76 (1.04)
02.08.86	H	4.16 (0.23)	17.23 (0.95)	52.15 (1.35)
	J	2.56 (0.14)	17.24 (1.02)	55.10 (1.39)
	I	1.50 (0.11)	15.83 (0.82)	53.83 (1.46)
	R	1.06 (0.08)	14.96 (0.66)	51.63 (1.40)
	V	0.92 (0.07)	12.57 (0.79)	55.55 (1.72)
	B	0.90 (0.05)	8.70 (0.29)	52.48 (0.90)
04.08.86	H	4.13 (0.19)	17.87 (0.78)	55.65 (0.99)
	J	2.59 (0.12)	16.27 (0.79)	53.01 (1.35)
	I	1.43 (0.07)	15.01 (0.68)	55.18 (1.34)
	R	1.07 (0.04)	12.87 (0.51)	56.74 (1.13)
	V	0.97 (0.09)	11.83 (0.46)	53.85 (1.36)
	B	0.85 (0.07)	8.61 (0.57)	56.64 (1.67)
	U	0.60 (0.06)	7.92 (0.67)	58.14 (1.13)
05.08.86	K	8.02 (0.74)	17.93 (1.57)	63.05 (2.51)
	H	4.16 (0.19)	19.56 (1.91)	55.29 (2.64)
	I	1.39 (0.13)	16.09 (1.14)	59.64 (2.03)
	B	0.89 (0.09)	7.73 (0.68)	57.20 (2.57)
06.08.86	H	3.98 (0.18)	20.27 (1.30)	57.27 (1.96)
	J	2.52 (0.12)	20.03 (1.23)	62.31 (1.75)
	I	1.44 (0.07)	14.98 (0.72)	57.00 (1.39)
	V	0.87 (0.04)	11.14 (0.81)	60.17 (2.30)
	B	0.90 (0.04)	8.65 (0.35)	57.24 (1.22)
07.08.86	H	4.05 (0.19)	18.59 (1.55)	60.72 (2.06)
	I	1.45 (0.07)	14.39 (0.81)	58.34 (1.61)
	B	0.91 (0.05)	9.60 (0.71)	58.48 (2.48)
Combined data	K	8.02 (0.74)	17.93 (1.57)	
	H	4.13 (0.08)	17.98 (0.46)	
	J	2.54 (0.07)	17.21 (0.54)	
	I	1.44 (0.03)	14.89 (0.32)	
	R	1.07 (0.04)	13.53 (0.37)	
	V	0.89 (0.04)	11.76 (0.33)	
	B	0.90 (0.04)	8.42 (0.16)	
	U	0.60 (0.04)	7.92 (0.67)	

(1979). The infrared zero points were obtained from Campins, Rieke & Lebofsky (1985). The galactic extinction in the region of sky around 3C 345 is negligible according to Burstein & Heiles (1982). The infrared observations were made through a 7.8-arcsec aperture and the optical observations through a 8.0-arcsec aperture. The data were corrected for bias in the measurement of the degree of polarization in the presence of noise (Wardle & Kronberg 1974), though for these observations this correction was small. The position angles were calibrated by observations of standard polarized stars from Serkowski (1974). The relative position angle offsets and efficiencies for the polarimeter in the different wavebands were measured by placing a polarizing prism into the beam of the instrument. The observed measurements are listed in Table 1.

3 Results

The mean B magnitude for the period August 1 to August 7 was 16.69 ± 0.02. Babadzhanyants *et al.* (1985) list photographic B magnitudes over the period 1973–83, which peak at $B=14.70$ in 1982 September. The faintest magnitude recorded was in 1973 with $B=17.15$. 3C 345 was thus 2 mag below its peak B flux, during the 1986 August observations. There is no evidence for any variability in the fluxes over the period of our observations. There is marginal evidence that the polarization may have increased in the infrared wavebands until August 6 before dropping slightly to the August 7 value. However, the difference between the H polarizations on August 1 and August 6 is 3.9 ± 1.8 per cent, a 2.2σ result. Given this low significance, the possibility of variability over the period of observations has been ignored and the data were combined to give the mean flux and polarization values listed at the bottom of Table 1, and shown in Fig. 1. There is no evidence that the position angle depends on frequency on any night. However, the position angle increased by 6° over the period of the observations. Table 2 shows the mean position angle for each night. Position angle rotations such as this are not uncommon in observations of blazars (e.g. Brindle *et al.* 1986).

Qualitatively the frequency behaviour of the degree of polarization that we observe is similar to that seen by SBHE. There is no evidence that the degree of polarization depends on frequency for the infrared data, but it steadily decreases with increasing frequency for the optical measurements. This behaviour cannot be simply explained in terms of the synchrotron emission mechanism thought to be responsible for the polarization of the power-law continuum in blazars. A simple uniform synchrotron emitter should show no frequency dependence of polarization, as is often observed in blazars. Unlike 3C 345, however, a number of blazars do show increasing polarization towards higher frequencies (e.g. Brindle *et al.* 1986).

SBHE considered a variety of methods of reproducing their data and concluded that dilution by a non-polarized component in the optical frequencies was the most likely explanation. There are several such optical components in a blazar spectrum. These are the stellar emission of an underlying galaxy, the quasar line emission, the Balmer and Paschen continua and the thermal 'blue bump' as proposed by Malkan & Sargent (1982). SBHE constructed a model in which the contribution of the first two of these components was estimated, and then they fitted a polarized power-law continuum and a blackbody component to the remaining flux and the corrected polarization data. In order to compare the 1986 August data with those of SBHE a similar model has been constructed. The galaxy flux is estimated from the C model of an evolving elliptical galaxy developed by Bruzuala (1983). This is evaluated at $z=0.595$, the redshift of 3C 345, taking $H_0=100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $\Omega_0=0$. The flux is scaled to be in agreement with the measured R flux of the 'fuzz' surrounding 3C 345 as observed by Hutchings, Crampton & Campbell (1984). This galaxy component contributes most at K (about 10 per cent of the total flux observed) and is negligible in the B and U wavebands. The helium line emission and the Balmer and Paschen continuum fluxes are estimated using the synthetic spectrum of Grandi (1982). The *ad hoc* model

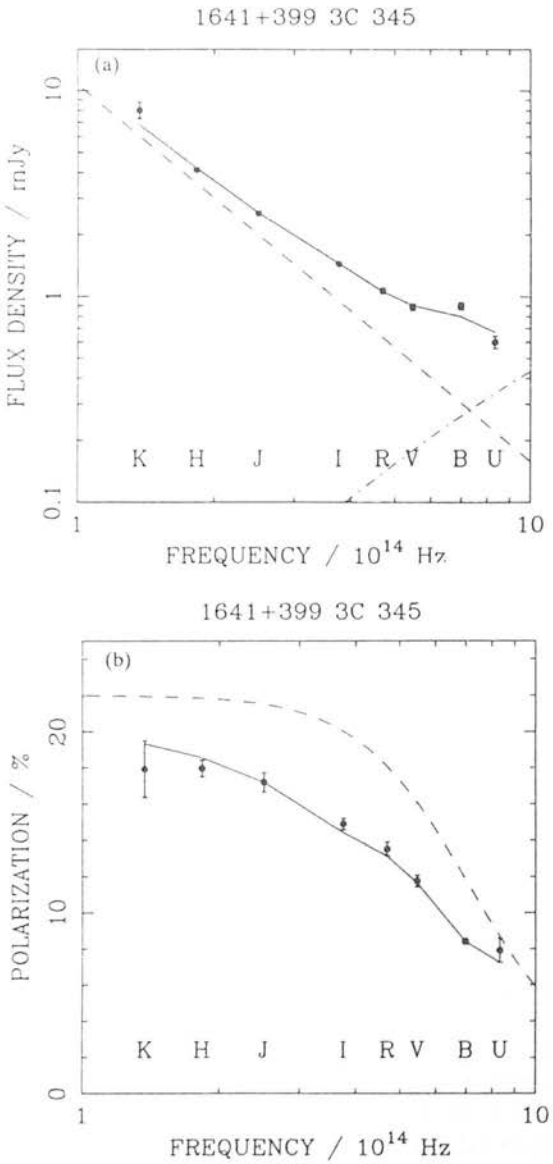


Figure 1. (a) The observed flux density is shown against frequency. The solid line connects the model estimates for the flux density at each observed frequency. The dashed lines show the flux density of the fitted power-law and blackbody components. (b) The observed polarization is shown against frequency. The solid line connects the model estimates for the polarization at each observed frequency. The dashed line shows the polarization predicted for the power-law and blackbody components combined. The difference between this curve and the observed values is due to the significant amount of line emission and starlight.

of Grandi (1981) is used to estimate the amount of Fe^+ emission. The Balmer series emission is estimated from the typical ratios quoted by Wills (1987), which agree with the observed ratios of the Hβ and Mg^+ λ2798 line fluxes given by Oke, Shields & Korycansky (1984) and Bregman *et al.* (1986). These line fluxes are used to scale the line and Balmer continuum emission, since the synthetic spectrum of Grandi (1982) gives all its line strengths and that of the Balmer continuum relative to Hα. The contributions to the non-polarized flux from these components are listed in

Table 2. Mean position angles for each night.

UT date	Position angle (°)
01.08.86	53.16 (0.73)
02.08.86	53.15 (0.53)
04.08.86	55.75 (0.47)
05.08.86	58.99 (1.20)
06.08.86	58.30 (0.71)
07.08.86	59.08 (1.13)

Table 3. This model for the contribution of line emission and starlight is essentially identical to that of SBHE. The differences occur in the estimation of the contribution of starlight (which SBHE estimate from observations of other galaxies) and in the Balmer series flux [for which SBHE use the Grandi (1982) prescription]. In order to check that these differences could not cause a systematic disagreement between the models fitted to the 1986 August data and the results of SBHE, the estimates presented here were used in fits to the data given by SBHE and gave results similar to those quoted by them, i.e. the differences are small.

These components do not appear to provide a sufficient amount of non-polarized flux to explain the polarization behaviour, and a further component of non-polarized radiation is required. Following SBHE, a power-law component with a frequency independent polarization and a non-polarized thermal component have been fitted as shown in Fig. 1. The best fitting parameters are a power-law component with a 10^{14} Hz flux density of 10.7 mJy, a spectral index ($S_\nu \propto \nu^{-\alpha}$) of 1.8 and a polarization of 22.0 per cent. The fitted blackbody has a temperature of 62 000 K [this temperature is the temperature of the emitting material, the apparent temperature being $T/(1+z)$], and an angular diameter of 0.2×10^{-6} arcsec. This angular diameter corresponds to a projected diameter of $\sim$0.003 pc or $\sim10^{14}$ m, if $H_0=100$ km s^{-1} Mpc^{-1} and $\Omega_0=0$. The χ^2 for this fit is 21.9 (10 degrees of freedom), implying a significance level of 0.016. The value of the fitted temperature is very high when compared to the values found by both Malkan & Sargent (1982) and SBHE. The peak flux for this blackbody lies at an observer's wavelength of 746 Å, well beyond the 3600 Å effective wavelength of the *U* filter. Consequently only the Rayleigh–Jeans ($B_\nu \propto \nu^2$) portion of the blackbody spectrum is being fitted. A temperature fixed at 26 000 K, was also fitted as this corresponds to the typical value found in observations of quasars and active galaxies (Edelson & Malkan 1986). For this fit the parameters of the power-law were essentially unchanged, while the fitted angular diameter was 0.4×10^{-6} arcsec, which corresponds to a projected diameter of $\sim$0.006 pc or $\sim2\times10^{14}$ m. Both of these fits give a *V* flux density of $\sim$0.2 mJy for the blackbody component. The χ^2 for the 26 000 K fit is 31.0 (9 degrees of freedom),

Table 3. Model contributions to observed fluxes in mJy.

Filter	Strong lines	Fe^+ lines	Balmer & Paschen continuum	Underlying galaxy
K	0.000	0.000	0.000	0.818
H	0.072	0.000	0.000	0.555
J	0.000	0.000	0.100	0.409
I	0.114	0.006	0.089	0.243
R	0.041	0.011	0.101	0.135
V	0.007	0.019	0.194	0.029
B	0.060	0.031	0.136	0.006
U	0.007	0.024	0.080	0.001

Bruzuala (1983) does not give a *V–I* colour for his model. The value shown here for the galaxy contribution in *I* is an interpolation between the *J* and *R* values.

which implies a significance level of 0.003. This level is totally unacceptable (assuming Gaussian statistics) and is due to the appreciable amount of curvature of this blackbody in the optical region. A Rayleigh–Jeans fit is all that is permitted by these data.

4 Discussion

SBHE found that their best fit blackbody parameters were an apparent temperature of ~16 000 K (corresponding to an emission temperature of 25 000 K) and an angular diameter of 0.6×10^{-6} arcsec. This angular diameter corresponds to a projected diameter of ~0.008 pc. The V flux density for this blackbody is ~0.4 mJy, i.e. a factor of 2 greater than that estimated for the 1986 August data. It is not unreasonable that this variability is due to a real variation of the parameters describing an accretion disc, given the expected variability time-scale of about a year.

It is important to note that variability in the line emission and Balmer continuum could also explain the difference between the 1986 August and the SBHE fits. The line fluxes used to scale the Grandi (1982) model to SBHE's data were obtained over the same period as their polarimetric observations (Oke *et al.* 1984; Bregman *et al.* 1986). These same line fluxes were used in the estimation of the line contribution to the 1986 August data. Consequently the fits to the 1986 August data will be incorrect if there has been any variability between the two epochs of observation. If the blackbody emission was constant over this period, then the variability in the line emission would have to change the V flux by ~0.2 mJy. However, no such variability has been recorded in past observations of this source.

Not just the variability, but even the existence of the blackbody component may be in doubt. The predicted contribution of line emission, Balmer continuum and starlight for the V-band is ~0.25 mJy. This is greater than the flux due to the fitted blackbody for the 1986 August data. Consequently any errors in the prediction of this contribution and those in other bands could greatly affect the fitted blackbody. The most uncertain of these contributions is the amount of Fe^+ line emission and Balmer continuum. The Grandi (1981) model is an *ad hoc* set of relative line strengths derived from fits to the spectra of quasars. Wills, Netzer & Wills (1985) present a set of fits to the optical/UV spectra of quasars using the model of Netzer & Wills (1983). This model should be an improvement on the Grandi (1981) model as it incorporates more Fe^+ multiplets. They find that all the excess flux above that which can be explained by a power-law continuum is fitted. This model is also used by Neugebauer *et al.* (1987) in analysing their 0.3–2.2 μm observations of the Palomar–Green sample of quasars. Neither of these investigations require blackbodies to explain continuum features in their spectra. These results suggest that the error in the estimation of the line emission and Balmer continuum contributions to the observed fluxes may be large enough that no blackbody contribution is required by the 1986 August data.

The fit to SBHE's data uses contemporaneous line flux data and hence the uncertainty in their fit lies mainly in the ability of the Grandi (1981) model to predict accurately the emission line contribution to the broad-band colours. The fitted blackbody is more luminous than that fitted to the 1986 August data, but the uncertainties in the model implied by the above results still apply. Since no blackbodies are found by fits using the improved model of Netzer & Wills (1985), it must be assumed that the inaccuracies in the Grandi (1981) model may be enough to explain all the flux contained in SBHE's blackbody.

These questions can only be resolved by detailed spectral fitting of the optical/UV continuum. Antonucci (1987) reports the results of spectropolarimetric observations of some low-polarization quasars and notes that the continuum polarization decreases into the blue. The excess unpolarized flux needed to explain this has a similar shape to the contribution of line emission and Balmer continuum in the fits of Wills *et al.* (1985). To find a solution to the problem of 3C 345, such spectropolarimetry is needed. The spectral resolution is required to make accurate estimates

of the contribution of the line-emitting components, while the polarimetry enables the power-law continuum component to be accurately determined.

Acknowledgments

A. Mead wishes to acknowledge the receipt of an SERC studentship.

References

Angel, J. R. P. & Stockman, H. S., 1980. *Ann. Rev. Astr. Astrophys.*, **18**, 321.
Antonucci, R. R. J., 1988. In: *Supermassive Black Holes*, Cambridge University Press, in press.
Babandzhanyants, M. K., Belokon, E. T., Denisenko, N. S. & Semenova, E. V., 1985. *Soviet Astr.*, **29**, 394.
Bessell, M. S., 1979. *Publs astr. Soc. Pacif.*, **91**, 589.
Bregman, J. N. *et al.* 1986. *Astrophys. J.*, **301**, 708.
Brindle, C., 1986. *PhD thesis*, Hatfield Polytechnic.
Brindle, C., Hough, J. H., Bailey, J. A., Axon, D. J. & Hyland, A. R., 1986. *Mon. Not. R. astr. Soc.*, **221**, 739.
Bruzuala, G., 1983. *Rev. Mex. Astr. Astrof.*, **8**, 63.
Burstein, D. & Heiles, C., 1982. *Astr. J.*, **87**, 1165.
Campins, H., Rieke, G. H. & Lebofsky, M. J., 1985. *Astr. J.*, **90**, 896.
Edelson, R. A. & Malkan, M. A., 1986. *Astrophys. J.*, **308**, 59.
Grandi, S. A., 1981. *Astrophys. J.*, **251**, 451.
Grandi, S. A., 1982. *Astrophys. J.*, **255**, 25.
Hutchings, J. B., Crampton, D. & Campbell, B., 1984. *Astrophys. J.*, **280**, 41.
Malkan, M. A., 1983. *Astrophys. J.*, **268**, 582.
Malkan, M. A. & Moore, R. L., 1986. *Astrophys. J.*, **300**, 216.
Malkan, M. A. & Sargent, W. L. W., 1982. *Astrophys. J.*, **254**, 22.
Netzer, H. & Wills, B. J., 1983. *Astrophys. J.*, **275**, 445.
Neugebauer, G., Green, R. F., Matthews, K., Schmidt, M., Soifer, B. T. & Bennett, J., 1987. *Astrophys. J. Suppl.*, **63**, 615.
Oke, J. B., Shields, G. A. & Korycansky, D. G., 1984. *Astrophys. J.*, **277**, 64.
Serkowski, K., 1974. In: *Methods of Experimental Physics*, Vol. 12, Part A, p. 361, ed. Carleton, N., Academic Press.
Smith, P. S., Balonek, T. J., Heckert, P. A. & Elston, R., 1986. *Astrophys. J.*, **305**, 484 (SBHE).
Wardle, J. F. C. & Kronberg, P. P., 1974. *Astrophys. J.*, **194**, 249.
Wills, B. J., 1988. In: *Physics of the Formation of Fe II Lines Outside LTE, IAU Coll. No. 94*, eds Viotti, R., Vittone, A. & Friedjung, M., p. 161.
Wills, B. J., Netzer, H. & Wills, D., 1985. *Astrophys. J.*, **288**, 94.

THE ASTROPHYSICAL JOURNAL, 253:19–27, 1982 February 1

SIMULTANEOUS OBSERVATIONS OF THE BL LACERTAE OBJECT I Zw 187

JOEL N. BREGMAN, A. E. GLASSGOLD, AND P. J. HUGGINS
New York University
J. T. POLLOCK, A. J. PICA, A. G. SMITH, AND J. R. WEBB
Rosemary Hill Observatory, University of Florida
WILLIAM H. -M. KU
Columbia Astrophysics Laboratory, Columbia University
RICHARD J. RUDY AND P. D. LEVAN
Center for Astrophysics and Space Sciences, University of California, San Diego
P. M. WILLIAMS
United Kingdom Infrared Telescope Unit of the Royal Observatory Edinburgh
P. W. J. L. BRAND
University of Edinburgh
G. NEUGEBAUER
Palomar Observatory
T. J. BALONEK AND W. A. DENT
University of Massachusetts
AND
H. D. ALLER, M. F. ALLER, AND P. E. HODGE
University of Michigan Radio Observatory
Received 1981 June 15; accepted 1981 August 18

ABSTRACT

Two sets of simultaneous spectra separated by 10 months were obtained for the X-ray–bright BL Lac object I Zw 187. The spectra consist of data obtained with radio, infrared, optical, ultraviolet, and X-ray telescopes. Repeated observations (nonsimultaneous) were made in several of these observing bands in order to detect flux variations. After contamination by galactic light is removed from the observations, the BL Lac component has a weak 3000 Å bump superposed upon an infrared-optical-ultraviolet spectrum of slope 0.9. The X-ray data fall on or close to an extrapolation of this power law; this result is consistent with the continuum from infrared through X-ray emission arising from a single synchrotron source. Optical and X-ray fluxes are observed to vary by no more than a factor of 3, and the shortest time scales of variability in these bands are comparable, about 1 week. No flux variations have been detected in the flat radio spectrum. Within the context of a synchrotron self-Compton model, the size of the optically thin region is $\sim 10^{16}$ cm, and the magnetic field is $\gtrsim 10^2$ *G*.

Subject headings: BL Lacertae objects — infrared: sources — radiation mechanisms — radio sources: galaxies — ultraviolet: spectra — X-rays: sources

I. INTRODUCTION

A clear understanding of the physical conditions in the continuum-emitting region is a primary goal of quasar research. Accurate spectra covering as wide a frequency range as possible are crucial to achieving this goal, but, unfortunately, they are difficult to obtain. Not only do broad-band observations of quasars, BL Lac objects, and active galactic nuclei require the use of many different telescope facilities, they exhibit irregular flux variations that require simultaneous measurements. This is the first in a series of papers which report and analyze nearly simultaneous spectra of BL Lac objects and optically violent variable quasars (OVVs).

There are two main reasons for obtaining simultaneous spectra. First, simultaneous spectra represent fundamental data which allow one to identify accurately the shape and position of spectral features that occur over several observing regions (e.g., the structure of low-frequency and high-frequency turnovers and the broad 3000 Å bump). Second, these data, when interpreted in terms of theoretical models, yield such properties as the size of the emitting region, its magnetic field, and the motion of the emitting plasma.

The BL Lac object reported on in this paper, I Zw 187 (I Zw 1727 + 502) is well suited for this analysis because it is sufficiently bright at all frequencies between radio and X-ray bands for accurate measurements to be made. Zwicky (1966) first drew attention to this compact galaxy by publishing a featureless spectrum. Photoelectric spectrophotometric observations (Oke *et al.* 1967) indicated that a nonthermal component was present which diluted the galactic spectrum. Sargent (1970) also recorded a featureless spectrum, but eventually Oke (1978) was able to detect weak galactic absorption lines from which he derived a redshift of

0.0554. Miller (1981) has also detected absorption lines at a redshift of 0.0546.

Observations obtained during the past 15 years reveal that I Zw 187 displays many properties that are typical of BL Lac objects. Temporal flux variations, possibly implying slope variations, have been observed in the nonthermal spectrum. Using photoelectric observations, no optical flux variations have been seen on the time scale of 1–3 days (Oke *et al.* 1967; Kinman 1980), but Oke *et al.* found a 0.5 mag change in the blue part of the spectrum for observations separated by 50 days while Sandage (1967) found $\Delta U \approx 0.5$ for observations separated by a month. Photographic monitoring has shown that flux variations of approximately 0.6 mag (B) occur over a 2 week period, and that the greatest range in B is 0.9 mag (Pica *et al.* 1980). In additon to flux variations, I Zw 187 displays optical polarization which varies between 4 and 6% (Kinman 1976).

Analysis of the optical data is complicated by the contamination of the spectrum by a large elliptical galaxy with about $M_v \approx -21.8$ ($H = 50$ km s^{-1} Mpc^{-1}; Oke 1978; Kinman 1978; Weistrop *et al.* 1981). When the underlying galactic component is removed from the composite spectra, the nonthermal component has a power-law slope between $\alpha = 1.6$ (Oke) and $\alpha = 0.80$–0.93 (Kinman 1978).

Although I Zw 187 has not been studied extensively at radio frequencies, it does possess a nearly flat radio spectrum with a flux density between 0.15 and 0.22 Jy in the frequency range from 1.4 to 8 GHz (Altschuler and Wardle 1975; LeSuéren, Biraud, and Lauqué 1972; Weiler and Johnson 1980); no flux variations are evident. VLBI measurements show that most of the radio emission at 5 GHz originates in a region larger than 1.2 milli-arcseconds (~ 1 pc; Weiler and Johnson).

In addition to optical and radio properties that are rather ordinary for a BL Lac object, I Zw 187 has one truly striking property—it is an extremely bright X-ray source. It was one of the few BL Lac objects detected with the *HEAO* A-1 experiment (Chubb 1978) and, of all BL Lac objects observed with the *Einstein* Observatory, I Zw 187 has the largest ratio of X-ray to optical luminosity (Ku 1982). Here we report on a series of observations which enable us to investigate the origin of the X-ray emission. Spectra that include data from the radio through the X-ray region were obtained at two different epochs. These spectra, along with monitoring studies at various frequency regions, are analyzed to calculate the properties of the emitting region.

II. THE OBSERVATIONS

I Zw 187 was observed in 1979 October and 1980 August in several different frequency regimes. The list of observers and the dates of observation are given in Table 1. The different observations are discussed in detail as a function of decreasing frequency.

TABLE 1

MULTIFREQUENCY OBSERVATIONS OF I ZW 187

Observers	Date (UT) (year/month/day)	Region	Observing Band	Raw Data	log υ	log F_υ
Ku	79/10/18.2	X-ray (IPC)	0.5–4.5 keV	0.642 ct/s	17.56	-28.96 ± 0.12
	80/8/10.2		0.5–4.5 keV	0.496 ct/s	17.56	-29.00 ± 0.12
Bregman, Glassgold, and Huggins	79/10/18.2	UV (*IUE*)	1250–1350 Å	2.59 (−15)[1]	15.36	-26.74 ± 0.04
			1350–1450 Å	1.82 (−15)	15.33	-26.84 ± 0.06
			1450–1550 Å	2.48 (−15)	15.30	-26.65 ± 0.07
			1550–1650 Å	1.93 (−15)	15.27	-26.70 ± 0.08
			1650–1750 Å	1.92 (−15)	15.25	-26.65 ± 0.05
			1750–1850 Å	2.00 (−15)	15.22	-26.59 ± 0.04
			1850–1950 Å	1.57 (−15)	15.20	-26.64 ± 0.03
	80/8/10.3	UV	1250–1350 Å	2.88 (−15)	15.36	-26.70 ± 0.04
			1350–1450 Å	2.69 (−15)	15.33	-26.67 ± 0.04
			1450–1550 Å	2.33 (−15)	15.30	-26.68 ± 0.05
			1550–1650 Å	2.49 (−15)	15.27	-26.59 ± 0.04
			1650–1750 Å	2.08 (−15)	15.25	-26.62 ± 0.03
			1750–1850 Å	2.10 (−15)	15.22	-26.57 ± 0.02
			1850–1950 Å	1.86 (−15)	15.20	-26.57 ± 0.02
			2300–2600 Å	1.38 (−15)	15.09	-26.48 ± 0.05
			2600–2900 Å	1.42 (−15)	15.04	-26.38 ± 0.05
			2900–3200 Å	1.40 (−15)	14.99	-26.31 ± 0.05
Pollock, Pica, Smith, and Webb	79/10/13	Optical	*B*	16.74 ± 0.11	14.83	-26.17 ± 0.05
	80/8/13		*U*	16.36 ± 0.09	14.92	-26.27 ± 0.02
			U	16.29 ± 0.04	14.92	
			B	16.74 ± 0.13	14.83	-26.23 ± 0.05
			B	16.94 ± 0.12	14.83	
			V	16.24 ± 0.10	14.74	-26.20 ± 0.05
Rudy and LeVan	79/10/18.2	IR	*J*	13.99 ± 0.25	14.38	-25.87 ± 0.10
			H	13.21 ± 0.10	14.26	-25.69 ± 0.05
			K	12.72 ± 0.10	14.12	-25.56 ± 0.05

TABLE 1.—(*Continued*)

Observers	Date (UT) (year/month/day)	Region	Observing band	Raw Data	log υ	log F_υ
Neugebauer	80/8/10	IR	r	16.23 ± 0.03	14.66	-26.15 ± 0.05
			J	14.24 ± 0.05	14.38	-25.81 ± 0.03
			H	13.53 ± 0.05	14.26	-25.71 ± 0.03
			K	13.03 ± 0.05	14.12	-25.62 ± 0.03
Williams and Brand	80/8/12	IR	J	14.11 ± 0.05	14.38	-25.89 ± 0.03
			H	13.32 ± 0.05	14.26	-25.71 ± 0.03
			K	12.86 ± 0.05	14.12	-25.62 ± 0.03
	80/8/19		K	12.86 ± 0.05	14.12	-25.62 ± 0.03
Balonek and Dent	79/10/26	Radio	89.6 GHz	0.14 ± 0.08[2]	10.95	$-23.85 \pm \frac{0.20}{0.37}$
	79/10/23		31.4 GHz	0.19 ± 0.05	10.50	-23.72 ± 0.07
	79/11/10		15.5 GHz	0.30 ± 0.07	10.19	-23.52 ± 0.10
	80/3/27		15.5 GHz	0.15 ± 0.09	10.19	$-23.82 \pm \frac{0.20}{0.40}$
	80/4/20		15.5 GHz	0.29 ± 0.08	10.19	$-23.54 \pm \frac{0.11}{0.14}$
	80/8/6		15.5 GHz	0.24 ± 0.07	10.19	$-23.62 \pm \frac{0.11}{0.15}$
	79/11/20		7.9 GHz	0.11 ± 0.06	9.90	$-23.96 \pm \frac{0.19}{0.34}$
	80/7/21		7.9 GHz	0.18 ± 0.11	9.90	$-23.74 \pm \frac{0.21}{0.41}$
Aller, Aller, and Hodge	79/9/4		14.5 GHz	0.36 ± 0.13	10.16	$-23.80 \pm \frac{0.14}{0.20}$[3]
	79/10/8		14.5 GHz	0.08 ± 0.11	10.16	
	79/10/22		14.5 GHz	0.14 ± 0.11	10.16	
	79/8/11		8.0 GHz	0.36 ± 0.08	9.90	-23.59 ± 0.03[4]
	79/10/5		8.0 GHz	0.27 ± 0.03	9.90	
	79/10/18		8.0 GHz	0.22 ± 0.02	9.90	
	79/11/1		8.0 GHz	0.25 ± 0.05	9.90	
	80/8/7		14.5 GHz	0.28 ± 0.14	10.16	-23.61 ± 0.03[5]
	80/8/13		14.5 GHz	0.20 ± 0.06	10.16	
	80/8/23		14.5 GHz	0.25 ± 0.04	10.16	
	80/8/25		14.5 GHz	0.29 ± 0.05	10.16	
	80/8/28		14.5 GHz	0.19 ± 0.08	10.16	
	80/8/2		8.0 GHz	0.20 ± 0.03	9.90	-23.62 ± 0.02[6]
	80/8/3		8.0 GHz	0.21 ± 0.14	9.90	
	80/8/10		8.0 GHz	0.22 ± 0.04	9.90	
	80/8/11		8.0 GHz	0.28 ± 0.05	9.90	
	80/8/15		8.0 GHz	0.24 ± 0.02	9.90	
	80/8/30		8.0 GHz	0.24 ± 0.02	9.90	
	80/9/1		8.0 GHz	0.28 ± 0.03	9.90	
	80/8/22		4.8 GHz	0.21 ± 0.03	9.68	-23.68 ± 0.06

[1] The ultraviolet raw data is given in units of ergs cm^{-2} s^{-1} Å.
[2] All raw radio fluxes are given in units of Jy.
[3] The measurement taken by Balonek and Dent on 1979 November 10 at 15.5 GHz was included when computing this average flux value.
[4] The measurement taken by Balonek and Dent on 1979 November 20 at 7.9 GHz was included when computing this average flux value.
[5] The measurement taken by Balonek and Dent on 1980 August 6 at 15.5 GHz was included when computing this average flux value.
[6] The measurement taken by Balonek and Dent on 1970 July 21 at 7.9 GHz was included when computing this average flux value.

a) X-Ray Measurements

X-ray data were obtained by the Columbia Astrophysics Laboratory with the Imaging Proportional Counter on the *Einstein* Observatory during the dates chosen for simultaneous observations. I Zw 187 was also the subject of a monitoring program designed to detect temporal flux variation.

There is an apparent decrease in flux between 1979 October 18.2 and 1980 August 10.2. However, the 1979 October measurement was made before the gain on the IPC was stabilized, so systematic errors may be the cause of the difference in counting rates. The monochromatic fluxes listed in Table 1 (which were derived from a spectral fitting procedure) are within 10% of the values obtained by imposing upon the data spectral slopes of 1.0 and 1.5 and a column density of neutral hydrogen of 3.25×10^{20} cm^{-2} (estimated from Tolbert 1972). The integrated fluxes (0.5–4.5 keV) are 1.11×10^{-11} ergs cm^{-2} s^{-1} (1979 October) and 9.8×10^{-12} ergs cm^{-2} s^{-1} (1980 August; measurements have a 30% uncertainty). The slopes are 1.0 (1979 October) and 2.4 (1980 August), but the uncertainty in the slopes is 0.5, so this difference in slopes may not be significant.

Between 1980 April 1 and April 14, 18 exposures were made in order to identify the shortest time scale of

variability. No significant changes over a time scale of hours were observed. The day-to-day fluctuations, if not due to instrumental effects, are consistent with a variability time scale similar to that found in the optical data (5 days). The greatest range in the X-ray count rate was slightly more than a factor of 2 between 1980 March and August (1.14–0.49 counts s^{-1}); the gain on the IPC was constant during this time.

b) Ultraviolet Observations

Ultraviolet spectra were taken with the *International Ultraviolet Explorer* (*IUE*) during the US1 shifts on the dates listed in Table 1. Although the exposures were approximately 5 hours for the short-wavelength camera and 2 hours for the long-wavelength camera, the spectra are weak. In reducing the data, saturated pixels, hits, and reseaux were removed from the line-by-line spectra. Five lines centered on the signal maximum define the signal and the adjacent five lines on either side define the background. After the background was filtered (11 point median filter) and smoothed (5 point running average), it was subtracted from the signal (the signal was not filtered or smoothed). Because of the weakness of the signal, the spectra are presented in 100 Å averages in the 1250–1950 Å region (short-wavelength camera) and in 300 Å averages in the 2300–3200 Å region (long-wavelength camera). The 1 σ errors in these averages are calculated in the usual way, assuming a normal distribution of data points about the mean. These errors may be larger than the true errors. An analysis of several *IUE* low-dispersion spectra indicates that, in addition to random errors due to photon statistics, there are systematic instrumental errors that do not form a normal distribution (to be discussed in detail in a future work). Because of the peculiar distribution of these systematic errors, the regular method of calculating the standard deviation of the mean yields a value that is too large. We estimate that the real uncertainty in the ultraviolet flux measurement listed in Table 1 is about 0.025 in log F_ν.

The reddening that was calculated, $E(B-V) = 0.025$ (after Burstein and Heiles 1978), is quite close to the reddening of several nearby objects listed by these authors. The extinction correction was made using the method described by Seaton (1980), and these data are presented in Table 1.

A power-law spectrum provides an extremely good fit to the data. The resulting spectral slopes are $\alpha = 1.00 \pm 0.07$ (1980 August) and $\alpha = 1.05 \pm 0.37$ (1979 October; due to the limited frequency range and poor quality of the spectra, the error in α is quite large). Determining the spectral slope from ultraviolet data has the important advantage that contamination by the galactic spectrum is negligible.

The difference in flux and spectral slope between the two short-wavelength spectra is not significant. The photometric uncertainty of well-exposed spectra may be as much as 10–15% in 100 Å bins, and peculiarities in the intensity transfer function at low-exposure levels may create additional errors (Holm 1980).

No emission lines are evident in these spectra.

c) Optical and Infrared Observations

An underlying galaxy appears on deep plates of I Zw 187 (Zwicky 1966). This galaxy contributes to the observed flux at optical and infrared frequencies, and its influence must be removed from the data. Several workers have estimated the brightness of the underlying galaxy. Oke (1978) determined that the galaxy, within an 8″ circular diaphragm, has a flux $F_\nu = 6.8 \times 10^{-27}$ ergs s^{-1} cm^{-2} Hz ($\lambda = 4326$ Å), with a 30% uncertainty. The spectral energy distributions for giant elliptical galaxies (Schild and Oke 1971; Whitford 1971) were used along with Oke's data to calculate $V(8'') = 16.97$ and $B(8'') = 18.06$, with uncertainties of 0.2–0.3 mag in each value. Weistrop *et al.* (1981) analyzed the brightness distribution of I Zw 187 and found that $V(8'') = 16.74 \pm 0.11$ and $B - V = 0.81 \pm 0.16$. Here, we adopt a value of $V(8'') = 16.85$, which is intermediate between the two determinations. Dr. T. Kinman has kindly given us his photoelectric *UBV* data of I Zw 187 (Kinman 1978, 1980) for which he has accurately measured I Zw 187 as a function of diaphragm size from 4″ to 30″ diameters. After applying the appropriate brightness corrections for seeing effects, we have determined V as a function of radius consistent with $V(8'') = 16.85$. The B magnitude at 8″ is determined less accurately than $V(8'')$ by Oke (1978) and Weistrop *et al.* because the BL Lac component contributes relatively more flux there. Therefore, we determined B and U for the galaxy by demanding that $B - V$ and $U - B$ be roughly constant as a function of radius; the color gradient effect (Strom *et al.* 1976) should be unimportant here. The results of our analysis are presented in Table 2 and are consistent with Kinman's (1978) original values (determined by a different analysis), as well as with an estimate of $V(8'')$ by Miller (1981). This radial brightness distribution is indistinguishable from those of some elliptical galaxies of comparable brightness in the rich clusters A426 (Perseus) and A2199 (K. M. and S. E. Strom 1978; S. E. and K. M. Strom 1978); it grows less rapidly than the standard growth curve for a giant elliptical (Sandage 1972).

The *UBV* data were obtained by measuring the brightness of photographic images with a 6″ fixed aperture rather than a variable aperture in the iris photometer (as was used in Pica *et al.* 1980). The use of a fixed aperture

TABLE 2

MAGNITUDES AND COLORS OF GALAXY COMPONENT OF I ZW 187

Aperture Size (″)	V	B	U	$B-V$	$U-V$	$U-B$
4	17.54	18.98[a]	...	1.44	...	...
6	16.95	18.02	18.17[a]	1.07	1.22[a]	0.15[a]
8	16.85	17.88	18.28	1.03	1.43	0.40
11	16.62	17.63	18.07	1.01	1.45	0.44
14	16.54	17.61	18.02	1.07	1.48	0.41
20	16.39	17.47	17.97	1.08	1.58	0.50
30	16.26	17.31	17.72	1.05	1.46	0.41
Average	...	...	...	1.05	1.48	0.43

[a] Error in these measurements is 0.2 mag or greater.

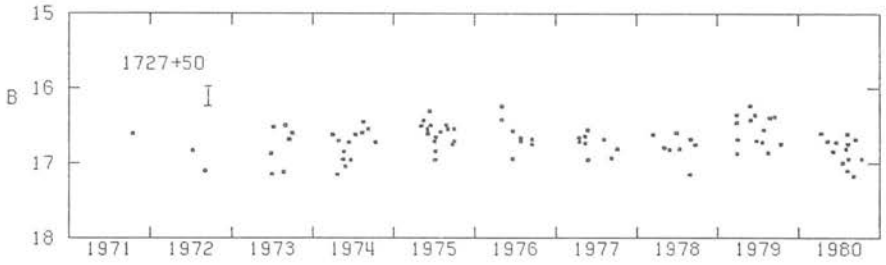


FIG. 1.—Optical monitoring of I Zw 187 (B mag) shows that its flux varies rapidly but is confined to a narrow range of about 0.85 mag. These photographic measurements were reduced with a fixed 6″ aperture iris photometer.

assures that the galaxy's contribution is constant from plate to plate; the 1979 October and 1980 August data, as well as data from other dates, were reduced by this alternative method and are presented in Figure 1. The data in Figure 1 are generally fainter by 0.1–0.2 mag than those presented in Pica *et al.* Substantial flux variations (0.6 mag) are seen in observations separated by 8–22 days. The most rapid variation, $\Delta B = 0.95$ in 8 days (5 σ effect, galaxy contribution removed), yields a variability time scale (Fdt/dF) of about 6 days, similar to the X-ray result. For the 1979 October and 1980 August measurements, the galaxy's contribution was removed and the data were dereddened (final values in Table 1 and Figures 2 and 3).

For the measurement taken at 0.655 μm ($r = 16.22$ corresponds to $F_\nu = 1.41 \pm 0.05 \times 10^{-26}$ ergs cm^{-2} s^{-1} Hz^{-1}; 5″ diaphragm), the galaxy contributes about 50% of the light (calculated from Whitford 1971, and Schild and Oke 1971). The deconvolved and dereddened flux of the BL Lac component is given in Table 1 and Figure 3.

Infrared data were obtained by three observers using photometers of different entrance apertures. In order to remove the galactic contribution properly, $V - J$, $V - H$, and $V - K$ colors for the galaxy are required. Because we have two sets of infrared data obtained through different apertures during one epoch, $V - K$ can be calculated directly, obviating the need to estimate it from typical galactic parameters. To calculate K, H, and $V - K$, we assume that $V - K$ for the galaxy is independent of diaphragm size and that $H - K = 0.21$ is a fair representation of this infrared color (Frogel *et al.* 1978); a mean value of K (or H) is then calculated. To find J, which is dominated by the galaxy more than H or K, we assume that the infrared color $J - H = 0.69$ is appropriate (Frogel *et al.*). The results of this method follow: galactic

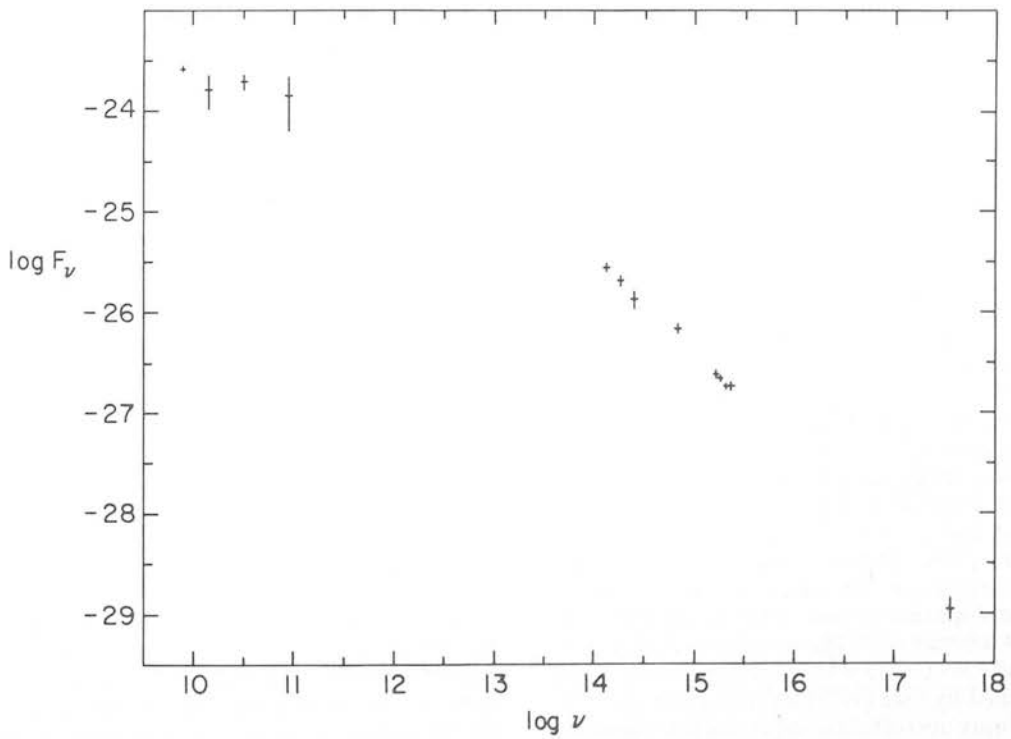


FIG. 2.—This nearly simultaneous spectrum of I Zw 187 was taken in 1979 October. The infrared-ultraviolet continuum connects smoothly with the X-ray measurement.

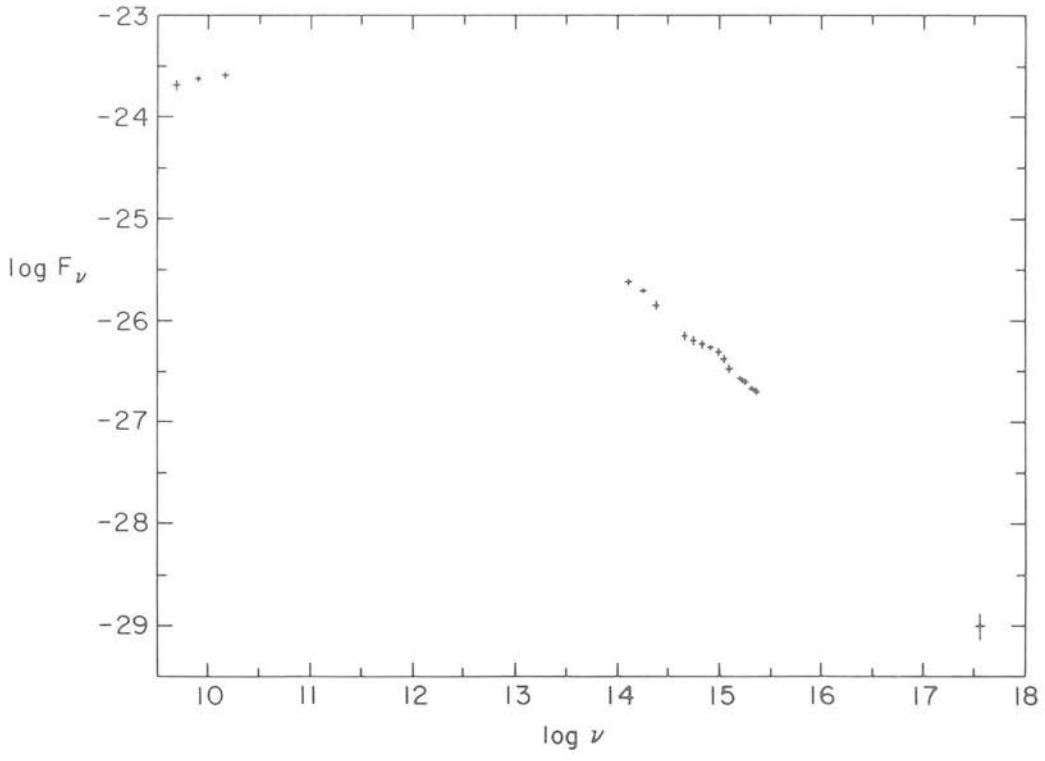


FIG. 3.—This nearly simultaneous spectrum of I Zw 187 was taken in 1980 August. Note the weak 3000 Å bump (log $\upsilon \approx 15$)

magnitudes within the 17″ diaphragm used by Rudy and LeVan are $J = 14.38$, $H = 13.69$, and $K = 13.48$; galactic magnitudes within a 10″ diaphragm used by Williams and Brand are $J = 14.60$, $H = 13.91$, and $K = 13.70$, and galactic magnitudes within the 5″ diaphragm used by Neugebauer are $J = 15.00$, $H = 14.31$, and $K = 14.10$. The galactic contribution to the total flux is between 50–70% in the 17″ diaphragm, 46–64% in the 10″ diaphragm, and 37–50% in the 5″ diaphragm. Of these three observing bands, the smallest galactic contribution occurs at 2.2 μm (K). The deconvolved, dereddened infrared data are presented in Table 1.

The infrared-optical spectra for the two epochs are, to within the errors, identical. The power-law fits to the dereddened, deconvolved data are characterized by slopes $\alpha = 0.85 \pm 0.10$ (1979 October) and $\alpha = 0.87 \pm 0.04$ (1980 August). These slopes are not only indistinguishable from each other, but are also indistinguishable from the ultraviolet power-law slopes. When a single ultraviolet-optical-infrared slope for each epoch is calculated, we find that $\alpha = 0.97 \pm 0.04$ (1979 October) and $\alpha = 0.88 \pm 0.02$ (1980 August). The formal difference between these values, a 2σ effect, is probably insignificant. These slopes are similar to the values 0.93–0.80 determined by Kinman (1978) for data taken during 1977 June 11–13 and 1977 October 15 but differ from the value of 1.6 determined by Oke (1978) for data taken on 1976 May 22–23. We are not certain whether this last difference is real or whether it reflects the uncertainty inherent in Oke's deconvolution of the BL Lac galaxy spectrum.

d) Radio Observations

Radio measurements in the GHz region were obtained with the 26 m paraboloid of the University of Michigan Radio Astronomy Observatory at 8 GHz and 14.5 GHz during both epochs; data reduction procedures are described in Aller (1970) and Aller, Aller, and Hodge (1981). The second epoch measurements, which are considerably more accurate than the first epoch, indicate that the spectrum is relatively flat. No flux difference is found between the two epochs.

Millimeter wave measurements were obtained at 10 mm and 3 mm during 1979 October with the National Radio Astronomy Observatory 11 m telescope at Kitt Peak National Observatory. These data indicate that the radio spectrum remains flat into the millimeter region, a property typical of BL Lac objects. Additional 7.9 GHz and 15.5 GHz data, obtained at Haystack Observatory, are consistent with the 8.0 GHz and 14.5 GHz data taken at the University of Michigan Radio Astronomy Observatory.

III. DISCUSSION

Sufficient data now exist to permit us to examine the nature of flux variations in several frequency regions. Unlike most BL Lac objects, which show optical variations of 2–5 visual magnitudes, I Zw 187 has not varied by more than $\Delta B = 0.85 \pm 0.11$ during the past 7 years (Fig. 1; $\Delta B = 1.25 \pm 0.13$ with the galaxy removed). Similarly, no dramatic flux variations are found in the

radio, infrared, or X-ray regions. X-ray measurements indicate flux variations of a factor of 2 (2.3 ± 0.3), which is similar to the range of optical variations (1.25 mag corresponds to a flux variation of 3.1). Consistent with this trend, no dramatic flux variations appear in the infrared data taken at Palomar Observatory between 1967 and 1980. Radio observations taken between 1972 and 1980 show no evidence at all for variation (1.4–8 GHz; LeSquéren, Biraud, and Lauqué 1972; Altschuler and Wardle 1975; Weiler and Johnson 1980; this work). To summarize, I Zw 187 has not yet displayed radio variability, and it appears to be less variable than most BL Lac objects in the infrared, optical, and X-ray regions.

Simultaneous observations are especially useful in revealing broad-band spectral features. We find evidence for a 3000 Å bump in the spectrum of 1980 August. If real, this is the first detection of the 3000 Å bump in a BL Lac object. It would suggest that this feature, which is frequently observed in QSO spectra (e.g., Neugebauer *et al.* 1979) is not associated with the emission line gas. Additional observations that allow a more thorough investigation of this feature in I Zw 187 will be presented in a future work (Bregman *et al.* 1982). A second peculiarity in the spectra, and one which is well substantiated, is the shallowness of the spectra. The infrared through ultraviolet slope of 0.88 (1980 August) is smaller than the average value for BL Lacs (Stein, O'Dell, and Strittmatter 1976), and the optical to X-ray slope of 1.0 is the smallest of any BL Lac object observed with the *Einstein* Observatory (the mean optical to X-ray slope for BL Lacs is 1.31 ± 0.05; Ku 1982).

The near equality of the infrared-ultraviolet slope and the optical–X-ray slope may have a simple explanation—we suggest that the infrared, optical, ultraviolet, and X-ray emission have a common origin in this object. When the power-law spectra defined by the infrared through ultraviolet data are extrapolated to higher energies, they pass slightly above the X-ray data point (1980 August); the difference between the extrapolated line and the X-ray data (0.38 dex) is a 3 σ effect. In the 1979 October spectrum, the extrapolated power law for infrared through ultraviolet data passes through the X-ray data point (less than a 1 σ difference). It is therefore reasonable to suggest that the X-rays are just a higher frequency part of the same power-law spectrum responsible for the infrared, optical, and ultraviolet emission. Because the optical polarization argues for a synchrotron origin for this continuous emission, the soft X-rays have a synchrotron rather than an inverse Compton origin. Using nonsimultaneous observations, Weistrop *et al.* (1981) arrive at a similar conclusion.

Our observations of the continuous spectrum provide strong evidence that the lack of strong emission lines is not caused by the lack of ionizing photons. According to current models of the emission line region (Kwan and Krolick 1981; Weisheit, Shields, and Tarter 1981), X-rays are an important source of ionization. The observed soft X-ray flux alone provides enough ionizing radiation to produce, at least, high-excitation emission lines. If the ultraviolet spectrum retains its power-law form at energies above a rydberg, sufficient ionizing radiation exists to produce emission lines at all levels of excitation. Therefore, the failure to detect lines implies that there is an insufficient amount of emitting gas rather than an insufficient number of ionizing photons.

By applying a theoretical model of the continuum emission to the data, one can determine the size, magnetic field strength, and other properties of the emitting region. The model used here is a modified version of the synchrotron self-Compton model of Jones, O'Dell, and Stein (1974*a*, *b*). Briefly, the emitting region is homogeneous so it is characterized by single values for the magnetic field and particle density. The only modification of the stationary model which we include here is that the emitting region may have a bulk velocity with respect to the rest frame of I Zw 187 directed toward or away from the observer; beaming effects can be included.

The five observables that uniquely define this model are: (1) the power-law slope of synchrotron radiation; (2) the frequency υ_b and (3) flux F_b at which the synchrotron radiation becomes optically thick; (4) a measure of the amount of inverse Compton radiation (E_υ^{sc}); and (5) the flux variability time scale (Fdt/dF).

The power-law slope is defined by the infrared through ultraviolet data ($\alpha = 0.9$). The X-ray data, which are probably not due to inverse Compton radiation, yield an upper limit to the inverse Compton contribution ($E_\upsilon^{sc} < 1$). The quantities υ_b and F_b are not observed directly, but the simultaneous spectra define their allowed ranges. If the rising infrared-ultraviolet power-law slope turns over and becomes flatter (presumably because of self-absorption), then the turnover point defines υ_b and F_b. In practice, one finds these quantities by extrapolating the radio spectra until they intersect the extrapolated infrared-ultraviolet power-law. Using this method, $\upsilon_b \approx 4\text{–}10 \times 10^{11}$ Hz and $F_b \approx 0.2\text{–}0.4$ Jy. The time scale for flux variation, about 5 days (§§ II*a* and II*c*), enables one to include the effects of source motion on the model. Because we suggested that $E_\upsilon^{sc} < 1$, one formally calculates lower limits to the magnetic field $B(G)$, size s(pc), and the motion of the emitting region with respect to the local cosmological rest frame (Lorentz factor γ; Table 3). However, these quantities are so weakly dependent upon E_υ^{sc} (e.g., $s \propto E_\upsilon^{sc-0.1}$) that these lower limits to s and γ are good estimates of their values as long as E_υ^{sc} falls in the range found for other QSOs and BL Lac objects. For X-ray–detected quasars (e.g., survey of Ku, Helfand, and Lucy 1980), $E_\upsilon^{sc} > 10^{-5}$ and typically $E_\upsilon^{sc} \approx 10^{-3}$ if the

TABLE 3

PROPERTIES OF THE NONTHERMAL CONTINUUM EMITTING REGION

E_υ^{sc}	$\log F_b$	$\log \upsilon_b$	θ (mas)	s (pc)	$B(G)$	γ
1.0	−23.4	11.6	5.0 (−3)	4.4 (−3)	8.9 (1)	1.001
1.0 (−3)	−23.4	11.6	8.6 (−3)	5.8 (−3)	1.1 (3)	1.053
1.0	−23.7	12.0	2.0 (−3)	2.8 (−3)	6.1 (2)	1.080
1.0 (−3)	−23.7	12.0	3.5 (−3)	3.7 (−3)	7.4 (3)	1.007

X-rays arise from the inverse Compton process (models in which $E_i^{sc} = 10^{-3}$ are presented in Table 3).

From this analysis, we find that the size of the emitting region is approximately a light-week ($\sim 10^{16}$ cm) and that there is no significant relativistic motion of the emitting region ($\gamma \approx 1$). The magnetic field is equal to or greater than 10^2 *G*, a value substantially larger than the milligauss fields found in earlier analysis of quasar data (Jones, O'Dell, and Stein 1974*b*). A calculation of the energy density of the magnetic field indicates that it is comparable to or greater than the energy density of emitting particles. Because the radio emission, which occurs at frequencies lower than the turnover frequency, does not behave like a self-absorbed source, we suggest that the emission originates in a "tapered" source (i.e., a density gradient in the source causes it to remain luminous but appear larger at lower frequencies). This interpretation is consistent with the VLBI size (> 1.2 milli-arcsec; 5 GHz) being larger than the theoretical size of the optically thin emitting region.

IV. CONCLUSIONS

Simultaneous radio, infrared, optical, ultraviolet, and X-ray observations of I Zw 187 were made during 1979 October and 1980 August. Additional nonsimultaneous data have been accumulated in these observing regions during the past 15 years. We find that:

1. I Zw 187 is less variable than most BL Lac objects at infrared, optical, and X-ray frequencies; it has not been observed to vary in the GHz range.

2. Nonsimultaneous observations show that the range of flux variations and the shortest time scale for variation are approximately the same in the X-ray and optical observing bands.

3. The infrared-optical-ultraviolet spectrum is characterized by a power law of slope 0.9, and the X-ray data fall close to an extrapolation of this power law. This result is consistent with the infrared, optical, ultraviolet, and X-ray emission all having a common origin, probably the synchrotron process.

4. There is a broad, weak bump in the continuum at about 3000 Å, the first indication that this feature may exist in BL Lac objects.

5. The nearly flat radio spectrum and the infrared-optical-ultraviolet spectrum would intersect in the $4–10 \times 10^{11}$ Hz range. The point of intersection may be a measure of the frequency at which the source becomes optically thick to synchrotron radiation.

6. When the simultaneous spectra are analyzed in terms of a synchrotron self-Compton model, we find that the size of the optically thin emitting region is approximately 10^{16} cm, that the magnetic field is equal to or greater than 10^2 *G*, and that there is no significant bulk motion of the emitting region.

Many people helped make these observations possible. G. Neugebauer would like to thank S. E. Persson and K. Mathews; infrared astronomy at Palomar Observatory is supported by NASA grants. Peter Forster and the staff at UKIRT provided valuable assistance to P. M. Williams and P. W. J. L. Brand. R. J. Rudy and P. D. LeVan were aided by the staff at Mount Lemmon. H. D. Aller, M. F. Aller, and P. E. Hodge thank the staff of the University of Michigan Radio Observatory for their assistance and the NSF for their support. T. J. Balonek and W. A. Dent thank the staff support at the 36′ NRAO telescope and at the Haystack Observatory; the Haystack Observatory is supported by the NSF. D. J. Helfand was of great assistance in the reduction and interpretation of the X-ray data. W. -M. Ku would like to thank the Columbia Astrophysics Laboratory and NASA for their support (Contract NAS 8-30753). J. T. Pollock, A. J. Pica, A. G. Smith, and J. R. Webb acknowledge the continuing support of the NSF; the current grant is AST 8000 246. Finally, J. N. Bregman, A. E. Glassgold, and P. J. Huggins are grateful to A. Boggess for the discretionary time to begin this project, and to C. C. Wu, A. Holm, S. Schiffer, and the staff of the *IUE* for their assistance. A. Kinney wrote the data reduction programs that made analysis of *IUE* spectra possible. *IUE* observations were made possible with the aid of NASA grant NAG 5-73. Finally, the authors would like to thank the referee for several helpful suggestions.

REFERENCES

Aller, H. D. 1970, *Ap. J.*, **161**, 1.
Aller, H. D., Aller, M. F., and Hodge, P. E. 1981, *A.J.*, **86**, 325.
Altschuler, D. R., and Wardle, J. F. C. 1975, *Nature*, **255**, 306.
Bregman, J. N., Ford, H. C., Glassgold, A. E., and Huggins, P. J. 1982, in preparation.
Burstein, D., and Heiles, C. 1978, *Ap. J.*, **225**, 40.
Chubb, T. A. 1978, *Pittsburgh Conference on BL Lac Objects*, ed. A. M. Wolfe (Pittsburgh: University of Pittsburgh), p. 163.
Frogel, J. A., Persson, S. E., Aaronson, M., and Matthews, K. 1978, *Ap. J.*, **220**, 75.
Holm, A. 1980, private communication.
Jones, T. W., O'Dell, S. L., and Stein, W. A. 1974*a*, *Ap. J.*, **188**, 353.
———. 1974*b*, *Ap. J.*, **192**, 261.
Kinman, T. D. 1976, *Ap. J.*, **205**, 1.
———. 1978, *Pittsburgh Conference on BL Lac Objects*, ed. A. M. Wolfe (Pittsburgh: University of Pittsburgh), p. 82.
———. 1980, private communication.
Ku, W. -M. 1982, in preparation.
Ku, W. -M, Helfand, D. J., and Lucy, L. B. 1980, *Nature*, **288**, 323.
Kwan, J., and Krolik, J. H. 1981, *Ap. J.*, **250**, 478.
LeSquéren, A. M., Biraud, F., and Lauqué, R. 1972, *Nature*, **240**, 75.
Miller, J. B. 1981, private communication.
Neugebauer, G., Oke, J. B., Becklin, E. E., and Matthews, K. 1979, *Ap. J.*, **230**, 79.
Oke, J. B. 1978, *Ap. J.* (*Letters*), **219**, L97.
Oke, J. B., Sargent, W. L. W., Neugebauer, G., and Becklin, E. E. 1967, *Ap. J.* (*Letters*), **150**, L174.
Pica, A. J., Pollock, J. T., Smith, A. G., Leacock, R. J., Edwards, P. L., and Scott, R. L. 1980, *A.J.*, **85**, 1442.
Sandage, A. 1967, *Ap. J.* (*Letters*), **150**, L177.
———. 1972, *Ap. J.*, **173**, 485.
Sargent, W. L. W. 1970, *Ap. J.*, **160**, 405.
Seaton, M. J. 1980, *M.N.R.A.S.*, **187**, 73P.
Schild, R., and Oke, J. B. 1971, *Ap. J.*, **169**, 209.

Stein, W. A., O'Dell, S. L., and Strittmatter, P. A. 1976, *Ann. Rev. Astr. Ap.*, **14**, 173.
Strom, K. M., and Strom, S. E. 1978, *A.J.*, **83**, 1293.
Strom, S. E., and Strom, K. M. 1978, *A.J.*, **83**, 732.
Strom, S. E., Strom, K. M., Goad, J. W., Vrba, F. J., and Rice, W. 1976, *Ap. J.*, **204**, 684.
Tolbert, C. R. 1972, *Astr. Ap. Suppl.*, **3**, 349.
Weiler, K. W., and Johnson, K. J. 1980, *M.N.R.A.S.*, **190**, 269.
Weisheit, J. C., Shields, G. A., and Tarter, C. B. 1981, *Ap. J.*, **245**, 406.
Weistrop, D., Shaffer, D. B., Mushotzky, R. F., Reitsema, H. J., and Smith, B. A. 1981, *Ap. J.*, **249**, 3.
Whitford, A. E. 1971, *Ap. J.*, **169**, 215.
Zwicky, F. 1966, *Ap. J.*, **143**, 192.

H. D. Aller, M. F. Aller, and P. E. Hodge: Department of Astronomy, University of Michigan, Ann Arbor, MI 48109

T. J. Balonek and W. A. Dent: Department of Physics and Astronomy, GR Tower B, University of Massachusetts, Amherst, MA 01003

P. W. J. L. Brand: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

Joel N. Bregman, A. E. Glassgold, and P. J. Huggins: Physics Department, New York University, 4 Washington Place, New York, NY 10003

W. -M. Ku: Columbia Astrophysics Laboratory, Columbia University, New York, NY 10027

P. E. LeVan and R. J. Rudy: Center for Astrophysics and Space Science (CASS), C-011, University of California, San Diego, La Jolla, CA 92093

G. Neugebauer: Hale Observatories, California Institute of Technology, Pasadena, CA 91125

A. J. Pica, J. T. Pollock, A. G. Smith, and J. R. Webb: Rosemary Hill Observatory, University of Florida, Gainesville, FL 32611

P. M. Williams: U.K. Infrared Telescope Unit, 900 Leilani Street, Hilo, Hawaii 96720

VARIABILITY OF BL LAC OBJECTS AT X-RAY AND OTHER FREQUENCIES

A.M.T.Pollock[1], P.W.J.L.Brand[2], J.L.Bregman[3] and E.I.Robson[4]
[1]Dept.of Space Research, University of Birmingham, U.K.
[2]Dept. of Astronomy, University of Edinburgh, Scotland
[3]National Radio Astronomy Observatory, Virginia, U.S.A.
[4]Dept. of Astronomy, Preston Polytechnic, U.K.

ABSTRACT. EXOSAT observations reveal that daily X-ray variability appears to be a common property of high luminosity BL Lac objects. Results are reported for PKS0735+178, OM280(=1147+245), OQ530(=1418+546), OV-236 (=1921-293) and OJ287(=0851+202). In OJ287, for which we have the most extensive coverage, the amplitude of variability at X-ray frequencies appears to be a factor of a few larger than at longer wavelengths.

1. INTRODUCTION.

The high variability and polarisation of the BL Lac objects observed throughout the radio, infra-red and optical shows without doubt the non-thermal, unscattered nature of this radiation (e.g.Angel and Stockman (1980)). These properties make them prime objects to search for the correlated changes at different wavelengths that have provided powerful constraints on models for the production of electromagnetic radiation by active galactic nuclei, of which the BL Lac objects provide the most unobscured view. Indeed Allen et al (1982) go further and suggest that a BL Lac object lies at the centre of every active galaxy.

This report is presented on behalf of many colleagues at our own and other institutes who have contributed observing time to two programmes of work. The first, involving Birmingham and Edinburgh, was made in an effort to see if the daily variability observed notably in the infra-red (Impey et al (1982) et seq.) is a common property at X-ray wavelengths. The second is a wider collaboration following the longer-term behaviour of one object, OJ287.

2. DAILY VARIABILITY OF HIGH-LUMINOSITY BL LAC OBJECTS

Observations have been made of the five 'high-luminosity' BL Lac objects detailed below with the 0.1-2 keV CMA telescope aboard EXOSAT on three consecutive days. Each was observed through 3000Å lexan filter (3LX), which is the most sensitive for weak sources such as these, with measurements also made with the aluminium/parylene (AL/P) and boron (B) filters

Space Science Reviews 40 (1985) 607–612. 0038–6308/85.15.

to get some spectral information, when the count rates permitted.

Object		Z	Date of EXOSAT Observations	Filters Used
PKS	0735+178	>0.424	1983 12-14 Nov	3LX
OM280	1147+245	?	1984 16-18 Jan	3LX
OJ287	0851+202	?0.306	1984 7-9 Feb	3LX,AL/P,B
OQ530	1418+546	?	1984 18-20 May	3LX,AL/P,B
OV-236	1921-293	0.3525	1984 12-14 Oct	3LX

On each occasion simultaneous infra-red observations were scheduled at UKIRT although weather and other problems prevented anything like complete coverage. The 3LX data are plotted in figure 1 with comments on the individual objects below. For each set of intensities a probability p(var), calculated following Pollock et al (1984), is quoted giving the probability that the object is variable.

(i) PKS 0735+178 p(var) = 0.92

Bregman et al (1984) have published several multifrequency observations of this object in which large changes up to 350% in the IR-UV continuum were accompanied by modest radio frequency changes and little or no change in the Einstein IPC X-ray flux (<50%). Radio, IR and optical measurements on day 3 show the source spectrum was very close to the faintest of Bregman et al's. The mean X-ray flux we detect, 0.15 μJy at 1 keV for a reasonable choice of spectral form, is also compatible with the earlier results. Although the source is weak in the CMA giving about 10 photons an hour there is modest evidence of variability.

(ii) OM280 p(var) = 0.40

1147+245 was constant over the three days with a 3LX count rate of $(0.30\pm0.03) \times 10^{-2}$ s^{-1}.

(iii) OJ287 p(var) = 0.998

The object is X-ray variable over the three day period at a high level of confidence resulting from a 30% increase from 0.7 to 0.9 μJy at 1 keV. Simultaneous IR measurements on the first two days showed an increase which averaged (6±2)% over J,H,K and L including 22.5 to 25.0 μJy at K. These data are discussed more fully below.

(iv) OQ530 p(var) = 0.991

There were four 3LX exposures including two separated by five hours. The intensity of the earlier of these two is about twice the level of the remaining three with variability well established. A weak detection in AL/P and an upper limit in an 11000 second B exposure show the spectrum is steep with number index <-2. The IR K-fluxes of 11.5±0.1 μJy on day 2 and 10.7±0.1 μJy on day 3 show the object at very nearly the same bright-

ness as the 1980 December multifrequency spectrum of Worrall et al (1984); an extrapolation of their UV spectrum to 1 keV also agrees well with the flux we have measured so the X-rays are probably of synchrotron origin.

(v) OV-236 p(var) = 0.993

1921-293 was significantly lower on day 2 by a factor of two.

Overall these results show that daily variability at X-ray wavelengths is a common property of these relatively distant members of their class. For those objects whose redshift is known the implied luminosity changes assuming isotropic emission fall far short of violating the Fabian-Rees limit (e.g. Fabian and Rees (1979)).

3. MORE EXTENSIVE OBSERVATIONS OF OJ287

In addition to the three day period in February OJ287 was observed on three other occasions with EXOSAT in the six-month period 1983 October - 1984 April. These data are combined and presented in figure 2 with many other measurements at UV, optical, IR and radio frequencies. Many were simultaneous. The optical data are from Webb and colleagues at the University of Florida; the radio data from M. Aller and colleagues at the University of Michigan. The intensity scales are logarithmic to allow a comparison at different frequencies of the amplitude of variability, which is undoubtedly strongly present throughout. Although the UV have only been analysed in a preliminary way it is clear that the X-ray flux falls below the extrapolation of the UV spectrum by a factor of a few indicating some curvature in this part of the spectrum. Worrall et al's (1982) non-simultaneous spectrum was lower in the UV and higher at 1 keV but the data were taken some months apart and are therefore probably not in disagreement with ours.

The multi-frequency data appear to show that all wavelengths except the radio vary in phase but with different amplitudes. The ratio of 5 reported above of the increases of X-rays and IR in a day is approximately reproduced as the X-rays subsequently fall by a factor of 6-7 by April 21 while the IR decrease is limited to a factor of 2. The X-ray variability is the greatest, followed somewhat surprisingly by the optical, and then the UV and IR. This is in complete contrast to the behaviour of PKS 0735-178 (Bregman et al (1984)). Exposures with three CMA filters at the X-ray peak show a steep power law slope of energy index $-(2\pm1)$ and $n_H \sim 3 \times 10^{20}$ cm^{-2}. Other filter ratios are consistent with this result although a steepening of the spectrum as the intensity falls cannot be ruled out at all.

The data seem consistent with a simple scheme in which synchrotron radiation is the mechanism responsible for the production of IR to X-ray radiation. Increases in flux follow fresh injections of relativistic particles while the energy dependent energy-loss time scales dictate the subsequent evolution of the spectrum. This could ensure variability of a similar nature for both upward and downward excursions. Although

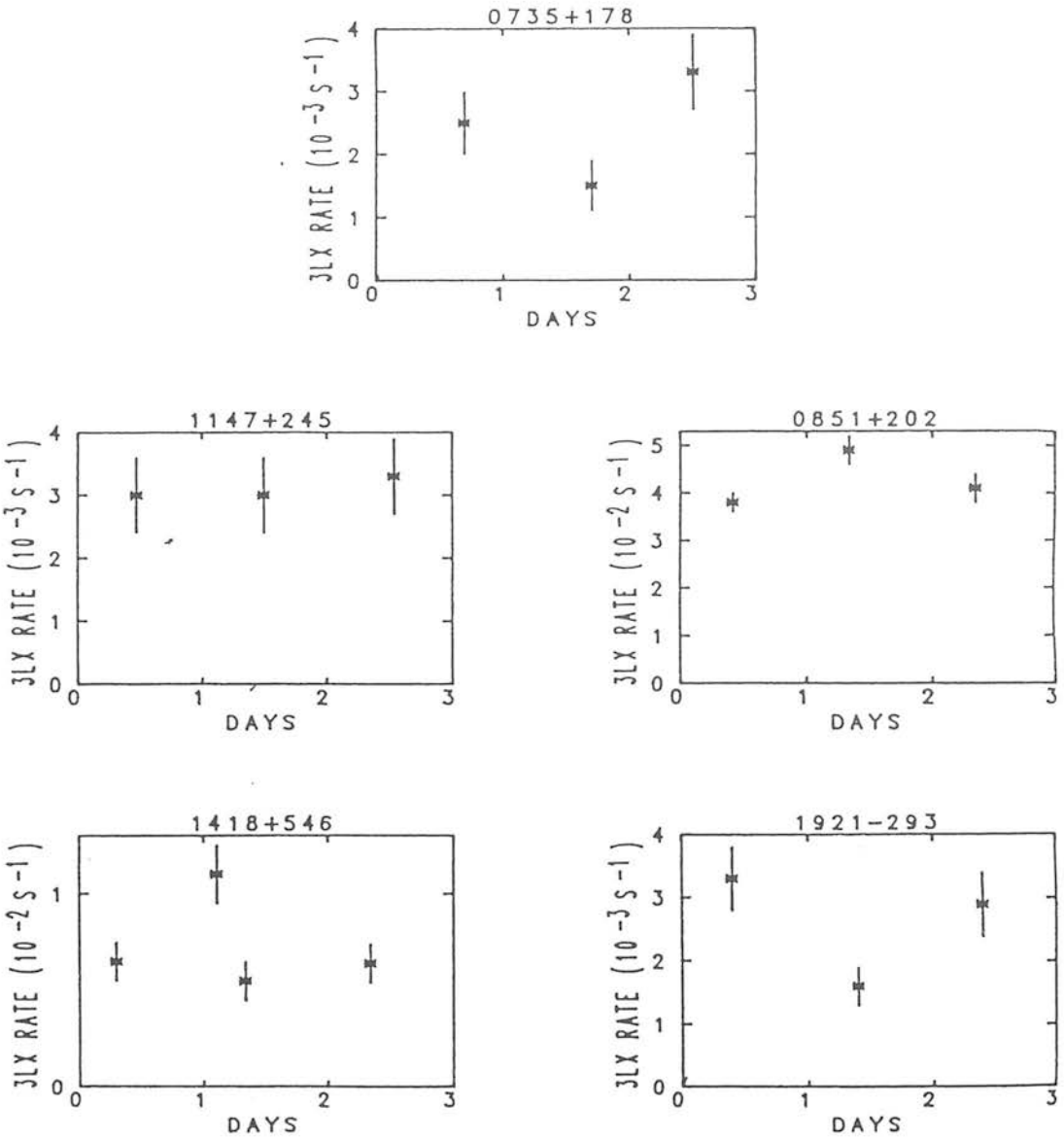


Figure 1 : Count rates observed in the 3000Å-lexan-filtered EXOSAT CMA instrument for the five BL Lac objects which were observed on three consecutive days.

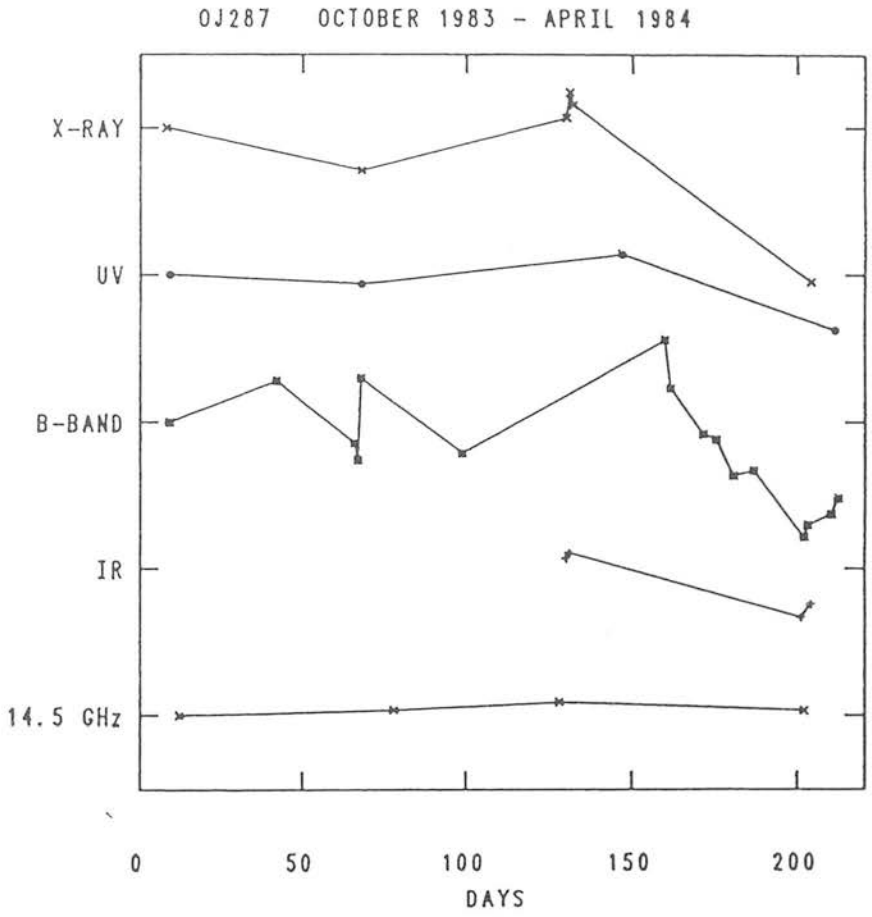


Figure 2 : Measurements at different frequencies of the BL Lac object OJ287 during the period October 1983 to April 1984. The X-rays are EXOSAT CMA count rates; the UV from IUE; B-band monitoring from the University of Florida; IR measurements from UKIRT; and the radio fluxes are courtesy M.Aller of the University of Michigan. The scales are logarithmic to allow comparison of the amplitude variations.

the coverage at different wavelengths in figure 2 is good it is not good enough to positively support such an interpretation and many more similar measurements are needed to establish the acceleration and cooling time scales which operate in OJ287.

ACKNOWLEDGEMENTS

Many thanks to Dr. Margo Aller of the University of Michigan for the use of the radio data and to the Rosemary Hill observers of the University of Florida for the use of the optical data.

REFERENCES

Allen D.A. et al, 1982. MNRAS, 199, 969
Angle J.R.P. and Stockman H.S., 1980. Ann.Rev.Astr.Astrophys., 18, 321
Bregman J.N. et al, 1984. Ap.J., 276, 454
Fabian A.C. and Rees M.J., 1979. X-ray Astronomy, ed. Baity and Peterson, p381
Impey C.D. et al, 1982. MNRAS, 200, 19
Pollock A.M.T. et al, 1984. Astron.Astrophys. In press
Worrall D.M. et al, 1982. Ap.J., 261, 403
Worrall D.M. et al, 1984. Ap.J., 284,

THE ASTROPHYSICAL JOURNAL, 313:662–673, 1987 February 15

SIMULTANEOUS MULTIFREQUENCY OBSERVATIONS OF THE BL LACERTAE OBJECT MARKARIAN 421

F. MAKINO,[1] Y. TANAKA,[1] M. MATSUOKA,[1] K. KOYAMA,[1] H. INOUE,[1] K. MAKISHIMA,[1] R. HOSHI,[2] S. HAYAKAWA,[3] Y. KONDO,[4] C. M. URRY,[5] S. L. MUFSON,[6] K. R. HACKNEY,[7] R. L. HACKNEY,[7] S. KIKUCHI,[8] Y. MIKAMI,[8] W. Z. WIŚNIEWSKI,[9] N. HIROMOTO,[10] M. NISHIDA,[11] J. BURNELL,[12] P. BRAND,[12] P. M. WILLIAMS,[13] M. G. SMITH,[13] F. TAKAHARA,[14] M. INOUE,[14] M. TSUBOI,[15] H. TABARA,[16] T. KATO,[16] M. F. ALLER,[17] AND H. D. ALLER[17]

Received 1986 February 24; accepted 1986 August 8

ABSTRACT

Simultaneous multifrequency observations of the BL Lac object Mrk 421 covering radio through X-ray wavelengths were performed on two occasions separated by 5 weeks in 1984 January and March, and each observation was coordinated for about 1 week. We obtained composite multifrequency spectra of the central nonthermal component at the two epochs after subtracting the optical and infrared light of the underlying galaxy. The spectra show the gradual steepening toward high frequency; the power law indices are ~ 0.1, ~ 0.6, and ~ 1.0 for radio, infrared-optical, and UV bands, respectively. The UV and optical-infrared fluxes decreased by $\sim 20\%$ in 5 weeks, while the radio flux remained stable. The X-ray flux decreased by a factor of ~ 2, and the change was more pronounced at hard X-rays, which suggest that X-ray emission possibly consists of two components. The degree of polarization at the optical band varied on the time scale of a few days, while the position angle remained unchanged.

Physical parameters of Mrk 421 are discussed in terms of the synchrotron self-Compton model. Taking the spectral turnover between infrared and radio for synchrotron self-absorption, the radio emission originates in a more extended region than the infrared to X-ray emission, the source size of which should be less than $\sim 10^{-2}$ milliarcseconds. Relativistic beaming is required if the angular size is smaller than a few times 10^{-3} milliarcseconds. A possible explanation of the spectral change during the two epochs is also discussed.

Subject headings: BL Lacertae objects — galaxies: individual — galaxies: X-rays — polarization — radiation mechanisms — radio sources: galaxies

I. INTRODUCTION

The central problem of active galactic nuclei (AGNs) is the emission mechanism of the continuum radiation at radio through gamma-ray frequencies, and the ultimate energy source of that radiation. Apparently these AGNs somehow generate very high luminosities (assuming cosmological distances and isotropic emission) inside very small volumes (as deduced from variability time scales). BL Lac objects are distinguished from other classes of AGNs by their rapid variability, high degree of polarization, and general lack of discrete spectral features. These properties are useful diagnostics of the radiation processes. Simplified models such as the synchrotron self-Compton model with or without relativistic beaming have been proposed to explain many of the observed characteristics of BL Lac objects and other AGNs. A high degree of polarization in the infrared and optical region may be evidence of a synchrotron process, a hypothesis that can be strengthened if correlated temporal changes in radio, infrared, and optical bands are detected. Many AGNs have been observed simultaneously in various frequency bands from radio to X-ray frequencies in order to search for correlated variations within an individual spectrum (Kondo *et al.* 1981; Bregman *et al.* 1982, 1984; Worrall *et al.* 1982, 1984*a*, *b*, *c*; Wills *et al.* 1983; Mufson *et al.* 1984; Glassgold *et al.* 1983; Hutter and Mufson 1986). The absence of discrete spectral features, often a nuisance when determining the physical state of the emitting region, in this case provides a conveniently uncontaminated broad-band spectrum.

Mrk 421 was first noted as a blue excess object which turned out to be an elliptical galaxy with a bright, pointlike nucleus. The redshift of the galaxy was measured to be $z = 0.03$ (Ulrich *et al.* 1975), but the spectrum of the nucleus was featureless. It also showed the optical polarization and flat radio spectrum, common to BL Lac objects (Maza, Martin, and Angel 1978; O'Dell *et al.* 1978*a*). The angular size of the radio source associated with Mrk 421 is less than 1 milliarcsecond (mas) at 6 cm (Weiler and Johnston 1980; Baath *et al.* 1981). The ultraviolet emission from Mrk 421 has been monitored over the past 7 yr with the *International Ultraviolet Explorer* (*IUE*) satellite. Nine *IUE* spectra obtained between 1978 and 1981 showed some variability, but no or only very weak correlation between flux level and spectral index (Ulrich *et al.* 1984). Since it is one of the brightest BL Lac objects known in the X-ray

[1] Institute of Space and Astronautical Science, Tokyo.
[2] Department of Physics, Rikkyo University.
[3] Department of Astrophysics, Nagoya University.
[4] NASA/Goddard Space Flight Center.
[5] Center for Space Research, Massachusetts Institute of Technology.
[6] Astronomy Department, Indiana University.
[7] Department of Physics and Astronomy, Western Kentucky University.
[8] Tokyo Astronomical Observatory, University of Tokyo.
[9] Lunar and Planetary Laboratory, University of Arizona.
[10] Radio Research Laboratory, Tokyo.
[11] Department of Physics, Kyoto University.
[12] Astronomy Department, University of Edinburgh.
[13] Royal Observatory Edinburgh.
[14] Nobeyama Radio Observatory, University of Tokyo. The Nobeyama Radio Observatory, a branch of the Tokyo Astronomical Observatory, University of Tokyo, is a cosmic radio observing facility open for outside users.
[15] Department of Astronomy, University of Tokyo.
[16] Faculty of Education, Utsunomiya University.
[17] Radio Astronomy Observatory, University of Michigan.

band, Mrk 421 has been observed fairly often with various X-ray instruments (see summary of observations and references in Urry, Mushotzky, and Holt 1986). X-ray observation of Mrk 421 with the *HEAO 1* satellite detected a change in spectral index from 2.1 to 3.9 and the disappearance of the hard component above 10 keV by comparing with earlier observations from the *OSO 8* and *SAS 3* satellites (Mushotzky *et al.* 1979). Because of the lack of simultaneity among the observations at radio through X-ray frequencies, correlations among intensity changes have been poorly studied.

We made simultaneous multifrequency observations of the BL Lac object Mrk 421 at radio, infrared, optical, ultraviolet, and X-ray frequencies on two occasions in 1984 January and March. In § II we summarize observations performed at each frequency band, and we describe the composite spectrum thus obtained at each of the two epochs. In § III interrelations among various observed bands and the comparisons with previous observations are discussed. In § IV an interpretation of those results in terms of the synchrotron self-Compton model is attempted. Conclusions will follow in § V.

II. OBSERVATIONS

The observation was conducted at the radio, infrared, optical, ultraviolet, and X-ray bands on two occasions separated by 5 weeks, 1984 January and March. The coverage of the observed frequency bands and the instruments utilized in the present observations are summarized in Table 1.

a) X-Ray Observations

X-rays were observed with two sets of the scintillation proportional counters (SPC-A and SPC-B) on board the *Tenma* satellite (Tanaka *et al.* 1984; Koyama *et al.* 1984). The observed energy range was 1.5–35 keV, but above ~ 10 keV, the flux was too faint to discriminate from the background. The day-to-day variations of the integrated energy flux over the energy range 1.5–10 keV are depicted in the lowest panel of Figure 1. The statistical accuracy of day-to-day data was not good enough to obtain the energy spectrum. Then we summed up the data for each observation period; the time-averaged energy spectra observed in January and March are shown in Figure 2.

The flux decreased by a factor of ~ 2 from January to March. Particular attention has been paid to the spectral analysis, since for faint sources such as Mrk 421 the spectral shapes depend critically on background subtraction. The properties of the background spectrum have long been studied carefully using extensive data sets compiled from *Tenma* (Koyama *et al.* 1984). Before summing the data from SPC-A and SPC-B, the spectral analysis was independently carried out for each datum to check the consistency of the results. The spectra obtained independently from SPC-A and SPC-B were combined to improve statistics, after confirmation of consistency. The best-fit power-law index for the January and March observations over 1.5–10 keV is 1.1 ± 0.2 and 1.8 ± 0.3, with the normalization factor $(1.01 \pm 0.17) \times 10^{-2}$ keV cm^{-2} s^{-1} keV^{-1} and $(8.7 \pm 2.4) \times 10^{-3}$ keV cm^{-2} s^{-1} keV^{-1} at 1 keV, respectively.

It is, however, to be noticed that the single power-law fit is only marginally acceptable, since the reduced χ^2 value is 2.3 (degree of freedom, 11) and 2.2 (degree of freedom, 7) for the January and March data, respectively. The deviations from the single power-law models are evident above 4 keV. The spectrum obtained in January is flatter above 4 keV as compared to March observations. Below 4 keV, on the other hand, the spectral index is 1.6 ± 0.1 and 1.7 ± 0.2, respectively, while the intensity varied by a factor of 2.

We notice that the *EXOSAT* satellite also observed Mrk 421 in the period 1984 February 1–6, 1 week after our January observation (Warwick, McHardy, and Pounds 1985). The intensity and spectral index are intermediate between those of our January and March observations.

b) UV Observations

Ultraviolet observations were made by the short-wavelength prime (SWP) and the long-wavelength prime (LWP) cameras of the *IUE* satellite on five occasions during the present program. The results are summarized in Table 2. The integrated fluxes are plotted in the fifth panel of Figure 1. The

TABLE 1

ORGANIZATION OF OBSERVATIONS

Observatory	Band	Observers
Tenma	X-ray 1.5–30 keV	F. M., Y. T., M. M., K. K., H. I., K. M., R. H., S. H.
IUE	UV 1200–3000 Å	Y. K., C. M. U., S. L. M., K. R. H., R. L. H.
Mt. Lemon	Optical *U, B, V, R, I*	W. Z. W.
Okayama ..... Dodaira	Optical *U, B, V, R* with polarimetry	S. K., Y. M.
Agematsu	IR *J, H, K*	N. H., M. N.
UKIRT	IR *J, H, K, L, L', M, N* with *J, H, K* polarimetry	J. B., P. B., P. M. W., M. G. S.
NRO	Radio 10, 22, 43 GHz	F. T., M. I., M. T., H. T., T. K
Michigan	Radio 4.8, 8.0, 14.5 GHz	M. F. A., H. D. A.

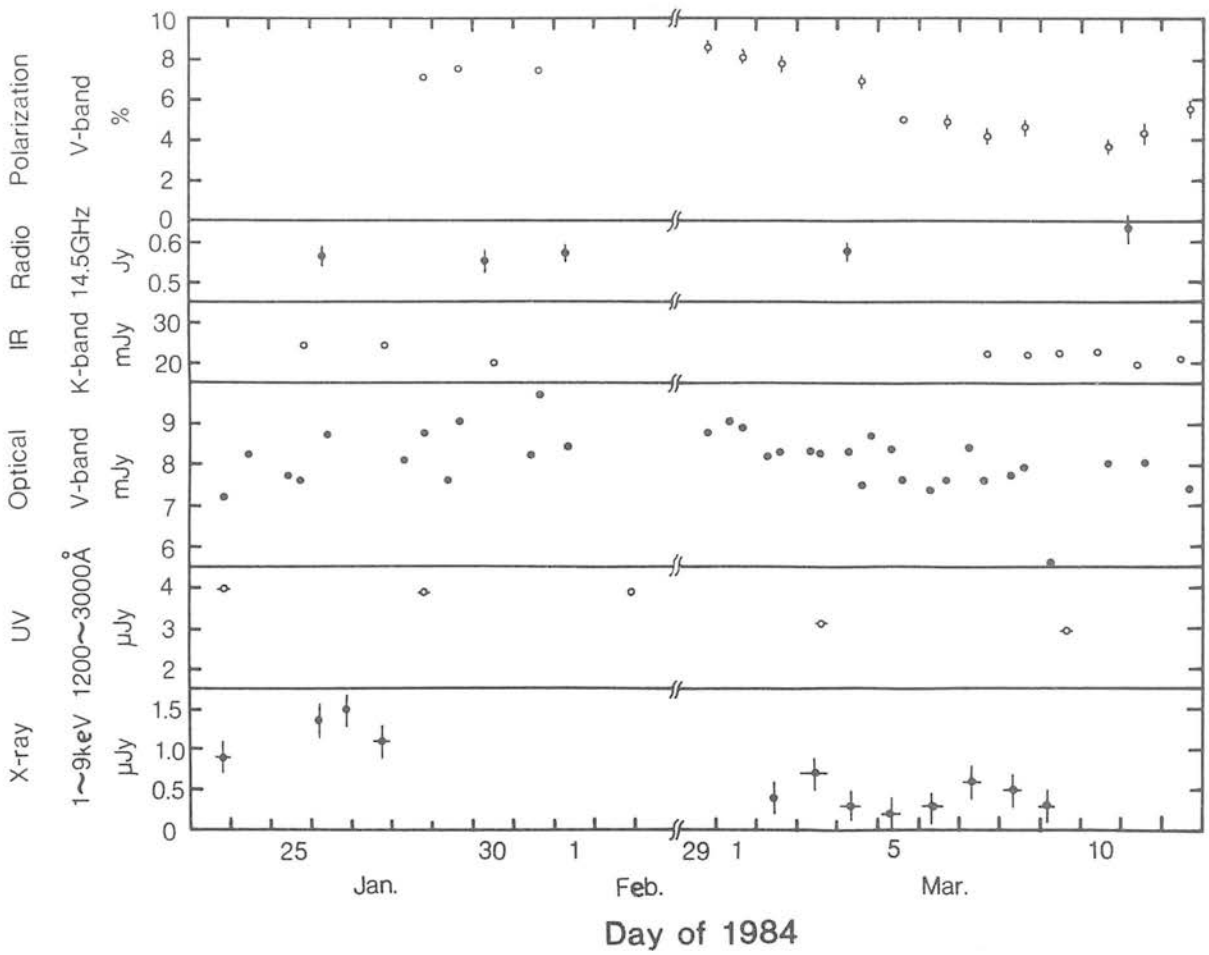


FIG. 1.—Light curves of the nonthermal component of Mrk 421. First panel shows that of the degree of polarization at the V band. Following panels show that fluxes in the radio (14.5 GHz), IR (K band), optical (V band), UV and X-ray frequencies, respectively.

TABLE 2

RESULTS OF UV OBSERVATIONS BY *IUE*

Data (1984)	Image	A^a (mJy)	α^b	χ^2/DOF	Integrated Flux[c]
Jan 23.....	LWP 2669 LWP 2700 SWP 22082	6.65 ± 0.12	1.04 ± 0.04	0.71	6.01 ± 0.11
Jan 28.....	LWP 2711 LWP 2712 SWP 22128	6.35 ± 0.20	1.01 ± 0.07	0.55	5.82 ± 0.11
Feb 2	LWP 2732	6.77 ± 0.15	1.13 ± 0.20	0.11	5.88 ± 0.18
Mar 3	LWP 2881 LWP 2882 SWP 22398	5.24 ± 0.18	1.10 ± 0.08	1.09	4.61 ± 0.16
Mar 9	LWP 2915 SWP 22445	5.04 ± 0.21	1.12 ± 0.08	0.47	4.38 ± 0.14

[a] Flux density at 3000 Å.
[b] Spectral index for 1200–3000 Å.
[c] Integrated over 1200–3000 Å and in units of 10^{-11} ergs cm^{-2} s^{-1}.

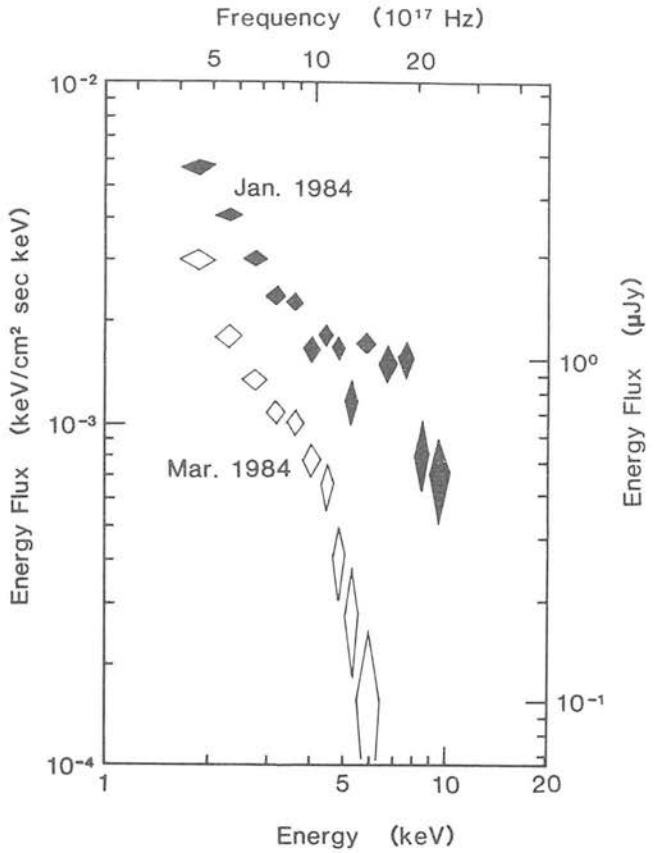


FIG. 2.—X-ray spectra obtained by *Tenma*. Data were integrated over each observation period.

spectra were extracted from the line-by-line file using a customized narrow slit to maximize the signal to noise. Data points contaminated by reseaux or cosmic-ray hits were removed, and the data were then binned in intervals of ~25 Å. The binned data points were weighted according to the rms deviations within the bins and were fitted to a power-law spectrum using a least-squares procedure for 1200–3000 Å range.

It is to be noted that the integrated ultraviolet flux decreased by 20%–30% from January to March, and the spectral index α shows a slight steepening (a change of α of ~0.1).

c) *Optical Observations*

Optical observations were made at the Okayama Astrophysical Observatory, the Dodaira Station, and the Mount Lemon Observatory. Five telescopes were used with various diaphragm sizes, which are listed in Table 3.

Mrk 421 is found at the core of a giant elliptical galaxy, and, in particular, optical and infrared fluxes are contaminated by galaxy emission. Therefore, we must subtract the galaxy component to obtain the flux from the central compact core. The distribution of the surface brightness of the galaxy was studied by Kinman (1978), Mufson *et al.* (1980), and Hickson *et al.* (1982), but some discrepancies remained among their results. Kikuchi and Mikami (1986) also made the multiaperture photometry of Mrk 421 and obtained a result similar to that by Kinman (1978). In the present study, the subtraction of the contribution from the underlying galaxy has been made using the results by Kikuchi and Mikami (1986). The corrections for the apertures used in the observations are also tabulated in Table 3. The colors of the underlying galaxy are taken to be constant as $U-B=0.46$, $B-V=1.05$, and $V-R=0.90$ after Kikuchi and Mikami (1985). These colors are consistent with those of a redshifted elliptical galaxy at $z=0.03$. For $R-I$, the mean value of 0.78 for elliptical galaxies (Johnson 1966) is assumed, since the measurement of the underlying galaxy has yet to be made at the I band.

The observed fluxes and the fluxes of the nonthermal component are tabulated in Table 4. The day-to-day variation of the flux of the nonthermal component at the V band is shown in the fourth panel of Figure 1. The nominal errors associated with the observed fluxes are typically 2%–3% but 1% for several cases at Dodaira in March. However, the derived flux of the nonthermal component shows a scatter greater than the observational errors, reaching ~10%, depending on the aperture and/or telescope. A part of this scatter is certainly caused by the incompleteness of the subtraction of the galaxy contribution. Although we have no reasonable explanation for the remaining scatter, we suspect that the seeing condition and the scattered light of a nearby 6 mag star affect the measurement of the extended source and of the sky brightness, respectively.

Continuum spectra for both observing periods, as well as those in the infrared region, are shown in Figure 3. In the January data of the fluxes in the B and V bands, we find a systematic difference between the data obtained by the 100 cm telescope at Mount Lemon and those by the 188 cm telescope at Okayama, although the spectral indices coincided well with each other. The difference amounts to 7%, and the origin of the difference is unknown. For the spectra in the March epoch, we used the results after March 4, since the optical flux abruptly decreased around March 3 but was stable thereafter.

The average V band intensities in January and March were 8.9 ± 0.5 mJy and 7.8 ± 0.5 mJy, respectively. The spectral

TABLE 3

LIST OF TELESCOPES FOR OPTICAL OBSERVATIONS AND APERTURE CORRECTION FOR THE UNDERLYING GALAXY

Observatory	Diameter (cm)	Aperture	Date (1984)	Aperture Correction V Magnitude
Okayama	91	37.″1	Jan 23	13.82
		27.9	Jan 29, Mar 3, 4, 6	14.05
	188	8.9	Jan 29, 31	14.98
		11.8	Jan 28	14.72
Dodaira	91	18	Feb 7, 29	14.39
		13	Feb 6, 9, Mar 1–12	14.65
Mt. Lemon	100	29.7	Jan 24–Feb 1, Mar 9	14.00
	150	12.4	Mar 1–8	14.68

TABLE 4

OBSERVED OPTICAL FLUXES AND DERIVED FLUXES OF NONTHERMAL COMPONENT

DATE 1984 (UT)	APERTURE	OBSERVED					NONTHERMAL				
		$F(U)$	$F(B)$	$F(V)$ (mJy)	$F(R)$	$F(I)$	$F(U)$	$F(B)$	$F(V)$ (mJy)	$F(R)$	$F(I)$
Okayama											
Jan 23.80	37".1	8.96	12.25	18.69	...	...	7.61	7.51	7.22	...	...
25.74	27.9	8.40	11.70	16.89	...	...	7.31	7.86	7.61	...	...
28.80	11.8	7.94	9.91	13.79	19.52	...	7.36	7.84	8.79	10.60	...
29.64	8.9	7.73	9.56	13.05	17.81	...	7.27	7.93	9.11	10.78	...
31.65	8.9	8.40	10.38	13.67	17.48	...	7.93	8.75	9.73	10.46	...
Mar 3.53	27.9	7.38	11.07	17.53	...	...	6.29	7.23	8.24	...	...
4.82	27.9	7.05	10.58	18.02	...	...	5.96	6.74	8.73	...	...
6.74	27.9	7.31	10.29	16.89	...	...	6.23	6.45	7.61	...	...
Mt. Lemon											
Jan 24.43	29.7	...	11.38	18.02	...	...	.....	7.36	8.30	...	...
25.40	29.7	...	11.07	17.53	...	...	...	7.05	7.81	...	...
26.43	29.7	...	11.70	18.52	...	...	...	7.68	8.80	...	...
28.33	29.7	...	10.77	17.85	...	...	...	6.75	8.13	...	...
29.39	29.7	...	11.38	17.37	...	...	...	7.36	7.65	...	...
31.44	29.7	...	11.38	18.02	...	...	...	7.36	8.30	...	...
Feb 1.34	29.7	...	11.70	18.18	...	...	...	7.68	8.46	...	...
Mar 1.35	12.4	7.25	9.73	14.31	20.44	27.64	6.64	7.58	9.12	11.18	12.31
2.26	12.4	6.79	9.38	13.42	20.26	27.39	6.18	7.23	8.22	11.00	12.06
3.34	12.4	6.67	9.04	13.54	19.52	27.14	6.06	6.89	8.35	10.26	11.81
4.31	12.4	6.67	8.88	13.54	19.34	26.64	6.06	6.73	8.35	10.09	11.31
5.36	12.4	...	9.04	14.05	...	...	...	6.89	8.85	...	...
6.30	12.4	6.25	8.64	12.58	18.82	25.45	5.64	6.49	7.38	9.56	10.11
7.26	12.4	6.73	8.88	13.67	19.70	25.68	6.12	6.73	8.47	10.45	10.35
8.30	12.4	6.25	8.56	12.93	18.65	25.92	5.64	6.41	7.74	9.39	10.59
9.29	29.7	...	9.38	15.27	...	...	...	5.36	5.54	...	...
Dodaira											
Feb 6.79	13.0	8.27	10.67	15.23	22.08	...	7.65	8.46	9.89	12.56	...
7.83	18.0	9.42	11.95	17.55	25.25	...	8.63	9.14	10.76	13.15	...
9.78	13.0	8.58	10.86	15.91	22.66	...	7.96	8.65	10.56	13.14	...
29.76	18.0	7.72	10.29	15.56	22.43	...	6.92	7.48	8.77	10.34	...
Mar 1.65	13.0	7.29	9.68	14.22	20.46	...	6.66	7.47	8.87	10.94	...
2.60	13.0	7.03	9.26	13.62	19.61	...	6.40	7.05	8.28	10.09	...
4.63	13.0	6.25	8.40	12.82	18.80	...	5.62	6.19	7.48	9.28	...
5.64	13.0	6.36	8.55	12.93	18.80	...	5.74	6.34	7.59	9.28	...
6.71	13.0	6.30	8.55	12.93	18.80	...	5.68	6.34	7.59	9.28	...
7.68	13.0	6.36	8.55	12.93	18.64	...	5.74	6.34	7.59	9.13	...
8.61	13.0	6.66	8.86	13.27	19.12	...	6.03	6.65	7.93	9.60	...
10.68	13.0	6.54	8.78	13.39	19.44	...	5.91	6.57	8.04	9.93	...
11.60	13.0	6.78	9.02	13.39	19.28	...	6.15	6.81	8.04	9.76	...
12.71	13.0	6.48	8.70	12.82	18.64	...	5.85	6.49	7.48	9.13	...

index in the *UVBR* region varied from 0.65 ± 0.15 to 0.76 ± 0.10 from January to March. Slight darkening and spectral steepening are barely significant, if we consider only the data in the optical region. It should be also noted that, in early February, a brightening to 10.4 ± 0.3 mJy in the V band with a spectral index of 0.72 ± 0.06 has been observed at Dodaira.

Optical polarimetry was also made simultaneously with the photometry by the 188 cm telescope at Okayama, and by the 91 cm telescope at Dodaira. The polarization of the nonthermal component showed no wavelength dependence in the *UBVR* region throughout the whole observing period, provided that the light from the underlying galaxy was unpolarized. The polarization properties of the nonthermal component through the U to K bands averaged over March 8–12 are given in Table 5. No wavelength dependence of the degree of polarization was found in the optical and near-infrared regions.

The variation of the degree of the polarization is shown in the first panel of Figure 1. The degree of polarization showed a considerable change during the March epoch; from February 29 to March 5 the degree of polarization was as high as 7%–8.5%, with a trend of the smooth weakening, and on March 6 it declined abruptly to 4%–5%, while the position angle remained constant at 170°–180°. This abrupt change of the polarization seems to be correlated with the flux decrease in the same period.

d) Infrared Observations

Infrared observations were made with the 3.8 m telescope at United Kingdom Infrared Telescope (UKIRT) at Mauna Kea, Hawaii, and the 1 m telescope at the Agematsu Infrared Observatory. At Agematsu we observed Mrk 421 on January 25 and 27 and March 7 and 8 with a rectangular aperture of 24″ in the

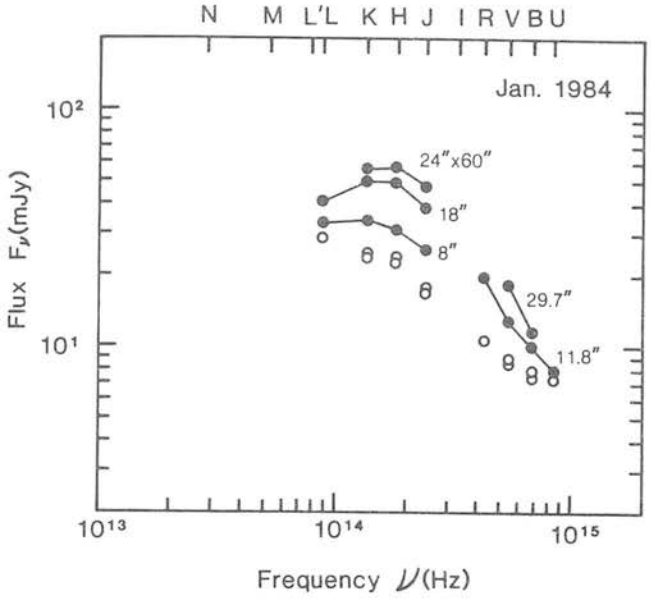


FIG. 3*a*

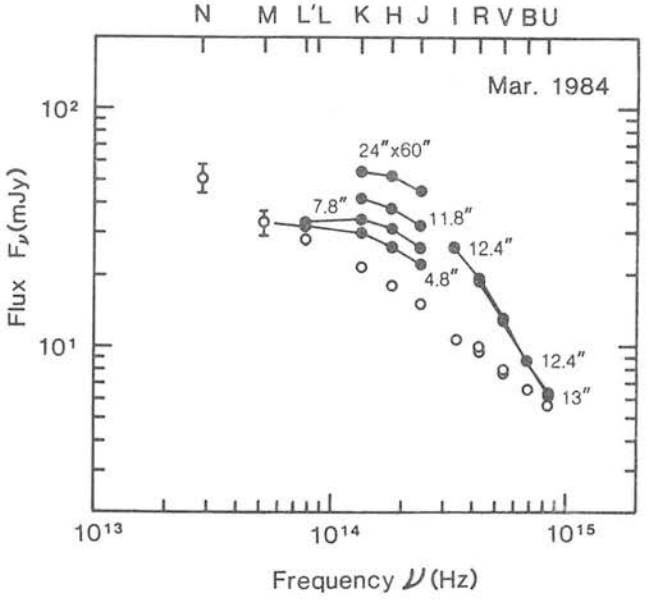


FIG. 3*b*

FIG. 3.—IR-optical spectra of Mrk 421 in 1984 January–February (*a*) and in March (*b*). *Filled circles with straight lines*: observed fluxes. Beam sizes are also shown numerically in the figure. *Open circles*: derived nonthermal fluxes.

right ascension by 60″ in the declination and 40″ chopping throw at the *J*, *H*, and *K* bands. The calibration was made by using ξ UMa as a standard star and SAO 062387 as a reference star. UKIRT observations were made on February 4 and 6 at the *J*, *H*, *K*, and *L* bands with a 18″ aperture using GL 299 as a reference star. On March 8–12, the *J*, *H*, *K*, and *L'* band observations were made with 4″.8, 7″.8, and 11″.8 apertures, on March 11–12 the *M* band observation was made with a 4″.8 aperture, and on March 8–10 the *N* band observation was made with a 6″ aperture. On January 30, photometry at *J*, *H*, *K*, and *L* bands was made with a 8″ aperture using the NASA infrared telescope, at Mauna Kea, Hawaii.

The observed data are tabulated in Table 6. The infrared flux decreased by 13% ± 4% from January to March, comparing the Agematsu observations at two epochs. As in the optical observations, the contribution of the underlying galaxy was subtracted to obtain the fluxes of the central nonthermal component for *J*, *H*, *K*, *L* (*L'*) bands. Since the observation with the largest beam size of 24″ × 60″ does not agree with the brightness of the model galaxy, we obtain contrariwise the brightness of the underlying galaxy for the 24″ × 60″ beam size by subtracting the nonthermal component derived from the observations with smaller beam sizes on March 8, and apply it to the photometry in the January epoch. The infrared colors of the underlying galaxy are obtained as $J-H=0.69$ and $H-K=0.40$ from the observations in March, which become almost consistent with those of typical elliptical galaxies (Frogel *et al.* 1975; Persson, Frogel, and Aaronson 1979; Glass 1984), taking account of the *K*-corrections (Frogel *et al.* 1978). We assumed $K-L=0.06$ (Aaronson 1978) and $V-K=3.40$ (Aaronson 1978; Persson, Frogel, and Aaronson 1979), consistent with those of typical elliptical galaxies including *K*-corrections. For the *M* and *N* bands the contribution of the galaxy is safely neglected. The resulting fluxes of the nonthermal component are summarized in the bottom panel of Table 6, and the flux at the *K* band is plotted in the third panel of Figure 1. The spectra obtained for respective epochs are shown in Figure 3, along with the optical spectra. The spectral index in the infrared band was 0.5 ± 0.2 and 0.57 ± 0.11 for January and March, respectively. The conversion of the magnitude to the absolute flux has been made using coefficients tabulated in Table 6, on the basis of the calibration by Blackwell *et al.* (1983).

The polarization measurements were also made on March 9, 11, and 12 at the *J*, *H*, and *K* bands by UKIRT. Those results are tabulated in Table 5, together with the optical polarization. The degree of polarization and the position angle of the nonthermal component in the infrared are 4.8% ± 0.4% and 180° ± 5°, which are fully consistent with those at optical bands.

TABLE 5

POLARIZATION OF NONTHERMAL COMPONENT OBSERVED IN 1984 MARCH 8–12

Parameter	*U*	*B*	*V*	*R*	*J*	*H*	*K*
Beam size	13″	13″	13″	13″	7″.8	7″.8	7″.8
Observed degree of polarization[a]	4.0 ± 0.3	3.3 ± 0.3	3.0 ± 0.2	2.7 ± 0.2	2.5 ± 0.4	2.9 ± 0.3	2.8 ± 0.6
Degree of polarization of nonthermal component[a]	4.3 ± 0.3	4.3 ± 0.4	4.7 ± 0.4	5.0 ± 0.4	4.4 ± 0.7	5.0 ± 0.4	4.5 ± 1.0
Position angle	177° ± 2°	175° ± 3°	174° ± 2°	175° ± 3°	179° ± 9°	180° ± 5°	184° ± 13°

[a] In percent.

TABLE 6

SUMMARY OF INFRARED PHOTOMETRIC OBSERVATIONS[a]

DATE (1984) (UT)	APERTURE	*J*	*H*	*K*	*L*	*L'*	*M*	*N*
	Central frequency:	240 THz	182 THz	136 THz	86.9 THz	78.5 THz	62.7 THz	28.6 THz
	Central wavelenth:	1.25 μm	1.65 μm	2.2 μm	3.45 μm	3.82 μm	4.78 μm	10.5 μm
	Absolute flux at zero magnitude:	1540 Jy	980 Jy	670 Jy	290 Jy	250 Jy	165 Jy	40 Jy
Jan 25/27	24″ × 60″	46.94	57.44	56.24	...	...	...	...
Jan 30	8	25.56	31.28	33.89	32.69	...	...	...
Feb 4/6	18	38.7	49.1	49.4	40.4	...	...	...
Mar 7/8	24 × 60	44.83	51.91	54.21	...	...	...	...
Mar 8	4.8	22.06	26.02	30.62	...	30.90	...	...
	7.8	26.76	31.86	35.49	...	31.76	...	...
	11.8	31.88	37.95	41.50	...	...	...	...
	6	...	...	...	...	...	...	58 ± 18
Mar 9	4.8	23.52	26.74	33.27	...	35.15	...	...
	7.8	26.27	31.28	34.52	...	32.65	...	...
	6	...	...	...	...	...	...	58 ± 12
Mar 10	7.8	26.03	30.71	32.97	...	29.8	...	...
	6	...	...	...	...	...	...	36.5 ± 7
Mar 11	4.8	22.26	26.26	29.52	...	30.9	38 ± 6	...
	7.8	26.03	31.57	34.20	...	30.3	...	...
Mar 12	4.8	21.65	26.26	25.95	...	30.0	29 ± 6	...
	7.8	26.27	31.86	34.52	...	35.5	...	...
Jan 25/27	...	17.7	23.9	24.2	...	...	...	...
Feb 4/6	...	16.6	22.5	23.4	28.5	...	...	...
Mar 8/12	...	15.2	18.1	21.6	...	27.7	33.2	50.7

[a] Lowermost three rows show the fluxes of the nonthermal component in units of mJy. Errors are less than 1% unless otherwise indicated.

e) Radio Observations

Radio observations were made with the 45 m telescope at the Nobeyama Radio Observatory and the 26 m telescope at the University of Michigan. At Nobeyama, flux measurements at 10, 22, and 43 GHz were made, while at Michigan flux was measured at 4.8, 8.0, and 14.5 GHz. At Nobeyama the calibration was done by using 3C 286 as the primary calibrator of which the fluxes are 4.7, 2.6, and 1.7 Jy at 10, 22, and 43 GHz, respectively, and 3C 274, 4C 39.25, and 1144+40 were observed as references. No significant daily variations above 10% level were found, so that only the averaged data over each epoch are used. The results are listed in Table 7, from which one sees that radio fluxes are rather stable and are 570, 510, and 520 mJy at 10, 22, and 43 GHz, respectively, throughout the January and March observations.

TABLE 7

A. NRO 45 METER RADIO OBSERVATIONS[a]

Epoch (1984)	10 GHz	22 GHz	43 GHz
Jan	0.59 ± 0.03	0.47[b]	0.49 ± 0.03
Mar	0.55 ± 0.03	0.55 ± 0.05	0.55 ± 0.1

B. UNIVERSITY OF MICHIGAN 26 METER RADIO OBSERVATIONS

Epoch (1984)	4.8 GHz	8.0 GHz	14.5 GHz
Jan	0.657 ± 0.034	0.604 ± 0.013	0.540 ± 0.017
Mar	0.659 ± 0.036	0.62 ± 0.03	0.571 ± 0.020

[a] Flux density is shown in units of janskys, and errors are the standard deviation from the mean daily flux.

[b] Data are available only for 1 day, and the error cannot be estimated.

The data obtained at Michigan are consistent with those at Nobeyama, viz., 660, 610, and 550 mJy at 4.8, 8.0, and 14.5 GHz, respectively. In the second panel of Figure 1, the data at 14.5 GHz are plotted.

f) Composite Spectra

Figure 4 shows the composite energy spectra emerging from the central nonthermal component of Mrk 421 in 1984 January and March. These composite spectra are similar to those of other BL Lac objects, i.e., a power-law spectrum with a gradually increasing spectral index toward high frequencies. In the January epoch, the spectral indices are ~0.1, ~0.5, ~0.7, ~1.0, and ~1.1 for radio, IR, optical, UV, and X-ray wavelengths, respectively, while in March these are ~0.1, ~0.6, ~0.8, ~1.1, and ~1.8.

III. COMPARISON WITH PREVIOUS OBSERVATIONS

At radio wavelengths, VLBI observations (Baath *et al.* 1981; Baath 1984) revealed an unresolved core of 240 ± 40 mJy with a size less than 0.3 mas, as well as a resolved component of 270 ± 50 mJy and 3.4 mas size at 5 GHz. Other VLBI observations (Weiler and Johnston 1980; Wehrle *et al.* 1984) also revealed a compact component of ~1 mas size and ~300 mJy strength both at 5 and 2.3 GHz. There also exists a ~0.1 Jy extended component of a few arcminute size (Kapahi 1979; Ulvestad, Johnson, and Weiler 1983). Our flux measurements at radio frequencies included all these components. But the total flux of Mrk 421 was rather stable at 1 centimeter wavelength, and our value is consistent with previously reported values (Margon, Jones, and Wardle 1978; O'Dell *et al.* 1978*a*). Therefore the flux of the core component is probably ~300 mJy.

O'Dell *et al.* (1978*a*) showed that the radio spectrum extends to at least 90 GHz with a flat shape. The extrapolation of the

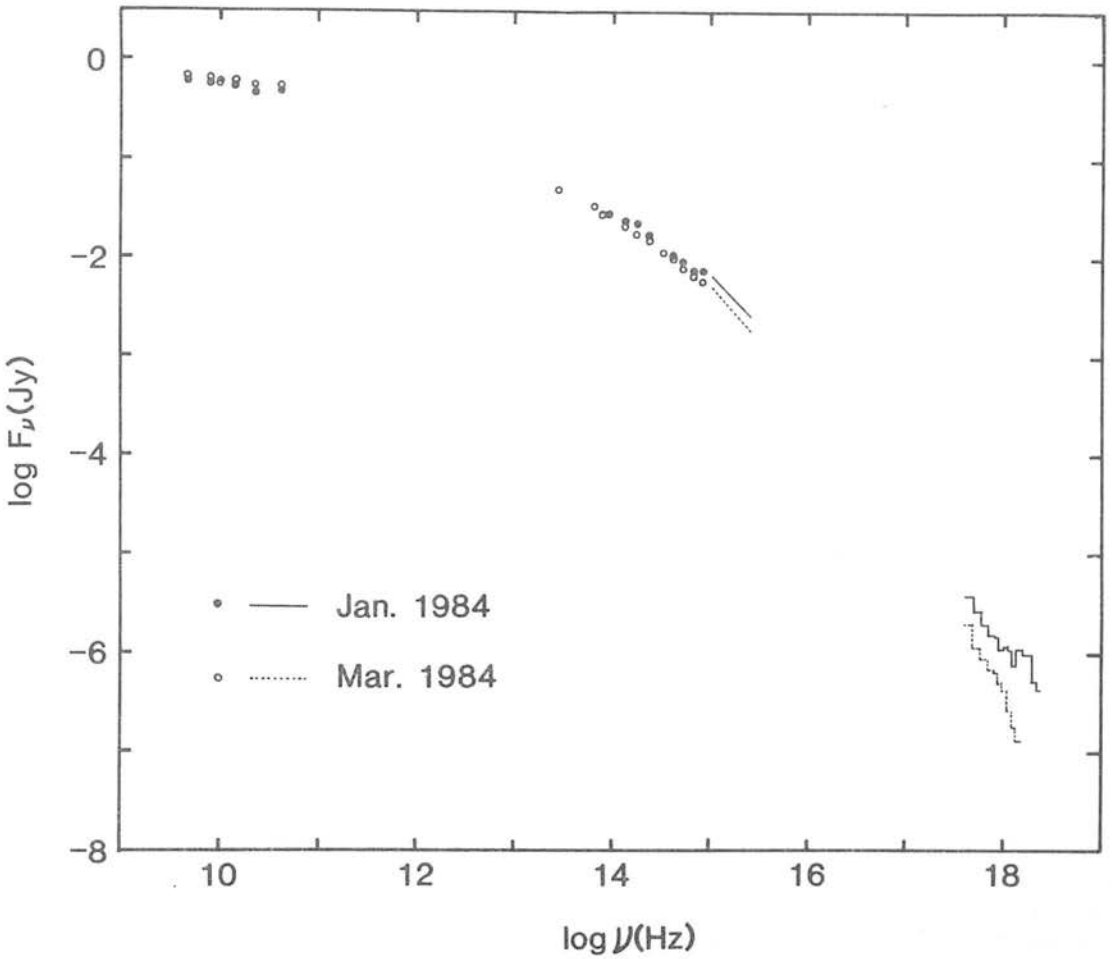


FIG. 4.—Composite spectra of the nonthermal component of Mrk 421 in 1984 January and March

flat radio spectrum of 300 mJy intersects the extrapolation of the IR-optical spectrum at $\sim 10^{12}$ Hz.

The optical fluxes were faint compared to the previous observations (Hagen-Thorn *et al.* 1983). Also the infrared fluxes of the nonthermal component were in a faint phase among previous observations at the *K* band (Ulrich *et al.* 1975; Allen 1976; Joyce and Simon 1976; O'Dell *et al.* 1978*b*; Bailey *et al.* 1981; Impey 1983). Thus we conclude that Mrk 421 was in a dark phase on two occasions of our coordinated observations.

The position angle of the optical polarization in the present observation agrees with the dominant direction in the long-term behavior (Hagen-Thorn *et al.* 1983). The polarization of Mrk 421 observed by Bailey *et al.* (1981) in 1980 December is nearly equal to that in the present observation in the position angle and the degree of polarization, although Mrk 421 was in a brighter phase then. However, Bailey *et al.* (1981) claimed that the polarization degree decreases as the wavelength increases, while our degree of polarization shows no wavelength dependence. We think it is not due to observational errors but due to the difference in the models of the underlying galaxy. Their model galaxy has the brightness by 0.3 mag smaller than that we used in the *V* band, which brings a decreasing degree of polarization as the wavelength increases through the dilution effect. Therefore, a decreasing polarization of Mrk 421 with wavelength is questionable.

It is to be noted that there exists a spectral break between IR-optical and UV bands at $\sim 10^{15}$ Hz. This spectral break or steepening seems to be common to BL Lac objects.

The UV light curve for the period 1978–1981 was analyzed by Ulrich *et al.* (1984), and they have shown that the flux varied on the time scale of a few years. In 1979 November the UV flux was in a minimum state with the spectral index of 1.0. Both the flux and the spectral index are similar to those of the present observation in 1984. Since the flux steadily increased and the spectrum hardened from 1979 to 1981, Mrk 421 might have turned to a decreasing and softening phase in 1982–1983.

The X-ray flux of this observation is significantly lower than those of previous observations by *SAS 3* (Hearn, Marshall, and Jernigen 1979) and *HEAO 1* (Mushotzky *et al.* 1979) by an order of magnitude. At those epochs X-ray fluxes were higher than or on the extrapolation from the UV spectrum. The extrapolation of the UV spectrum toward shorter wavelengths should be compared to the X-ray spectra obtained by *Tenma*. The extrapolation from the UV spectrum lies at 5.47 and 2.26 μJy at 10^{18} Hz at January and March epochs, respectively, while the observed fluxes were 1.1 and 0.51 μJy, respectively. Then the observed X-ray fluxes are lower than the extrapolation from the UV spectrum by an order of magnitude in the present observation. The spectral slope between UV and X-ray is 1.3–1.4 and is apparently steeper than the UV slope of ~ 1. On the other hand, the X-ray spectral index changes from ~ 1

at the January epoch to ~ 2 at the March epoch. Below 4 keV the X-ray spectral index was ~ 1.5 for both periods.

This result implies that X-ray emission is represented by two components. These variations of flux and spectral shape at 5 week interval are very important feature. *HEAO 1* observations in 1978 May showed a remarkable spectral change from *OSO 8* observations in 1977 May (Mushotzky *et al.* 1979). It has been suggested that the X-ray spectra of BL Lac objects have two components, a soft and a hard component (Riegler, Agrawal, and Mushotzky 1979; Worrall *et al.* 1981). We may infer that the hard component varies more violently than the soft component.

IV. IMPLICATIONS ON THE MODEL

a) Source Parameters

The continuum emission from BL Lac objects may be due to synchrotron radiation, inverse Compton scattering, or both (Blandford 1984; Konigl 1981). In this subsection we derive physical parameters of the source using the data obtained in the present observations within the context of the homogeneous and spherical symmetric synchrotron self-Compton (SSC) model (Gould 1979). The SSC model assumes the presence of relativistic electrons with power-law distribution, $N(\gamma) \propto \gamma^{-p}$ (γ is the Lorentz factor, and p is the power-law index), which emit synchrotron radiation having the power-law spectrum $S_{\rm syn} \propto \nu^{(1-p)/2}$. Since the synchrotron absorption coefficient is proportional to $\nu^{-(4+p)/2}$, the self-absorption turnover is expected at a certain frequency ν_a (ν_* GHz) with the corresponding flux density S_* Jy.

Common to those BL Lac objects, a break or bending of the spectrum appears at the break frequency $\nu_b \approx 10^{15}$ Hz. The acceleration rate of relativistic electrons responsible for synchrotron radiation above ν_b may be smaller than the energy loss rate, making the energy spectrum steeper. Here the Lorentz factors of electrons which emit the synchrotron radiation at the turnover and break frequencies are denoted by γ_a and γ_b, respectively. Soft X-rays in the January epoch and the whole X-ray in the March epoch may be produced by the synchrotron radiation by electrons in the high-energy tail. The relatively flat, hard X-ray component seen in the January spectrum (see Fig. 2) may be attributed to the self-Compton photons. This self-Compton flux S_C has a power-law spectrum with the same spectral index as the synchrotron spectrum with

$$S_C \propto \nu_X^{-(p-1)/2}\theta_*^{-2(p+2)}\nu_*^{-(3p+7)/2}S_*^{(p+3)} \times [\delta/(1+z)]^{-(p+3)} \ln(\nu_b/\nu_a)\,, \quad (1)$$

where ν_X is the frequency of a scattered photon, δ is the kinematical Doppler factor in the relativistic beaming model, z is the cosmological redshift, and the apparent angular diameter of the source is θ_* mas.

Energy densities of magnetic fields $u_{\rm mag}$ and synchrotron photons $u_{\rm syn}$, and the lower limit to the energy density of relativistic electrons $u_{\rm rel}^{\rm min}$, and γ_a are given as

$$u_{\rm mag} \propto \theta_*^{8}\nu_*^{10}S_*^{-4}[\delta/(1+z)]^2\,, \quad (2)$$

$$u_{\rm syn} \propto \theta_*^{-2}\nu_*\, S_*[\delta/(1+z)]^{-4}\,, \quad (3)$$

$$u_{\rm rel}^{\rm min} \propto \theta_*^{-9}\nu_*^{-7}S_*^{4}[\delta/(1+z)]^{-5}(1+z)^2 d_{L^*}^{-1}\,, \quad (4)$$

and

$$\gamma_a \propto \theta_*^{-2}\nu_*^{-2}S_*[\delta/(1+z)]^{-1}\,, \quad (5)$$

respectively (Marscher *et al.* 1979; Marscher 1983). Here d_{L^*} G pc is the luminosity distance. The proportional coefficients in equations (1)–(5) are complicated functions of p. It is to be noted that resultant numerical values depend critically on the choices of θ_*, S_*, and ν_*, but the essential features described below are the same for plausible choices.

Since the X-ray flux by the self-Compton photons should be lower than the observed X-ray flux, we obtain the lower limit to the beaming factor $\delta_{\rm min}$ from equation (1) as

$$\delta_{\rm min} \propto (1+z)S_*\left[\frac{\ln(\nu_b/\nu_a)}{\nu_X^{(p-1)/2}\theta_*^{2(p+2)}\nu_*^{(3p+7)/2}S_*}\right]^{1/(p+3)}. \quad (6)$$

The synchrotron self-absorption turnover was not directly observed, but presumably ν_* lies at $\sim 10^3$, the intersection of straight extrapolation from radio and that from IR spectra. Assuming that the flux at the turnover is not much different from that of the smallest unresolved radio component (0.3 Jy with an angular size less than 0.3 mas), we adopt $S_* = 0.3$. The turnover frequency ν_* is determined so that the synchrotron flux gives the observed UV flux at 10^{15} Hz (6.5 mJy for the January epoch and 5.1 mJy for the March epoch), with the spectral index $\alpha = 0.625$ ($p = 2.25$) the well-fitted value to the IR-optical data. The source parameters are calculated as a function of θ_*, which are shown in Figure 5 for the January data. We notice that there are two critical values of θ_*, referred

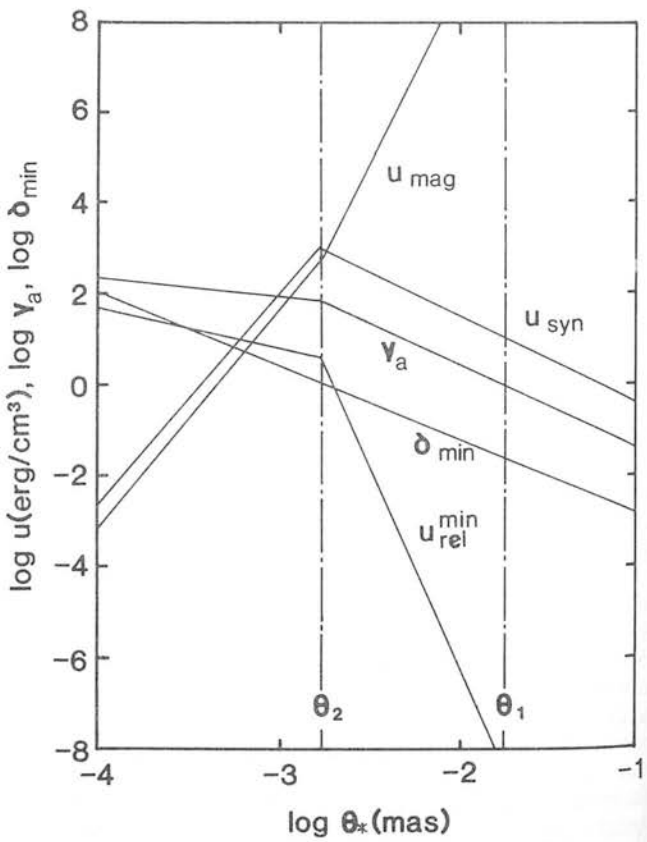


FIG. 5.—Source parameters obtained by the homogeneous synchrotron theory using the January data. The energy density of magnetic fields $u_{\rm mag}$, that of synchrotron photons $u_{\rm syn}$, and the lower limit to the energy density of relativistic electrons $u_{\rm rel}^{\rm min}$ are shown as a function of the angular diameter of the source θ_* mas. The minimum beaming factor $\delta_{\rm min}$ and the Lorentz factor of electrons γ_a which emit the synchrotron radiation at the self-absorption turnover frequency are also depicted.

TABLE 8

SOURCE CHARACTERISTICS DEDUCED FROM SYNCHROTRON SELF-COMPTON MODEL

PARAMETER	JANUARY OBSERVATION	MARCH OBSERVATIONS		
		1	2	3
ν_* (GHz)	2.2×10^3	1.5×10^3	1.5×10^3	1.5×10^3
θ_1 (mas)	1.6×10^{-2}	2.3×10^{-2}	2.3×10^{-2}	2.3×10^{-2}
θ_2 (mas)	1.8×10^{-3}	2.7×10^{-3}	2.7×10^{-3}	3.1×10^{-3}
δ	1	1	1.2	1.7
θ_* (mas)	1.8×10^{-3}	2.7×10^{-3}	2.4×10^{-3}	2.2×10^{-3}
t_{1c} (s)	7.6×10^4	1.2×10^5	1.0×10^5	9.7×10^4
γ_a	80	76	80	63
B (G)	1.3×10^2	9.5×10^1	7.0×10^1	7.9×10^1
$u_{\rm mag}$ (ergs cm^{-3})	6.4×10^2	3.6×10^3	2.0×10^2	2.5×10^2
$u_{\rm syn}$ (ergs cm^{-3})	1.2×10^3	4.0×10^2	2.1×10^2	5.7×10^1
$u_{\rm rel}$ (ergs cm^{-3})	3.1	1.2	1.4	3.9×10^{-1}
t_a (s)	3.3×10^2	1.1×10^3	1.8×10^3	8.5×10^3
t_b (s)	1.6×10^1	3.9×10^1	8.4×10^1	3.9×10^2

to as θ_1 and θ_2. For $\theta_* > \theta_1$, γ_a is less than unity. Since the synchrotron flux with a power-law spectrum is valid for $\gamma \gg 1$, thus θ_1 provides an optimistic upper limit to the angular diameter. While for $\theta_* < \theta_2$, the expected inverse Compton flux at ν_X deduced from the observed synchrotron flux exceeds the observed flux if $\delta = 1$. Conversely, such relativistic effects are required if θ_* is found virtually smaller than θ_2. In the second and third columns of Table 8, the results for the January and March data are tabulated for $\delta = 1$ and $\theta_* = \theta_2$, taking $\nu_X = 10^{18}$ Hz.

Ulrich *et al.* (1975) reported the absorption redshift of the underlying galaxy, $z = 0.03$, implying the distance of $90h^{-1}$ Mpc; 1 mas corresponds to $0.44h^{-1}$ pc, and the light crossing time of the source is $t_{1c} = 4.5 \times 10^7 \theta_* h^{-1}$ s, where h is the Hubble parameter in units of 100 km s^{-1} Mpc^{-1}. In Table 8 the cooling times, t_a and t_b, of electrons with Lorentz factors of γ_a and γ_b, are given, respectively, as well as t_{1c}. The cooling time of electrons responsible for the IR-optical radiation is shorter than the light crossing time, implying *in situ* acceleration of the relativistic electrons. The energy density of magnetic fields, $u_{\rm mag}$, dominates over that of relativistic electrons, if $u_{\rm rel}^{\rm min}$ is not so different from the true energy density of relativistic electrons, since γ_a is relatively small. This property is consistent with the observed behavior of polarization in the IR-optical bands that the position angle remained constant although the degree of polarization and flux varied with time.

Since VLBI experiments have shown ~ 1 mas structure at 5 GHz, radio structure should be more extended than IR-optical-UV component. This is consistent with the model of flat spectrum radio sources which can be expressed in terms of inhomogeneous synchrotron source (Marscher 1977; Konigl 1981). In this model the apparent angular size is inversely proportional to the frequency in the flat portion of the spectrum, and the size at 10^{12} Hz becomes $\sim 5 \times 10^{-3}$ mas, consistent with above analysis.

Mrk 421 was rather faint in the epochs of the present observation, and the total luminosity was $\sim 10^{44}$ ergs s^{-1}. At the most luminous epoch in the past it was ~ 10 times brighter. If we adopt the conclusion that the ultimate energy release occurs as a result of accretion onto the supermassive black hole and 10^{45} ergs s^{-1} corresponds to the Eddington luminosity, the mass of the black hole is $10^7 M_\odot$ and the accretion energy is reprocessed in the size of 10^{13} cm, which amounts to $\theta_* \approx 10^{-5}$. Thus the size of the synchrotron source is $\sim 10^2$ times larger than the size of the ultimate energy release region.

b) *Spectral Variation*

Our January and March observations of Mrk 421 suggest that the self-Compton flux corresponding to the relatively flat, hard X-ray tail seen in the January observations decreased by more than factor of ~ 2 during 5 weeks, while the synchrotron flux decreased by 15%–25% in the IR, optical, and UV bands in the same interval. The radio flux remained essentially constant during those two observational periods. There seem to be several ways to explain the change of the spectra obtained in those two epochs with the homogeneous and spherical symmetric SSC model, as will be discussed below.

The simplest model is that the change was due to a slight steepening of the energy spectrum of electrons, while θ_*, S_*, and δ remained constant. Quantitatively, 22% decrease of synchrotron flux at 10^{15} Hz corresponds to the change of p from 2.25 to 2.33. This leads to decreases of $u_{\rm rel}^{\rm min}$ and $u_{\rm syn}$ by $\sim 10\%$ and $\sim 13\%$, respectively, while $u_{\rm mag}$ is increased by $\sim 3\%$ and ν_* remained essentially constant. The decrease of $S_{\rm C}$ depends on θ_* and is $\sim 20\%$ for $\theta_* \approx 3 \times 10^{-3}$, which is insufficient to explain the observed spectral change.

We also examined the cases where the spectral shape of electrons remained unchanged with $p = 2.25$. It is convenient to express S_* and ν_* in terms of θ_*, δ, $u_{\rm mag}$, and $u_{\rm rel}^{\rm min}$ from equations (2) and (4). Because the synchrotron flux $S_{\rm syn}$ is proportional to $S_* \nu_*^{(p-1)/2}$, then we obtain

$$S_{\rm syn} \propto (u_{\rm rel}^{\rm min})^{(p+4)/6} u_{\rm mag}^{(2p+5)/12} \theta_*^{(p+16/6} \delta^{(p+5)/2} . \quad (7)$$

On the other hand, the inverse Compton flux $S_{\rm C}$ is given from equation (1) as

$$S_{\rm C} \propto (u_{\rm rel}^{\rm min})^{(p+4)/3} u_{\rm mag}^{(p+7)/12} \theta_*^{(p+10)/3} \delta^{(p+5)/2} . \quad (8)$$

Let us first assume that δ and θ_* remained constant during the two epochs. The decreases of $S_{\rm syn}$ and $S_{\rm C}$ are due to those of $u_{\rm rel}^{\rm min}$ and $u_{\rm mag}$. A decrease of $u_{\rm mag}$, however, would lead to a larger decrease of $S_{\rm syn}$ than that of $S_{\rm C}$, contrary to the observed data. While a decrease of $u_{\rm rel}^{\rm min}$ would lead to a larger decrease of $S_{\rm C}$ than that of $S_{\rm syn}$. Quantitatively, $\sim 21\%$ decrease of $u_{\rm rel}^{\rm min}$ would decrease $S_{\rm syn}$ and $S_{\rm C}$ by $\sim 22\%$ and $\sim 40\%$, respectively. In this case S_* and ν_* would also decrease by $\sim 18\%$ and

$\sim 8\%$, respectively. The decrease of S_C is slightly smaller than the observed change.

We next consider the case of adiabatic expansion with $u_{\rm mag} \propto \theta_*^{-4}$ and $u_{\rm rel}^{\rm min} \propto \theta_*^{-4}$. We obtain $S_{\rm syn} \propto \theta_*^{-(7p+10)/6}$ and $S_C \propto \theta_*^{-(4p+13)/3}$. Then an expansion by $\sim 6\%$ would decrease $S_{\rm syn}$ and S_C by 22% and $\sim 35\%$, respectively. In this case S_* and ν_* would also decrease by $\sim 21\%$ and $\sim 13\%$, respectively. The decrease of S_C is also smaller than the observed change. A common feature to the above models is the decrease of the energy density of relativistic electrons.

Of course, there may exist more complicated changes involving the beaming factor δ. As an example, we assume that S_* remained constant because radio flux remained unchanged, which requires a 33% decrease of ν_* with constant $p = 2.25$. For $u_{\rm mag} \propto \theta_*^{-4}$, we obtain $S_C \propto \nu_*^{(p-1)/6}\delta^{-(2p+7)/3}$. In order to decrease the self-Compton flux by a factor of 2.2, the kinematical Doppler factor is required to increase by 20% within 5 weeks, and the source would expand during the above time interval by 34%. Accompanying this $u_{\rm mag}$ and $u_{\rm rel}^{\rm min}$ would decrease by a factor of 3.2 and 2.3, respectively. Thus the increase in the relativistic motion of the source might account for the observed property of the spectral variation separated by 5 weeks. This case is rather artificial since the increase of relativistic beaming compensates for a large decrease of intrinsic fluxes. In the third column of Table 8, the source parameters in the March epoch are shown assuming that the source in January was in a state of $\theta_* = \theta_2$ and $\delta = 1$. Since the hard X-ray spectrum was very steep in March, we obtain more severe constraint if we apply the same argument at higher ν_X. In the fourth column of Table 8 are tabulated the numerical results when we take $\nu_X = 1.4 \times 10^{18}$ Hz.

Virtually, the spectral variation in our two observational epochs may be explained by one or some combination of above alternative possibilities. It is also to be noted that the source should have some inhomogeneity. Even within the homogeneous source model discussed here, the radio emission arises in a more extended region than the central region where IR to UV emission emanates. In various inhomogeneous jet models, emissions at different frequencies are ascribed to different regions, implying different time scales of variations for different wave bands (Konigl 1981; Marscher 1980; Reynolds 1982; Ghisellini, Maraschi, and Treves 1985). However, our observations lack such detailed frequency dependence of time scales of flux variations, which does not enable us to set some constraints to inhomogeneous jet models. Only information we have obtained is the frequency independence of the degree of polarization at the March epoch, which is rather favorable to the homogeneous model. Although we have not applied the inhomogeneous jet models to determine source parameters and to explain the spectral variation, the analyses of our observations with such models are also awaited.

V. SUMMARY

The results of the present observations are summarized as follows:

1. Mrk 421 was in a relatively faint state in 1984 January and March at frequencies from IR through X-rays.

2. From January to March, Mrk 421 decreased its flux in the IR to X-ray bands. X-ray emission decreased by a factor of 2–3, while IR-UV radiation decreased by 15%–25%. Radio emission remained essentially stable.

3. The change in the X-ray spectral shape suggests that X-ray emission has two components: the soft component is probably the synchrotron radiation, and the hard component was more violently variable and may be ascribed to other mechanisms such as the inverse Compton scattering of synchrotron photons.

4. The break was observed at a frequency of $\sim 10^{15}$ Hz; the power-law index of IR-optical part is 0.5–0.8, while for the UV range it is ~ 1.0.

5. At the optical band, the degree of polarization showed a factor of 2 variation on the time scales of a few days, while the position angle remained constant during the present observations. The degree of polarization in the IR-optical region shows no wavelength dependence.

6. The radio spectrum is flat, and the extrapolation from the radio spectrum intersects the extrapolation from IR at a frequency of a few times 10^{12} Hz, below which the synchrotron self-absorption plays an important role.

7. Applying the homogeneous synchrotron source theory to the observed results, physical source parameters are deduced. Source size should be less than $\sim 10^{-2}$ mas. If the source size is less than a few times 10^{-3} mas, relativistic beaming is required. The source is magnetic field dominated, and the relativistic electrons with relatively low energy are responsible for the synchrotron radiation. The cooling time of electrons is shorter than the light crossing time, implying that there must be *in situ* acceleration of the electrons.

8. The spectral variations in the 5 week interval may be interpreted as a result of a slight steepening of the electron spectrum, the decrease of the energy density of relativistic electrons, an expansion of the source with or without an increase of beaming factor, or some combinations of these alternatives, although it is hard to discriminate those possibilities.

Some of the optical observations at Okayama were kindly made for this program by K. Misawa, Y. Fujime, and M. Kondo. The IR data at January 30 and February 4 and 6 were kindly provided by E. I. Robson. We appreciate their cooperation. The work at the University of Michigan was supported by the National Science Foundation under grant AST 83-01234. The authors would like to appreciate the referee for suggestions regarding emission models.

REFERENCES

Aaronson, M. 1978, *Ap. J. (Letters)*, **221**, L103.
Allen, D. A. 1976, *Ap. J.*, **207**, 367.
Baath, L. B. 1984, in *IAU Symposium 110, VLBI and Compact Radio Sources*, ed. R. Fanti, K. Kellerman, and G. Setti (Dordrecht: Reidel), p. 127.
Baath, L. B., Elgered, G., Lundqvist, G., Graham, D., Weiler, K. W., Seielstad, G. A., Tallqvist, S., and Schilizzi, R. T. 1981, *Astr. Ap.*, **96**, 316.
Bailey, J., Cunningham, E. C., Hough, J. H., and Axon, D. J. 1981, *M.N.R.A.S.*, **197**, 627.
Blackwell, D. E., Leggett, S. K., Petford, A. D., Mountain, C. M., and Selby, M. J. 1983, *M.N.R.A.S.*, **205**, 897.
Blandford, R. D. 1984, in *11th Texas Symposium on Relativistic Astroophysics*, ed. D. S. Evans (*Ann. NY Acad. Sci.*), **422**, 303.
Bregman, J. N., *et al.* 1982, *Ap. J.*, **253**, 19.
———. 1984, *Ap. J.*, **276**, 454.
Frogel, J. A., Persson, S. E., Aaronson, M., Becklin, E. E., Matthews, K., and Neugebauer, G. 1975, *Ap. J. (Letters)*, **200**, L123.
Frogel, J. A., Persson, S. E., Aaronson, M., and Matthews, K. 1978, *Ap. J.*, **220**, 75.
Ghisellini, G., Maraschi, L., and Treves, A. 1985, *Astr. Ap.*, **146**, 204.
Glass, I. S. 1984, *M.N.R.A.S.*, **211**, 461.
Glassgold, A. E., *et al.* 1983, *Ap. J.*, **274**, 101.
Gould, R. J. 1979, *Astr. Ap.*, **76**, 306.
Hagen-Thorn, V. A., Marchenko, S. G., Smehacheva, R. I., and Yakovleva, V. A. 1983, *Astrophyzika*, **19**, 199.

Hearn, D. R., Marshall, F. J., and Jernigen, J. R. 1979, *Ap. J.* (*Letters*), **227**, L63.
Hickson, P., Fahlman, G. G., Auman, J. R., Walker, G. A. H., Menon, T. K., and Ninkov, Z. 1982, *Ap. J.*, **258**, 53.
Hutter, D. J., and Mufson, S. L. 1986, *Ap. J.*, **301**, 50.
Impey, C. D. 1983, *M.N.R.A.S.*, **202**, 397.
Johnson, H. L. 1966, *Ap. J.*, **143**, 87.
Joyce, R. R., and Simon, M. 1976, *Pub. A.S.P.*, **88**, 870.
Kapahi, V. K. 1979, *Astr. Ap.*, **74**, L11.
Kikuchi, S., and Mikami, Y. 1986, *Pub. Astr. Soc. Japan*, in press.
Kinman, T. D. 1978, in *Pittsburgh Conference on BL Lac Objects*, ed. A. M. Wolfe (Pittsburgh: University of Pittsburgh), p. 82.
Kondo, Y., *et al.* 1981, *Ap. J.*, **243**, 690.
Konigl, A. 1981, *Ap. J.*, **243**, 700.
Koyama, K., *et al.* 1984, *Pub. Astr. Soc. Japan.*, **36**, 659.
Marscher, A. P. 1977, *Ap. J.*, **216**, 244.
———. 1980, *Ap. J.*, **235**, 386.
———. 1983, *Ap. J.*, **264**, 296.
Marscher, A. P., Marshall, F. E., Mushotzky, R. F., Dent, W. A., Balonik, T. J., and Hartman, M. F. 1979, *Ap. J.*, **233**, 498.
Margon, B., Jones, T. W., and Wardle, J. F. C. 1978, *A.J.*, **83**, 1021.
Maza, J., Martin, P. G., and Angel, J. R. P. 1978, *Ap. J.*, **224**, 368.
Mufson, S. L., *et al.* 1980, *Ap. J.*, **241**, 74.
———. 1984, *Ap. J.*, **285**, 571.
Mushotzky, R. F., Boldt, E. A., Holt, S. S., and Serlemitsos, P. J. 1979, *Ap. J.* (*Letters*), **232**, L17.
O'Dell, S. L., Puschell, J. J., Stein, W. A., Owen, F., Porcas, R. W., Mufson, S., Moffet, T. J., and Ulrich, M. H. 1978*a*, *Ap. J.*, **224**, 22.
O'Dell, S. L., Puschell, J. J., Stein, W. A., and Warner, J. W. 1978*b*, *Ap. J. Suppl.*, **38**, 267.
Persson, S. E., Frogel, J. A., and Aaronson, M. 1979, *Ap. J. Suppl.*, **39**, 61.
Reynolds, S. P. 1982, *Ap. J.*, **256**, 13.
Riegler, G. R., Agrawal, P. C., and Mushotzky, R. F. 1979, *Ap. J.* (*Letters*), **233**, L47.
Tanaka, Y., *et al.* 1984, *Pub. Astr. Soc. Japan*, **36**, 641.
Ulrich, M. H., Hackney, K. R. H., Hackney, R. H., and Kondo, Y. 1984, *Ap. J.*, **276**, 466.
Ulrich, M. H., Kinman, T. D., Lynds, C. R., Rieke, G. H., and Eckers, R. D. 1975, *Ap. J.*, **198**, 261.
Ulvestad, J. S., Johnston, K. J., and Weiler, K. W. 1983, *Ap. J.*, **266**, 18.
Urry, C. M., Mushotzky, R. F., and Holt, S. S. 1986, *Ap. J.*, **305**, 369.
Warwick, R. S., McHardy, I. M., and Pounds, K. A. 1985, *Space Sci. Rev.*, **40**, 597.
Wehrle, A. E., Preston, R. A., Meier, D. L., Gorenstein, M. V., Shapiro, I. I., Rogers, A. E. E., and Rius, A. 1984, *Ap. J.*, **284**, 18.
Weiler, K. W., and Johnston, K. J. 1980, *M.N.R.A.S.*, **190**, 269.
Wills, B. J., *et al.* 1983, *Ap. J.*, **274**, 62.
Worrall, D. M., Boldt, E. A., Holt, S. S. Mushotzky, R. F., and Serlemitsos, P. J. 1981, *Ap. J.*, **243**, 53.
Worrall, D. M., *et al.* 1982, *Ap. J.*, **261**, 403.
———. 1984*a*, *Ap. J.*, **278**, 521.
———. 1984*b*, *Ap. J.*, **284**, 512.
Worrall, D. M., Puschell, J. J., Rodrigues-Espinosa, J. M., Bruhweiler, F. C., Miller, H. R., Aller, M. F., and Aller, H. D. 1984*c*, *Ap. J.*, **286**, 711.

H. D. ALLER and M. F. ALLER: Department of Astronomy, Dennison Building, University of Michigan, Ann Arbor, MI 48109-1090

P. BRAND and J. BURNELL: Astronomy Department, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, Scotland, UK

K. R. HACKNEY and R. L. HACKNEY: Department of Physics and Astronomy, Western Kentucky University, TCCW119, Bowling Green, KY 42101

S. HAYAKAWA: Department of Astrophysics, Nagoya University, Furo-cho Chikusa-ku, Nagoya 464, Japan

N. HIROMOTO: Radio Research Laboratory, 2-1, Nukuikitamachi 4-chome, Koganei, Tokyo 184, Japan

R. HOSHI: Department of Physics, Rikkyo University, 34-1, Nishi-Ikebukuro 3-chome, Toshima-ku, Tokyo 171, Japan

H. INOUE, K. KOYAMA, F. MAKINO, K. MAKISHIMA, M. MATSUOKA, and Y. TANAKA: Institute of Space and Astronautical Science, 6-1, Komaba 4-chome, Meguro-ku, Tokyo 153, Japan

M. INOUE and F. TAKAHARA: Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo, Nobeyama, Minamisaku, Nagano 384-13, Japan

T. KATO and H. TABARA: Faculty of Education, Utsunomiya University, 350, Minemachi, Utsunomiya, Tochigi 321, Japan

S. KIKUCHI and Y. MIKAMI: Tokyo Astronomical Observatory, University of Tokyo, 21-1, Osawa 2-chome, Mitaka, Tokyo 181, Japan

Y. KONDO: Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Code 685, Greenbelt, MD 20771

S. L. MUFSON: Department of Astronomy, Indiana University, 319 Swain West, Bloomington, IN 47405

M. NISHIDA: Department of Physics, Kyoto University, Kitashirakawa-Oiwakecho, Sakyo-ku, Kyoto 606, Japan

M. G. SMITH and P. M. WILLIAMS: Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, Scotland, UK

M. TSUBOI: Department of Astronomy, University of Tokyo, 11-16, Yayoi, 2-chome, Bunkyo-ku, Tokyo 113, Japan

C. M. URRY: Center for Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139

W. Z. WISNIEWSKI: Lunar and Planetary Laboratory, University of Arizona, Tucson, AZ 85721

THE ASTROPHYSICAL JOURNAL, 352:574–586, 1990 April 1

12

MULTIFREQUENCY OBSERVATIONS OF BL LACERTAE[1]

JOEL N. BREGMAN,[2] A. E. GLASSGOLD, AND P. J. HUGGINS
New York University; and the National Radio Astronomy Observatory

G. NEUGEBAUER, B. T. SOIFER, K. MATTHEWS, AND J. H. ELIAS
Palomar Observatory, California Institute of Technology

J. R. WEBB, J. T. POLLOCK, R. J. LEACOCK, AND A. G. SMITH
Rosemary Hill Observatory, University of Florida

H. D. ALLER, M. F. ALLER, AND P. A. HUGHES
University of Michigan Radio Astronomy Observatory

D. MACCAGNI, B. GARILLI, AND P. GIOMMI
CNR and ESTEC

J. S. MILLER AND S. STEPHENS
Lick Observatory

T. J. BALONEK, W. A. DENT, AND W. KINSEL
University of Massachusetts and Colgate University

W. Z. WISNIEWSKI
Steward Observatory, University of Arizona

P. M. WILLIAMS AND P. W. J. L. BRAND
United Kingdom Infrared Telescope Unit, Royal Observatory Edinburgh; and University of Edinburgh

AND

W. H.-M. KU
Columbia Astrophysics Laboratory, Columbia University
Received 1989 May 30; accepted 1989 September 28

ABSTRACT

We present 20 years of optical, infrared, and radio monitoring data for BL Lac as well as four simultaneous multifrequency spectra covering the 10^9–10^{18} Hz range. Although there is no time delay between the optical and infrared variability, the high-frequency radio variations precede lower frequency variations, but only by weeks. The optical variability precedes the radio variability by a few years. The structure function for the radio variations is nearly that of shot noise for time scales less than 600 days, in contrast to the optical variation, which is similar to flicker noise. These results indicate that, although there are fundamental differences between the optical and radio emitting regions, they are related, possibly by the propagation of shocks between regions.

The multifrequency spectra show that the power per logarithmic bandwidth has a well-defined peak in the near-infrared and a sharp cutoff in the optical-ultraviolet region. This cutoff is like those seen in a few other blazars and is attributed to synchrotron losses that prevent particle acceleration from exceeding a critical energy. The X-ray continuum is not smoothly connected to the optical-ultraviolet emission and has a flatter slope, similar to that of the infrared-millimeter region. In addition, the X-ray emission varied in the same sense as the infrared-millimeter emission but opposite that of the optical-ultraviolet emission. These X-ray properties are those expected from the synchrotron–self-Compton process. The best model suggests that the plasma radiating at $\sim 10^{11.5}$ Hz has a size of $\sim 10^{-2.5}$ pc, a Doppler parameter of $\delta = 2$–3, and a magnetic field of $B = 2$–40 G.

Subject headings: BL Lacertae objects — infrared: sources — radiation mechanisms — radio sources: galaxies

I. INTRODUCTION

The nonthermal continua of active galactic nuclei (AGNs) are most prominent in optically violent variable quasars and BL Lacertae objects, collectively referred to as blazars. These sources are polarized and variable, with many of them bright enough for study in all wavebands. In our previous investigations, we measured the continuous spectrum of several of these objects, each distinguished by some particular property such as superluminal motion, flaring, or unusual X-ray brightness (1413+135, Bregman *et al.* 1981; I Zw 187, Bregman *et al.* 1982; 1156+295, Glassgold *et al.* 1983; 0725+178, Bregman *et al.* 1984; 3C 345, Bregman *et al.* 1986; and 3C 446, Bregman *et al.* 1988). Related studies have been undertaken by several other groups that included some of the present authors (e.g., Kondo *et al.* 1981; Worrall *et al.* 1982; Mufson *et al.* 1984; Maraschi *et al.* 1985; Landau *et al.* 1986; Makino *et al.* 1987). Here we investigate BL Lacertae, archetype of its class and one

[1] In view of the large number of collaborators that are required for multifrequency programs, we list the authors by research groups, but emphasize that each group has made important contributions to the data and analysis of this study.

[2] Work also carried out at the National Radio Astronomy Observatory, which is operated by Associated Universities, Inc., under contract with the National Science Foundation.

of the most interesting and important sources for studies of this kind.

BL Lac, which lies in a giant elliptical galaxy at a redshift of 0.0688 (Miller, French, and Hawley 1978), is one of the best-studied blazars at optical and radio wavelengths. Nearly all the radio flux is emitted from a region smaller than 100 mas (Stannard and McIlwrath 1982). Faint extended emission (only 29 mJy) comes from a region 48 kpc in extent ($H_0 = 75$ km s^{-1} Mpc^{-1}; used throughout). Superluminal components are seen frequently on the milliarcsecond scale, and during 1980–1984, four components emerged with apparent velocities of $3.3c$ (Mutel and Phillips 1987). Both the core and the knots are sometimes unresolved, even with resolution better than 0.5 mas (Baath *et al.* 1981; Wardle and Roberts 1988). The emission from the superluminal components is polarized, and the angle of the derotated polarization vector (5°–25°) is the same as the position angle of the VLBI structure (8°–18°; Hughes, Aller, and Aller 1985, 1988; Phillips and Mutel 1988; Mutel and Phillips 1988; Wardle and Roberts 1988). This has led to the suggestion that a weak shock propagates along a jet, locally compressing the field nearly perpendicular to the jet axis and yielding significant polarized emission in the post-shock region (Marscher and Gear 1985; Hughes, Aller, and Aller 1989*a*, *b*). To accommodate the simple relativistic aberration model for superluminal motion, a model is used in which the shocked downstream fluid has a Lorentz factor (observer's frame) of $\gamma = 4.0$, the shock propagates at $\gamma = 2.5$, and the jet axis lies about 38° from the line of sight (Hughes, Aller, and Aller 1989*b*). From the data presented here, shocks and bulk relativistic motion are also inferred.

The optical and infrared emission of BL Lac is variable and polarized, with some correlation indicated between radio and optical polarizations (Sitko, Schmidt, and Stein 1985). The distribution of optical polarization angles has a peak near 20° and a broad width, which becomes more narrow as the brightness of BL Lac decreases (Moore *et al.* 1982; Brindle *et al.* 1985; Sitko, Schmidt, and Stein 1985; Smith 1986; Hagen-Thorn and Yacovleva 1987; Moore, Schmidt, and West 1987). Changes in the polarized flux of 1° hr^{-1} and the polarization angle of 5° hr^{-1} have led to the suggestion that several small (≥ 1 lt-hr), variable, and highly polarized components produce the net observed emission (Moore, Schmidt, and West 1987; see also Bjornsson 1986). The variability behavior suggests a total size of the optical emitting region of approximately a few light days.

In this work, we present both monitoring data and multifrequency spectra for BL Lac. The monitoring data at optical, infrared, millimeter, and centimeter wavelengths cover a 20 year period and permit a detailed search for correlated variability. In addition, we determine the character of the variability by using structure functions (§ II). The multifrequency spectra were obtained during outbursts as well as during more normal periods. These spectra include ground-based observations as well as X-ray data from the *Einstein* and *EXOSAT* satellites, ultraviolet data from the *IUE*, and infrared data from *IRAS* (§ III). These multifrequency spectra permit us to investigate the origin of the X-ray emission and to determine the physical parameters of the emitting regions (§ IV). As a result, a picture of the emitting region that extends over a wide range of size scales is realized.

II. MONITORING DATA

Because of its distinction as the object that defines its class, BL Lac has been extensively studied for a period of about 20 years in several wavebands. Here we present the most complete set of optical, near-infrared, and radio monitoring data available for BL Lac, and we discuss the nature of its variability by employing structure functions and correlation analysis. Corrections for galactic absorption have not been applied to any of the optical monitoring data, some of which has previously been published elsewhere (Pica *et al.* 1980; Pollock 1983; Webb *et al.* 1988; Aller *et al.* 1985; Hagen-Thorn, Marchenko, and Yacovleva 1985).

a) Methodology

Variability of data can be described in a statistical manner by the use of structure functions (e.g., Simonetti, Cordes, and Heeschen 1986). As used here, the structure function provides a measure of the mean difference in the flux densities as a function of the separation in the sampling interval. Formally, it is defined as

$$SF(\Delta t) = \langle [F(t) - F(t + \Delta t)]^2 \rangle .$$

A power spectrum analysis gives information similar to a structure function analysis, but the latter method has advantages when examining nonperiodic data. There is a simple correspondence between power laws in a Fourier analysis, fractal analysis, or structure function analysis. If $SF \propto t^a$, then the fractal function would be proportional to $t^{a/2}$ and the Fourier power spectrum $P \propto f^{-(a+1)}$, where f is frequency.

A second method of analysis used here is one in which we search for a correlation in the variability at two different wavelengths. A simple approach is taken (Bregman *et al.* 1988) in which the two data sets are scaled and normalized such that the total range is between -0.5 and 0.5 with a mean of 0. One data set is shifted in time (Δt) with respect to the other, and the mean variance is calculated. This is formally expressed as:

$$MV = \frac{1}{n} \sum_{1}^{n} [F(t + \Delta t) - G(t)]^2 ,$$

where F and G are the two data sets. In practice, an observed value of $G(t)$ may not have been made at a particular shifted time $t + \Delta t$. In these cases, interpolated values are used, where the interpolation was carried out for the better sampled and more slowly varying of the two data sets.

b) Radio Data

The radio data were obtained at the University of Michigan Radio Observatory (Aller *et al.* 1985), at Haystack Observatory, at the NRAO (Balonek 1982; Johnston *et al.* 1984), and with the Aerospace 4.6 m millimeter wave telescope (Epstein *et al.* 1982). Figure 1 shows that the behavior of the flux densities in the different radio bands is quite similar, and this leads to similar structure functions. The most uniform and accurate data over the longest time scale are at 8 GHz, for which the structure function has a slope $a = 0.8$ for 50 days $< \Delta t <$ 600 days and flattens to nearly zero slope for longer time scales (Fig. 2). The results for 4.8 GHz and 15 GHz are not significantly different ($a = 0.8$–1.0). Using a shorter time base, Johnston *et al.* 1984) and Fiedler *et al.* (1987) analyzed 2.7 and 8.1 GHz data closely spaced in time and found that $a = 0.9$ for 2 days $< \Delta t <$ 100 days (Johnston *et al.* 1984) and $a = 1.4$, 1.3 (for 2.7 and 8.1 GHz, respectively) for 10 days $< \Delta t <$ 200 days (Fielder *et al.* 1987). These results indicate that the structure function has a similar slope for $\Delta t = 2$–600 days and that this slope is nearly that of shot noise ($a = 1$). It also implies that there is no unique minimum time scale. The shortest time scale

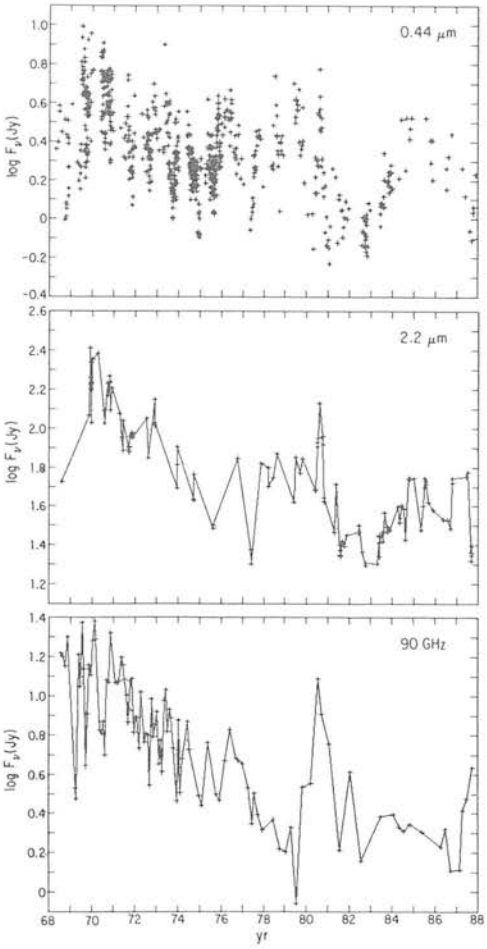


FIG. 1*a*

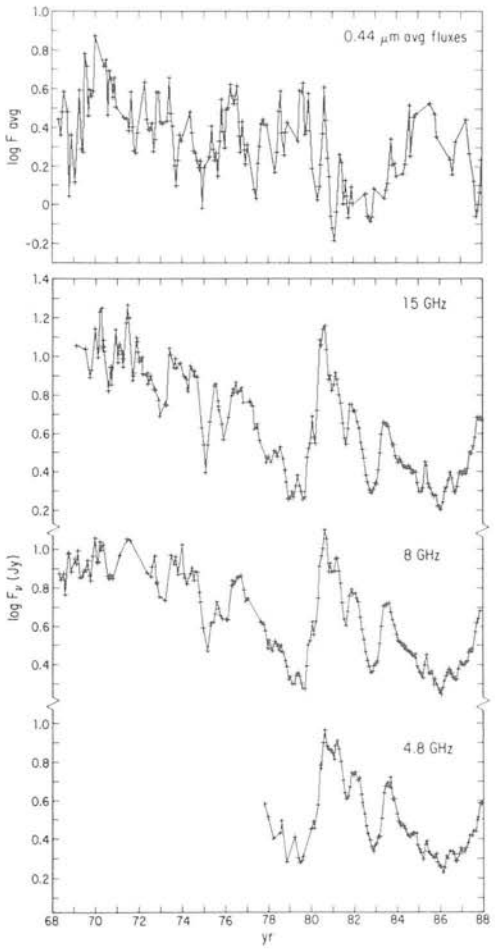


FIG. 1*b*

FIG. 1.—(*a*) Monitoring data at 0.44 μm, 2.2 μm, and 90 GHz during 1968–1988 for BL Lac. (*b*) Monitoring data during 1968–1988 for BL Lac. The 0.44 μm data is averaged into 3 month bins, while monthly averages are given for the 4, 8, 8, and 15 GHz measurements.

(dt/d log F) yet seen is 20 days (determined by observations where $\Delta t = 2$ days). If this time scale is a measure of the size of the radiating region, an enormous brightness temperature is implied. This dilemma is often resolved by appealing to Doppler boosting, by a special geometry, or both. The flattening of the structure function at $\Delta t \simeq 600$ days ($a = 0.0$–0.4) shows that the most common variability occurs with a time scale less than 2 yr.

Correlation analyses of the radio observations show that the data at any two frequencies are well correlated with little time delay (Fig. 3). The 90 GHz data seem to precede the 15 GHz data by about three weeks, the 15 GHz data precede the 8 GHz data by about two weeks, and the 8 GHz data precede the 5 GHz data by about two weeks (delays less than 30 days are difficult to measure; Fielder *et al.* 1987 find the 8.1 GHz variation precede those at 2.7 GHz by 31 days). The small time delay in flux variation (noted previously by Johnston *et al.* 1984; Aller, Aller, and Hughes 1985) suggests that the radio region is remarkably transparent most of the time. In particular, the data rule out the expanding source model in which the opacity decreases with time, leading to outbursts at later times for lower frequency radiation. However, the observations are well-fitted by a propagating shock model involving *in situ* particle acceleration (e.g., Hughes, Aller, and Aller 1989*a*, *b*).

c) Optical-Infrared Data

Most of the optical data were obtained at the Rosemary Hill Observatory by Smith and collaborators using photographic techniques (Webb *et al.* 1988). Additional optical monitoring data is available in the literature for BL Lac (Pollock 1983, and references therein), and this has been included in the analysis. To check that the results are not influenced by the inhomogeneous nature of this data set, analysis was also carried out using only the Rosemary Hill data and using weighted monthly averages of the complete data set. Most of the infrared data were obtained at Palomar observatory by Neugebauer and collaborators, providing a uniform data set with errors of typically 5% or less. However, beam size effects can be important when BL Lac is faint. The observed infrared flux can differ by 10%–15% between measurements made with 5″ and 10″ diameter apertures. The characteristics of the observing system, the calibration standards, and the stability of the system are described in Neugebauer *et al.* (1989).

The infrared and optical data have a greater amount of rapid flicker than the radio data (Fig. 1), and this is verified by a difference in the structure functions (Fig. 2); rapid infrared polarization variations have been discussed by Impey *et al.* (1984). For the optical and infrared data sets, the structure

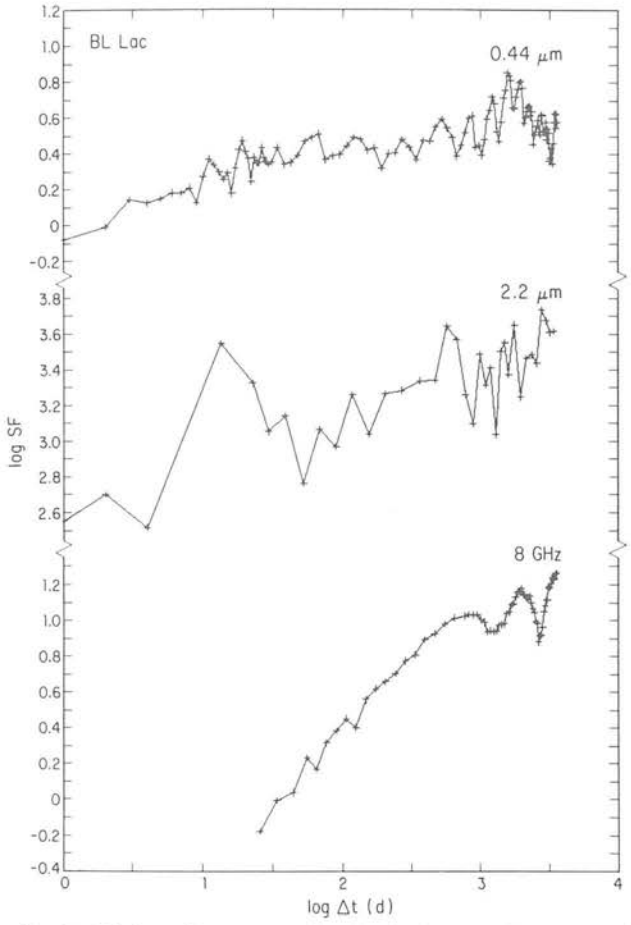


FIG. 2.—The log of the structure function is plotted against the log of the time interval for optical, infrared, and radio frequencies. The large scatter at 2.2 μm is due to the paucity of data; the feature at 10 days may not be statistically significant. Notice the similarity between the optical and infrared structure functions and their difference with that at 8 GHz.

function is fairly flat for $\Delta t > 100$ days with a slope $a = 0.0$–0.3, while for $\Delta t < 100$ days, $a = 0.2$–0.4 (Fig. 2). This demonstrates that there exists considerably more power at short time scales than in the radio data. The amount of power in short-term variations can change with time, and Smith, Leacock, and Webb (1987) have noted that since 1980, flickering is less prominent (although still with much more power than radio flickering).

Cross-correlation studies were performed to determine the connection between optical and infrared variability. A strong correlation is found between the 0.44 and 2.2 μm fluxes with no measurable delay (Fig. 3); formally, the 2.2 μm fluxes lag behind with $\Delta t = 1.5 \pm 2.0$ days.

The multicolor infrared observations permit a study of the spectral variation. Between 1976 and 1988, the flux density was measured at 1.25, 1.65, and 2.2 μm sequentially on the same night on 68 different occasions. In general, the 1.25–2.2 μm slope lies in the range -1.0 to -1.3 (Fig. 4), and most of the dispersion is probably due to the uncertainty in the slope (typically 0.14). However, during the outburst in 1980, when BL Lac achieved its greatest infrared brightness in the 1976–1988 period (Fig. 1), the spectrum was noticeably steeper. Steepening also occurred in the 2.2–10.1 μm range, with α decreasing by 0.15. Thus, the entire infrared spectrum softened during the outburst.

d) The Optical-Radio Connection

For this correlation study, the 0.44 μm flux densities were averaged into 1 month intervals in order to avoid the problem that sampling was more frequent during the first 9 years; monthly averages of the radio data were also used. Correlation analyses of the optical-infrared data with the radio data indicate that the variation of the radio fluxes lags behind that of the optical-infrared fluxes by several years (Fig. 3). There is a broad but significant minimum in the mean variance between 1.0–4.5 yr for the correlation of both the optical and infrared fluxes with the 8 GHz data. The minimum is not as broad when the optical data is compared to the 15 GHz fluxes. Only one well-defined minimum is seen near 4.0 yr.

Although the structure functions for radio and optical data are significantly different for time scales less than a few hundred days, they are more similar for longer time scales (600 days $< \Delta t <$ 3500 days). The comparison is hampered because the data extend to only about 7000 days, so a few outbursts with spacings of years add features to the structure functions. When combined with the correlation analyses, we conclude that the long time scale variations at both the optical-infrared and radio regions are caused by the same outbursts.

III. MULTIFREQUENCY SPECTRA

Multifrequency spectra were obtained on four dates: 1980 June, 1983 June, 1983 December, and 1984 May. Radio, millimeter, infrared, optical, and ultraviolet data were obtained on each occasion, far infrared/submillimeter data were obtained for two spectra (1983 June and 1983 December), while X-ray measurements were obtained during 1980 June and 1983 December (Table 1, Fig. 5). Because of the large number of observatories and instruments used, the details of each observation are not given when the procedures are standard and have been described elsewhere.

The reddening to BL Lac has been determined by several approaches. Du Puy *et al.* (1969) examined the reddening of halo K giants near the line of sight to BL Lac and obtained $E(B-V) = 0.31$, while Thuan, Oke, and Gunn (1975) estimated $0.28 < E(B-V) < 0.40$. We also estimated the reddening from the H I column density in that direction (Burstein and Heiles 1978) to obtain $E(B-V) = 0.26 \pm 0.06$. For the purposes of this work, we adopt the value $E(B-V) = 0.31$ when correcting the infrared, optical, and ultraviolet observations and values of $A(\lambda)/E(B-V)$ given by Seaton (1980). To correct the X-ray observations for absorption, $\log N_{\rm H} = 22.1$ is used, which is the value determined directly from a spectral fit to the pulse-height data of the Einstein Observatory IPC observation. This is greater than the galactic value determined from H I maps ($\log N_{\rm H} = 21.3$; Burstein and Heiles 1978) and suggests that additional absorbing material may lie in the host galaxy.

a) General Features of a Single Spectrum

The multifrequency spectra, displayed as power per logarithmic bandwidth (νF_ν) versus $\log \nu$, are shown in Figure 5. The striking peak near $\log \nu = 14.3$ is caused by a remarkable steepening of the optical-ultraviolet continuum, which is nearly an exponential function of frequency. This is seen more clearly in Figure 6, which shows the mean spectral slope of all multifrequency spectra. If the optical-ultraviolet spectrum is extrapolated to higher frequency, no observable X-ray emis-

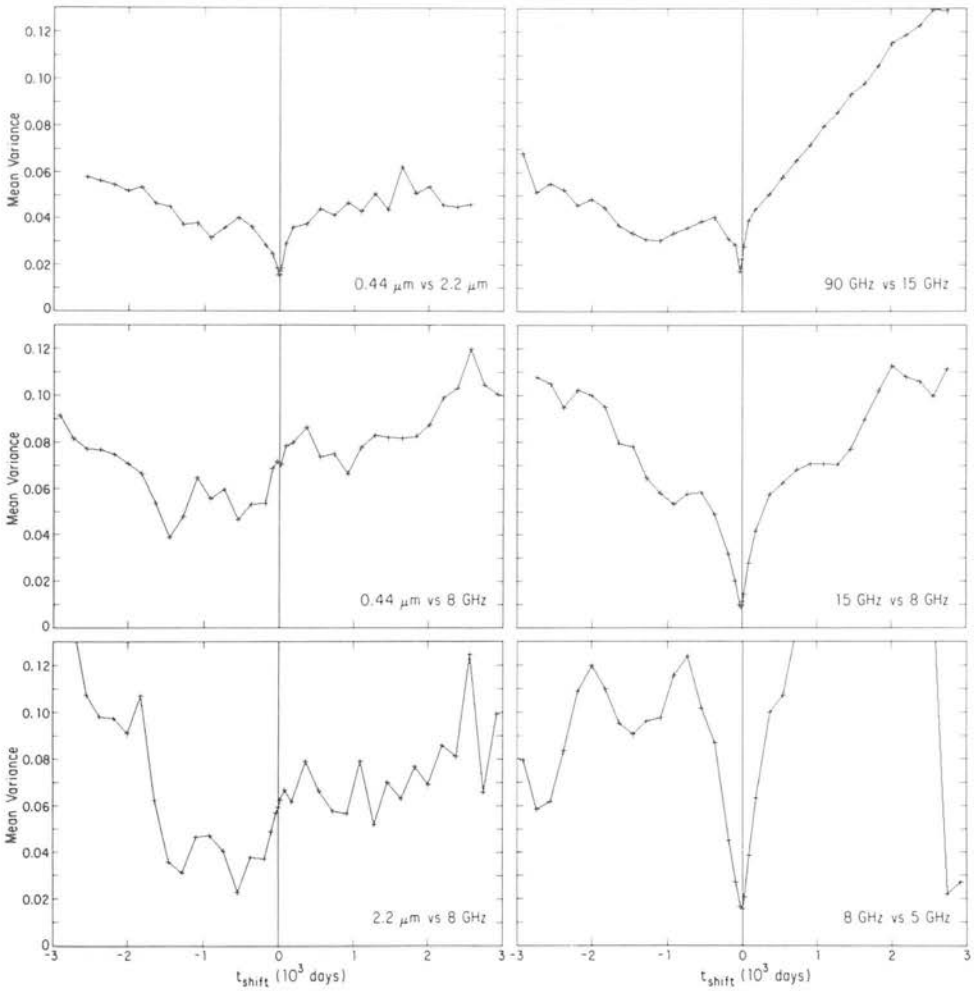


FIG. 3.—The mean variance (MV) of different pairs of fluxes is shown as a function of time shift. A minimum in MV at negative Δt indicates that a correlation exists, with the first listed quantity preceding the second by that amount. Most of the figures have values for MV calculated at $\Delta t = 0, \pm 30$, and ± 90 days, and then at half-year intervals beginning with 0.5 yr.

sion is expected. X-ray emission was, however, clearly detected on two occasions, indicating that the X-rays are not an extension of the synchrotron emission in the radio-ultraviolet region (Fig. 5).

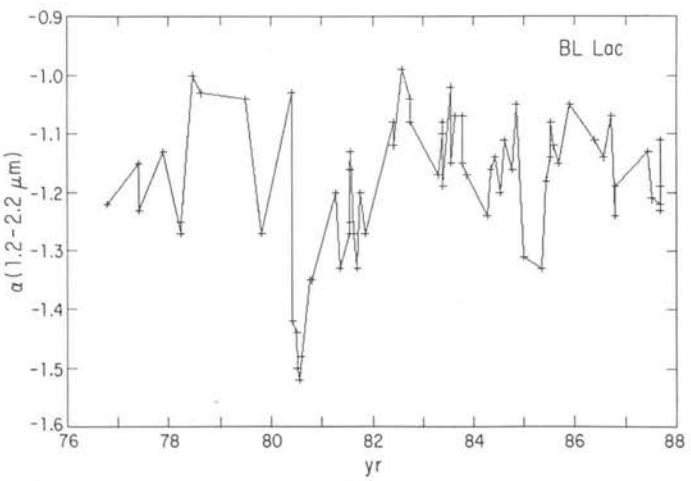


FIG. 4.—The infrared slope for BL Lacertae is determined with a typical error of 0.14. The spectral steepening that occurred in 1980 is associated with an outburst when BL Lac became brighter than at any other time during 1976–1988.

Most of the radiated power emerges in the infrared-optical region, with only minor contributions from other wavebands. It is difficult to determine accurately the total X-ray luminosity because the peak in the power per logarithmic bandwidth in the 10–500 keV range is not observed. To estimate the integrated X-ray flux, we assumed that the X-ray continuum is described by a power law of slope $\alpha = -0.7$ (suggested by the data; see below) to an energy of 100 keV. Under these assumptions, the total integrated X-ray flux is about 10% of the radio-UV flux (Table 2); extending the X-ray continuum to 500 keV (the electron pair production energy) increases the integrated X-ray flux to 16% of the radio-ultraviolet flux.

b) Specific Spectra

1980 June.—The source was undergoing a radio outburst at this time, although there was no simultaneous optical outburst. As is frequently seen during the brightening phase of radio outbursts, the radio spectrum is inverted in the cm region and flattens in the mm region.

The X-ray spectrum is composed of measurements from two instruments, the IPC and the MPC of the *Einstein Observatory* (integration time of 4323 s; sequence number 5693). The spectral fit to the IPC data (0.1–4 keV) shows a slope of $\alpha = -2.2 \pm 1.0$, the slope determined from the MPC data (2–10 keV) is $\alpha = -0.8 \pm 0.5$, and the spectrum connecting the MPC

TABLE 1

MULTIFREQUENCY DATA FOR BL LAC

Name	Date	Region	logν (Hz)	Flux	log F_ν[1]	log σ
Ku	8 Jun 80	1 keV	17.38	1.60E-29[2]	-28.80	0.06
Columbia		5.6 keV	18.13	5.00E-30	-29.30	0.12
Bregman, Glassgold,	8 Jun 80	2525[3]	15.075	3.69E-16[4]	-26.20	0.05
and Huggins		2850	15.022	5.95E-16	-26.06	0.03
NYU and NRAO						
Webb, Pollock, Pica,	11 Jun 80	3600	14.92	0.44[5]	-25.75	0.06
Smith, and Leacock		4400	14.83	1.36	-25.36	0.06
Univ. Florida		5500	14.74	2.51	-25.22	0.06
Miller	8 Jun 80	4200	14.85	0.83	-25.55	0.02
Lick Obs.		4500	14.82	1.08	-25.47	0.02
		5000	14.78	1.66	-25.34	0.02
		5500	14.74	2.36	-25.24	0.02
		6000	14.70	3.04	-25.17	0.02
		6500	14.66	3.72	-25.11	0.02
		7000	14.63	4.68	-25.04	0.02
		7500	14.60	5.72	-24.98	0.02
Williams, Brand	1 Jun 80	1.25[6]	14.38	25.0	-24.51	0.02
UKIRT		1.65	14.26	39.1	-24.35	0.02
		3.80	13.90	100	-24.00	0.02
Neugebauer, Soifer	2 Jun 80	1.25	14.38	19.7	-24.61	0.02
Palomar Obs.		1.65	14.26	31.0	-24.45	0.02
		2.2	14.13	46.2	-24.30	0.02
Balonek, Dent, and	Jun 80	89.6[7]	10.95	9.20[8]	-22.04	0.03
Barvanis	Jun 80	31.4	10.50	12.50	-21.90	0.01
Univ. Mass.						
Flett and	2 Jun 80	33.5	10.53	15.0	-21.82	0.02
Henderson (1981)						
Aller, Aller,	6 Jun 80	14.5	10.16	12.9	-21.889	0.013
and Hodge	8 Jun 80	8.0	9.90	9.2	-22.036	0.014
Univ. Mich.	7 Jun 80	4.8	9.68	6.11	-22.214	0.014
Bregman, Glassgold,	11 Jun 83	2600[3]	15.062	3.61E-16[4]	-26.26	0.07
and Huggins		2700	15.046	4.63E-16	-26.16	0.07
NYU and NRAO		2800	15.030	8.08E-16	-25.92	0.07
		3000	15.000	1.00E-15	-25.84	0.08
		3100	14.986	9.38E-16	-25.85	0.08
Webb, Pollock, Pica,	9.4 Jun 83	4400	14.83	16.50[9]	-25.47	0.05
Smith, and Leacock	10.4 Jun 83	3600	14.92	16.29	-25.82	0.04
Univ. Florida		4400	14.83	16.57	-25.47	0.03
		5500	14.74	15.34	-25.16	0.05
	11.4 Jun 83	3600	14.92	16.59	-25.80	0.07
		4400	14.83	16.58	-25.47	0.03
		5500	14.74	15.26	-25.16	0.04

TABLE 1—*Continued*

Name	Date	Region	logν (Hz)	Flux	log F_ν	log σ
Sitko et al. 1985	7 Jun 83	3600	14.92	0.40[5]	-25.79	0.02
		4400	14.83	1.14	-25.43	0.02
		5500	14.74	2.51	-25.22	0.02
		6400	14.67	4.04	-25.07	0.02
		7900	14.58	6.71	-24.93	0.02
	8 Jun 83	3600	14.92	0.40	-25.79	0.02
		4400	14.83	1.06	-25.46	0.02
		5500	14.74	2.18	-25.28	0.02
		6400	14.67	3.69	-25.11	0.02
		7900	14.58	6.06	-24.98	0.02
Neugebauer, Soifer	21 May 83	1.25[6]	14.38	12.70	-24.80	0.03
Palomar Obs.	-25 May 83	1.65	14.26	18.30	-24.68	0.02
		2.2	14.13	25.30	-24.56	0.02
		3.7	13.91	40.2	-24.40	0.04
		10.1	13.47	97.7	-24.01	0.07
Neugebauer, Miley,	30 May -	12	13.40	128	-23.89	0.02
and others	16 Jul 83	25	13.08	242	-23.62	0.01
IRAS		60	12.70	466	-23.33	0.01
Dent, Kinsel, and Balonek	15 Jun 83	83.45[7]	10.92	2.42[8]	-22.62	0.03
Univ. Mass.						
Howard	11 Jun 83	89.60	10.95	3.40	-22.47	0.05
NRAO						
Aller, Aller,	13 Jun 83	14.5	10.16	4.39	-22.357	0.004
and Hodge	2-23 Jun 83	8.0	9.90	5.04	-22.298	0.004
Univ. Mich:	1-24 Jun 83	4.8	9.68	4.72	-22.326	0.004
Maccagni, Garilli,	8 Dec 83	1 kev	17.38	8.00E-30[2]	-29.10	0.06
and Giommi						
CNR, EXOSAT						
Bregman, Glassgold,	8 Dec 83	2650[3]	15.054	1.75E-15[4]	-25.57	0.12
and Huggins		2750	15.038	1.67E-15	-25.60	0.06
NYU and NRAO		2850	15.022	1.27E-15	-25.73	0.09
		2950	15.007	1.46E-15	-25.66	0.07
Miller and Stephens	6 Dec 83	4500	14.82	3.09E-15	-25.18	0.02
Lick Obs.		5000	14.78	3.13E-15	-25.14	0.02
		5500	14.74	3.32E-15	-25.09	0.02
		6000	14.70	3.58E-15	-25.02	0.02
		6500	14.66	3.68E-15	-24.97	0.02
		7000	14.63	3.84E-15	-24.91	0.02
Webb, Pollock, Pica,	6 Dec 83	4400	14.83	16.05[9]	-25.26	0.05
Smith, and Leacock	7 Dec 83	3600	14.92	16.01	-25.53	0.06
Univ. Florida		4400	14.83	15.92	-25.21	0.04
		5500	14.74	14.92	-25.02	0.05
	8 Dec 83	3600	14.92	16.00	-25.52	0.04
		4400	14.83	16.01	-25.24	0.04
		5500	14.74	14.91	-25.02	0.03
Wisniewski	16 Dec 83	3600	14.92	16.19	-25.60	0.01
Steward Obs.		4400	14.83	16.26	-25.34	0.01
		5500	14.74	15.21	-25.14	0.01
		7000	14.63	14.34	-25.01	0.01
		9000	14.52	13.35	-24.80	0.01

TABLE 1—*Continued*

Name	Date	Region	logν (Hz)	Flux	log F_ν	log σ
Neugebauer, Elias,	19 Nov 83	1.25[6]	14.37	15.0[5]	-24.73	0.03
Soifer, and Matthews		1.65	14.26	21.5	-24.61	0.03
Palomar Obs.		2.2	14.13	30.1	-24.49	0.03
		3.7	13.91	44.1	-24.35	0.03
		10.1	13.47	111.0	-23.94	0.06
Roellig	1-4 Dec 83	1.25	14.37	15.8	-24.71	0.02
NASA, Palomar Obs.		1.65	14.26	23.3	-24.57	0.02
		2.2	14.13	31.9	-24.47	0.02
		3.7	13.91	47.1	-24.32	0.02
	20 Jan 84	1000	11.48	1300.0	-22.19	0.48
Harvey et al. 1984	29 Oct 83	50	12.78	<180	<-23.03	upper
		100	12.48	<400	<-22.87	limits
Balonek, Dent, and Kinsel	19 Jan 84	89.6[7]	10.95	2.50[8]	-22.60	0.02
Aller, Aller,	3 Dec 83	14.50	10.16	3.040	-22.517	0.006
and Hodge	5 Dec 83	8.00	9.90	3.540	-22.451	0.011
Univ. Mich.	13 Dec 83	4.80	9.68	3.520	-22.453	0.006
Bregman, Glassgold,	30 May 84	2250[3]	15.125	1.23E-15[4]	-26.52	0.18
and Huggins		2450	15.088	9.02E-16	-25.79	0.13
NYU and NRAO		2550	15.071	1.02E-15	-25.79	0.05
		2650	15.054	7.73E-16	-25.94	0.05
		2750	15.038	9.01E-16	-25.88	0.04
		2850	15.022	1.19E-15	-25.76	0.04
		2950	15.007	1.17E-15	-25.76	0.04
		3050	14.993	9.81E-16	-25.83	0.05
Miller and Stephens	31 May 84	4500	14.82	2.65E-15	-25.25	0.02
Lick Obs.		5000	14.78	3.18E-15	-25.14	0.02
		5500	14.74	3.48E-15	-25.06	0.02
		6000	14.70	3.71E-15	-25.00	0.02
		6500	14.66	3.92E-15	-24.94	0.02
		7000	14.63	4.11E-15	-24.88	0.02
Neugebauer, Elias,	8 Jun 84	1.25[6]	14.37	21.5[5]	-24.58	0.03
Soifer, and Matthews		1.65	14.26	30.2	-24.46	0.03
Palomar Obs.		2.2	14.13	40.4	-24.36	0.03
		3.7	13.91	53.0	-24.27	0.03
Dent, Kinsel, and	29 Apr 84	89.60[7]	10.95	2.20[8]	-22.66	0.03
Balonek	12 Jul 84	89.60	10.95	2.07	-22.68	0.05
Univ. Mass.						
Aller, Aller	2 Jun 84	14.50	10.16	2.637	-22.579	0.004
Univ. Mich.	4 Jun 84	8.00	9.90	3.036	-22.518	0.006
	10 Jun 84	4.80	9.68	2.669	-22.574	0.016

[1] Corrected for $E(B-V) = 0.31$ or $\log N_H = 22.1$.
[2] In ergs cm^{-2} s^{-1} Hz^{-1}.
[3] In angstroms.
[4] In ergs cm^{-2} s^{-1} A^{-1}.
[5] Flux density in mJy.
[6] Wavelength in μm.
[7] Frequency in GHz.
[8] Flux density in Jy.
[9] Magnitudes.

TABLE 2

THE RADIO-ULTRAVIOLET CONTINUUM

Date	Integrated Flux (10^{-10} ergs cm^{-2} s^{-1})	Luminosity ($H_0 = 75$) (10^{45} ergs s^{-1})
1980 Jun	4.5	4.2
1983 Jun	2.2	2.1
1983 Dec	2.1	2.0
1984 May	2.5	2.5

and IPC data is $\alpha = 0.68 \pm 0.18$ (between energies 5.6 keV and 1.0 keV). There is no evidence for brightness variation during the observing period, and a 1 σ limit of 30% can be set on minute-to-minute variations. Fourier analysis suggests that there may be periodic variations in the data of 53.4 and 9.65 s (at the 95% level), although more data are needed to produce a secure result.

1983 June.—Monitoring data show evidence for a minor radio outburst with a peak near 1983 June, in contrast to the infrared and optical observations, which show low values at this time (Fig. 1). *IRAS* data were obtained during a period of a few months that encompass 1983 June (Impey and Neugebauer 1988); the flux determination at 100 μm is confused by IR cirrus and is not included. No variation was measured in any of the *IRAS* bands. The NRAO 12 m telescope had been

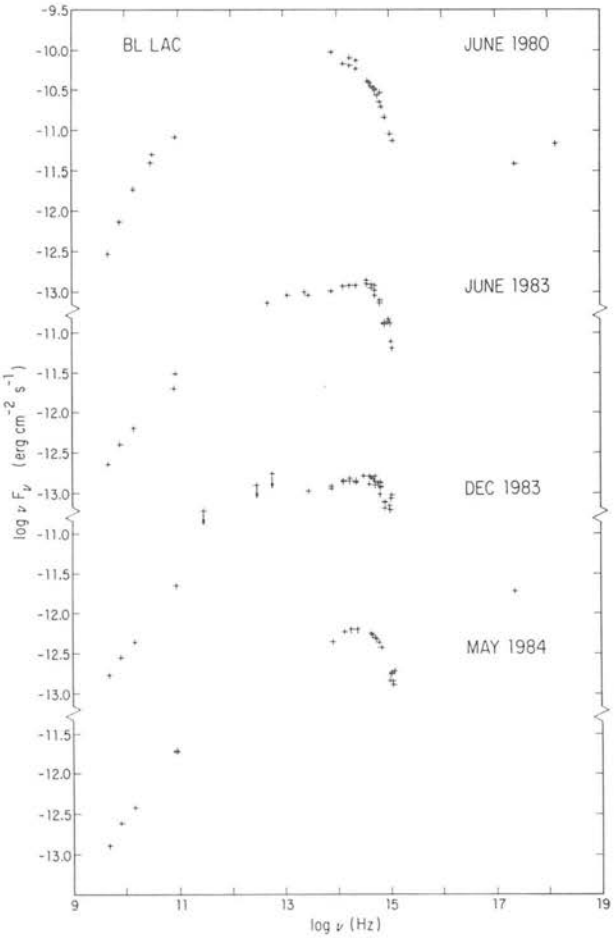


FIG. 5.—The multifrequency spectra of BL Lac on four separate occasions are plotted in power per logarithmic bandwidth vs. frequency. X-ray measurements were taken in 1980 June and 1983 December.

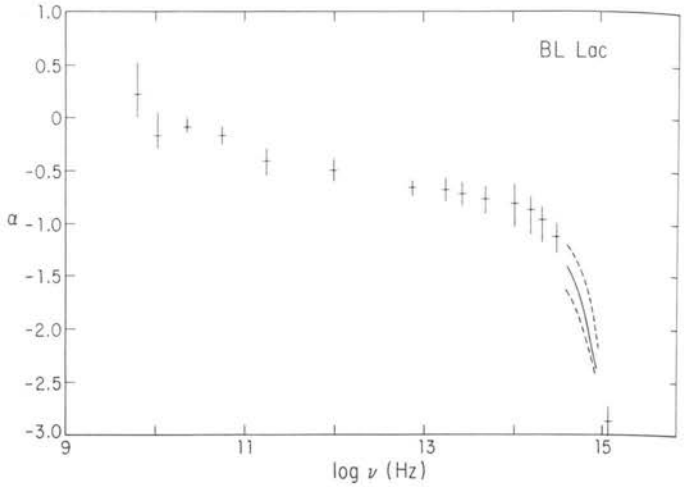


FIG. 6.—The average slope for the four multifrequency spectra. The vertical bars indicate the 1 σ range of the slopes and not the error in determining an individual slope in a multifrequency spectrum.

resurfaced shortly before the 90 GHz observations, and pointing uncertainties probably led to the discrepancy between the two observations.

The spectral slope changes from a value of about $\alpha = -0.3$ in the radio-mm region to a value of about -0.8 in the infrared-optical region. The data suggest that this change occurs near 10^{12} Hz, and this may denote the location where the synchrotron source changes from opaque to transparent.

1983 December.—The monitoring observations show a steady decrease in the radio flux density from mid-1983 at the same time that there is a slow increase in the mean optical-infrared brightness. 1983 December is the only data in addition to 1980 June when simultaneous X-ray data were obtained. The *EXOSAT* observations were processed using the spectral index and absorption column deduced from the *Einstein Observatory* IPC observations. The value of the flux density, which depends upon the assumed spectral parameters, may be uncertain by 50% or more. However, when identical spectral parameters are used for the reduction of the *Einstein Observatory* and *EXOSAT* measurements, the flux density obtained during 1983 December is always fainter.

Like the 1980 June spectrum, an extension of the optical-ultraviolet continuum into the X-ray region falls far below the X-ray emission. Relative to the 1980 June data, flux decreases are evident in the X-ray, infrared, millimeter, and radio data, while flux increases occur in the optical and ultraviolet regions (Fig. 7). This suggests that *the X-ray emission is likely to be related to the infrared-radio emission rather than to the optical-ultraviolet emission.*

The radio continuum is characterized by a slope of $\alpha = -0.1$, which steepens to $\alpha = -0.5$ in the infrared. This change in slope occurs in the 10^{11}–10^{12} Hz region and may be the result of increasing transparency.

1984 May.—The trend of decreasing mean radio flux density and increasing mean optical-infrared brightness continues from mid-1983 to mid-1984 (Fig. 1). In this spectrum, the peak in the power per logarithmic bandwidth is extremely well defined (near log $\nu = 14.3$–14.4), and the optical-ultraviolet steepening is quite sharp. Like the 1983 June and 1983 December spectra, this spectrum shows a radio spectrum with slope $\alpha > -0.5$ that steepens in the region log $\nu = 11$–12.

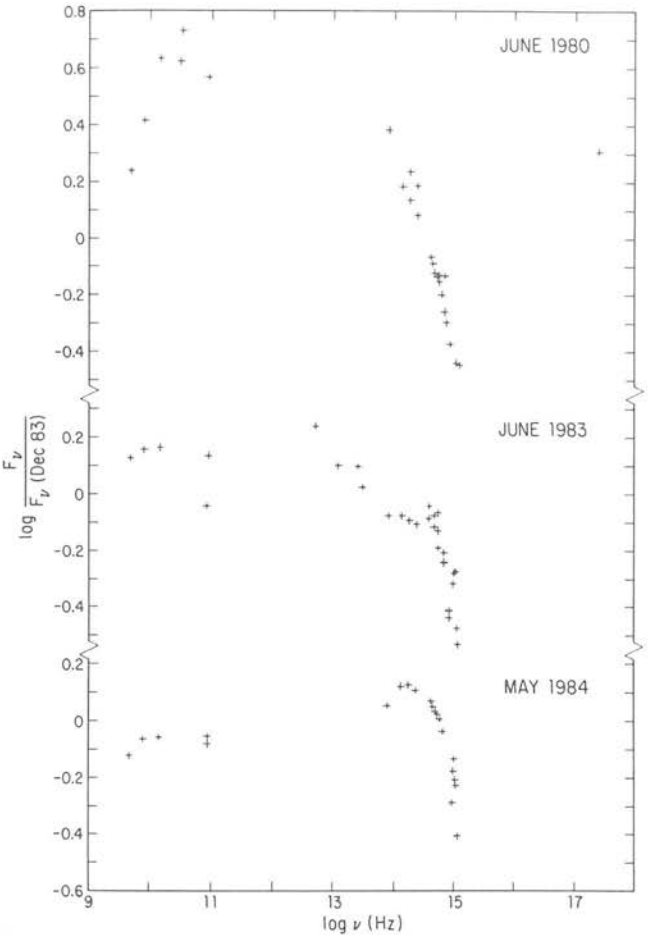


FIG. 7.—The multifrequency spectra are plotted relative to the 1983 December spectrum in order to reveal spectral changes. In 1980 June, the X-ray, infrared, and radio emission were brighter while the optical and ultraviolet continuum was fainter.

c) *Differences Between Spectra*

To see detailed changes in spectra more clearly, we divided the spectra by the spectrum of 1983 December. For the 1983 December spectrum, the exact values of the radio, millimeter and X-ray region were used, and a smooth curve was fit through the infrared-ultraviolet region. The ratios of the fluxes are shown in Figure 7.

Unlike some BL Lac objects and violently variable quasars, whose spectral shapes show little time variation (e.g., Bregman *et al.* 1986, 1988), there are clear changes in the spectrum of BL Lac. Between 1983 June and 1984 May, the fractional brightening in the ultraviolet region is greater than in the optical region. Because of the rapid decrease in the optical-ultraviolet spectrum, a shift in the frequency of spectral turnover could account for the dramatic ultraviolet variations, which, however, contain little power. The decrease in the flux density of the radio-mm continuum between 1983 June and 1984 May reflects the decay of the outburst that peaked near mid-1983. The comparison of the 1983 December with the 1980 June spectrum indicates that the variation of the X-ray emission occurs in the same sense as the radio-infrared variations, but not as the optical-ultraviolet variations.

d) *The Peak of the Power Distribution*

An extensive set of multicolor infrared measurements near the peak of the power per logarithmic bandwidth was obtained at the Palomar Observatory. On 19 occasions, observations were made on the same night in five filter bands: 1.25, 1.65, 2.2, 3.7, and 10.1 μm. In order to show the change in spectral shape, we scaled each spectrum in flux density so that they all have the same 1.25–10.1 μm integrated power. The uncertainty in the data points of a single spectrum is typically $\log \nu F_\nu = 0.03$, which is comparable to or smaller than the characteristic variation in the spectrum (i.e., the vertical bars in Fig. 8). In the majority of spectra, the peak in the power per logarithmic bandwidth is remarkably well defined and occurs either between 2.2 and 1.65 μm. In about one-quarter of the cases, the power per logarithmic bandwidth is evenly distributed from 1.65–10.1 μm. The mean 2.2–10.1 μm slope is $\langle\alpha\rangle = -0.85$, while in the 1.25–2.2 μm range, $\langle\alpha\rangle = -1.19$, which indicates a steepening of $\Delta\alpha = -0.34$ over a frequency range of $\Delta \log \nu = 0.45$. This well-measured and rapid steepening should serve as a strong constraint in the development of successful radiation models.

IV. DISCUSSION

a) *Sharp Spectral Cutoffs*

The remarkable cutoff in the continuum of BL Lac in the infrared-ultraviolet region is similar in shape to that seen in quasars with an unusually large near-infrared to optical brightness ratio (red quasars; Bregman *et al.* 1981; Beichman *et al.* 1981*a*, *b*; Rieke, Lebofsky, and Wisniewski 1982; Rieke, Lebofsky, and Kinman 1979). However, in red quasars, several of which are blazars, the cutoff occurs in the 1–10 μm region rather than at 0.1–1 μm (Fig. 9). X-ray data suggest there may exist a similar cutoff at yet higher frequencies for certain X-ray bright BL Lac objects. For example, Mrk 421 (Makino *et al.* 1987) has an IR-optical-ultraviolet spectrum that extends smoothly into the X-ray region, where the spectrum steepens (Ohashi 1988), implying a rapid cutoff to the emission. Thus, a continuum cutoff may be a common feature that occurs over a range of frequencies.

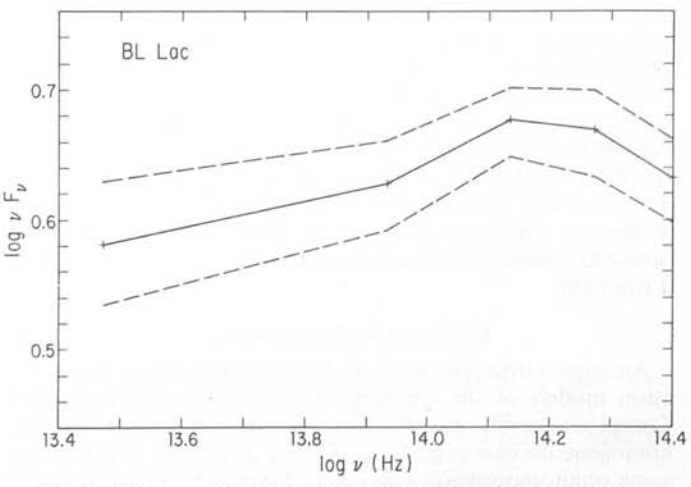


FIG. 8.—The mean infrared spectrum of BL Lac, which was obtained by averaging 19 spectra that were normalized to the same integrated flux. Two-thirds of the spectra are contained within the region defined by the dashed lines. The normalization of the vertical scale is arbitrary.

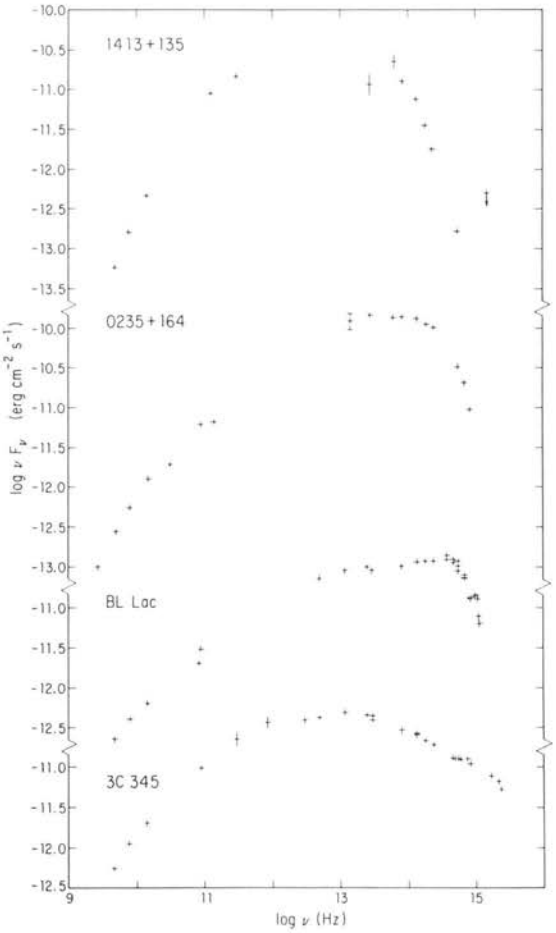


FIG. 9.—The multifrequency spectrum of BL Lac is compared to that of three blazars: the red quasar 1413+135, the BL Lac object AO 0235+164, and the violently variable quasar 3C 345. The sharp infrared-optical turnover occurs at progressively higher frequencies in the sequence proceeding from 1413+135 to AO 0235+164 to BL Lac. No such turnover is observed in 3C 345.

These spectral cutoffs probably reflect a limit in the energy to which electrons are easily accelerated before synchrotron losses become important. To explain this feature, models have been developed in which electrons are accelerated near a shock by first-order Fermi acceleration (e.g., Blandford and Eichler 1987). When synchrotron losses are included, the electrons reach a limiting energy when the acceleration time scale exceeds the synchrotron time scale (Webb, Drury, and Biermann 1984; Bregman 1985; Heavens and Meisenheimer 1987; Biermann and Strittmatter 1987; Fritz 1989). These processes produce a sharp cutoff in the synchrotron continuum, and there has been some success in fitting these models to the data (Fritz 1988).

b) Test of Radiation Models

An important application of the observations is to test radiation models of the emitting region. The synchrotron–self-Compton model makes two simple predictions in the homogeneous case (e.g., Jones, O'Dell, and Stein 1974) and in some nonhomogeneous cases (e.g., Marscher 1977; Ghisellini, Maraschi, and Treves 1985): (1) the continuum of the seed photons and of the Compton emission should be similar, and (2) the seed photons should vary in concert with the Compton emission. For BL Lac, we examine whether the X-ray emission is produced by the Compton process. The 1980 June spectrum indicates that the X-ray slope of -0.68 ± 0.18 is the same as that in the 10^{11}–10^{14} Hz region, which we identify as the seed photon region. The variation in the spectrum between 1980 June and 1983 December shows that the X-ray emission decreased as did the emission in the $10^{9.5}$–$10^{14.4}$ Hz region. The optical and ultraviolet continuum, which are not identified as the seed photons, changed in the opposite sense to that of the X-ray emission. These observations qualitatively support the synchrotron–self-Compton model, and additional multifrequency spectra with accurate X-ray continua are being obtained to investigate this issue more deeply. We note that in a competing model for the production of the X-ray emission, the pair cascade model (e.g., Fabian 1988), interactions between the γ-ray photons and the optical-ultraviolet photons lead to the X-ray emission. In this model, one expects the X-ray emission to be positively correlated with the optical-ultraviolet emission, the opposite of what we observe.

The physical properties of the emitting region can be estimated by fitting the synchrotron–self-Compton model to the data. An important parameter for the model is the frequency at which opacity effects become important (ν_m). The presence of a time delay in the variability between 90 GHz and 15 GHz suggests that the source is not entirely transparent until frequencies greater than 10^{11} Hz are reached. The data also indicate that a change in slope occurs in the 100–1000 GHz range, so we have fit homogeneous models under the assumption that opacity effects first become significant in this spectral range.

A time scale of variability that reflects the size of the emitting region is also needed for the model in order to constrain the Doppler boosting parameter. The turnover in the power spectrum of the optical variability suggests that the size of the entire optical emitting region is about a light-day (Moore, Schmidt, and West 1987). In the radio region, outbursts can develop on a time scale of weeks. In choosing the variability time scale for the 100–1000 GHz region, we have adopted values of half a week and of 1 day.

The models are fit to the two spectra with X-ray measurements, which is essential for constraining the Compton scattering contribution, and the results are presented in Table 3. The best-constrained parameter is the size of the emitting region, which falls in the range 0.003–0.02 pc. Every model indicates that modest relativistic boosting occurs, with $\delta = 2$–16, where $\delta = [\gamma(1 - \beta \cos \theta)]^{-1}$. For the least Doppler boosting (model 1983 December C), the bulk motion of the emitting plasma must be within 30° of our line of sight, while for the largest value of δ (model 1980 June D), the motion must be within 3°.6 of our line of sight. The selection of a preferred model is aided by Hughes, Aller, and Aller (1989*b*), who fit a propagating shock model to an independent data set, the

TABLE 3

SYNCHROTRON SELF- COMPTON MODELS FOR BL LAC

Model	ν_m (GHz)	$F_\nu(\nu_m)$ (Jy)	t_{var} (weeks)	B (G)	r (pc)	n (cm^{-3})	δ
1980 Jun A	100.	8.71	0.5	0.088	0.021	68	14.8
1980 Jun B	316.	6.00	0.5	1.61	0.010	110	7.2
1980 Jun C	1000.	3.50	0.5	30.	0.0046	240	3.3
1980 Jun D	316.	6.00	0.14	0.51	0.0062	190	15.6
1983 Dec A	100.	2.30	0.5	0.13	0.013	200	9.4
1983 Dec B	316.	1.26	0.5	2.4	0.0058	460	4.2
1983 Dec C	1000.	0.68	0.5	44.	0.0026	1080	1.9
1983 Dec D	316.	1.26	0.14	0.75	0.0036	740	9.0

VLBI, and the radio monitoring observations. Their best model has $\delta \approx 1$ and $\theta \simeq 38°$. A few of our models have $\delta \lesssim 4$, but none have $\delta \approx 1$. This difference with the propagating shock model could easily be due to the simplicity of our model or a slight underestimation of ν_m. Our models with the smallest values of δ are 1980 June C, 1983 December B, and 1983 December C (Table 3), which suggest that the plasma emitting at frequency ν_m has a magnetic field strength of 2–40 G. The inability to specify ν_m accurately prevents a comparison of the physical parameters for the two separate dates. The energy density associated with the magnetic field is considerably greater than that associated with the particles, which implies that the rapid radiative losses are balanced by particle reacceleration.

Nonhomogeneous synchrotron-Compton models (e.g., Ghisellini, Maraschi, and Treves 1985) suggest that the plasma emitting at higher frequencies, such as the optical emission, comes from a region smaller by an order of magnitude or more, with a correspondingly more intense magnetic field strength and greater particle density.

c) *Flux Density Correlations*

The monitoring data (§ II) yield information about the structure of the emitting plasma in BL Lac. It is not surprising that the variations in the infrared-optical emission show no time delay, which is expected if the emission is produced cospatially by the synchrotron process in a transparent region. With regard to the radio emission, the time delay between variations in neighboring wavebands is short, suggesting that optical depth effects are small. If the time delays measured between 4.8 GHz and 89.6 GHz can be extrapolated to the infrared-optical region, we estimate that the infrared-optical flux variations should precede the radio variations by about 4 months. However, a much longer time delay is determined from the correlation analysis, 1.5–4 yr. Furthermore, the structure function analysis shows that the behavior of the flux density variations is quantitatively different between the infrared-optical and the radio bands. Both of these results suggest that the radio emission region is not simply the optically emitting plasma that eventually becomes transparent.

The observations might be understood in terms of a model in which disturbances (e.g., input of fresh plasma) are created by outbursts associated with optical emission. The disturbance propagates outward, producing little radiation until it creates a shock in more distant plasma, giving rise to the radio emission. The similarity between the observed time delay of a few years and the light crossing time of the measured VLBI size support this picture. The light crossing time of the radio region is longer than the time separating rapid disturbances in the optical region (that produce optical flickering). This lessens the amplitude of short time scale radio variability, causing the radio structure function to be steeper than its optical counterpart, as is observed.

It may be possible to use the optical monitoring data to plan VLBI observing strategies. The correlation between the optical and radio behavior and the correlation between radio outbursts and superluminal components suggests that superluminal behavior can be linked to optical outbursts. This result was also found for 3C 345 (Bregman *et al.* 1986).

d) *Final Comments*

The data show that the emitting region is inhomogeneous, with a small inner region of high density and magnetic field (the optically emitting region) that eventually becomes a large, partially opaque radio emitting region with a low density and magnetic field. The X-ray emission appears to be produced by Compton scattering of the infrared-submillimeter photons. The synchrotron–self-Compton model was fit to the multifrequency spectra to determine the plasma parameters near $\nu_m \sim 10^{11.5}$ Hz. Hughes, Aller, and Aller (1989*b*) fit a propagating shock model to the radio variations at cm wavelengths (10^{10} Hz) to determine the physical conditions in this region. Because each of these approaches is limited in scope, we point out the need for a more general model that can explain the full assortment of observations presented here: the overall spectral shape as well as the time variability data (i.e., the structure function, the shape of outbursts, the frequency-dependent time delays).

One of the most important results of this work is that the two simple predictions of the synchrotron–self-Compton model are reproduced in the multifrequency spectra. Since the variability data supporting this conclusion rests on two such spectra (taken with two different instruments), there is concern that this was simply a chance coincidence. Consequently, we are obtaining additional multifrequency data which include X-ray data with greatly improved spectral information (from the GINGA observatory) and will report upon the results in a future work.

We would like to thank Rick Howard for using test time to make millimeter measurements of BL Lac and Anne Kinney for her assistance with *IUE* data reduction. *IUE* observing at NYU was supported by NASA grant NAG 5-73; J. N. B. would like to acknowledge support from grant NAG-5-1059; radio astronomy at the University of Michigan is supported by the NSF through grant AST-8501093 and AST-8815678; J. S. M. and S. S. wish to acknowledge Lick Observatory and the NSF for their support; W. A. D., and W. K. have been aided by the NSF and the University of Massachusetts; infrared astronomy at Caltech gratefully acknowledges continued support by the NSF; funding for *IRAS* observations and data reduction were provided by ESA, SERC, and NASA; W. Z. W. would like to thank Steward observatory for their support. The long-term University of Florida observations have been supported by a series of grants from the NSF; the current grant is AST-8516269.

REFERENCES

Aller, H. D., Aller, M. F., and Hughes, P. A. 1985, *Ap. J.*, **298**, 296.
———. 1988, in *The Impact of VLBI on Astrophysics and Geophysics*, ed. M. J. Reid and J. M. Moran (Dordrecht: Reidel), p. 83.
Aller, H. D., Aller, M. F., Latimer, G. E., and Hodge, P. E. 1985, *Ap. J. Suppl.*, **59**, 513.
Baath, L. B., Elgered, G., Lundqvist, G., Graham, D., Weiler, K. W., Seielstad, G. A., Tallqvist, S., and Schilizzi, R. T. 1981, *Astr. Ap.*, **96**, 316.
Balonek, T. J. 1982, Ph.D. thesis, University of Massachusetts.
Beichman, C. A., Neugebauer, G., Soifer, B. T., Wootten, H. A., Roellig, T., and Harvey, P. M. 1981*a*, *Nature*, **293**, 711.
Beichman, C. A., Provdo, S. H., Neugebauer, G., Soifer, B. T., Matthews, K., and Wootten, H. A. 1981*b*, *Ap. J.*, **247**, 780.
Biermann, P. L., and Strittmatter, P. A. 1987, *Ap. J.*, **322**, 643.
Bjornsson, C.-I. 1986, in *Continuum Emission in Active Galactic Nuclei*, ed. M. L. Sitko (Tucson: NOAO), p. 181.
Blandford, R. D., and Eicher, D. 1987, *Phys. Rept.*, **154**, 1.
Bregman, J. N. 1985, *Ap. J.*, **288**, 32.
Bregman, J. N., Lebofsky, M. J., Aller, M. F., Rieke, G. H., Aller, H. D., Hodge, P. E., Glassgold, A. E., and Huggins, P. J. 1981, *Nature*, **293**, 714.
Bregman, J. N., *et al.* 1982, *Ap. J.*, **253**, 19.

———. 1984, *Ap. J.*, **276**, 454.
———. 1986, *Ap. J.*, **301**, 708.
———. 1988, *Ap. J.*, **331**, 746.
Brindle, C., *et al.* 1985, *M.N.R.A.S.*, **214**, 619.
Burstein, D., and Heiles, C. 1978, *Ap. J.*, **225**, 40.
Du Puy, D., Schmitt, J. L., McClure, R., van den Bergh, S., and Racine, R. 1969, *Ap. J. (Letters)*, **156**, L35.
Epstein, E. E., Fogarty, W. G., Mottmann, J., and Schneider, E. 1982, *A.J.*, **87**, 449.
Fabian, A. C. 1988, in *Supermassive Black Holes*, ed. M. Kafatos (Cambridge: Cambridge University Press), p. 234.
Fielder, R. L., *et al.* 1987, *Ap. J. Suppl.*, **65**, 319.
Fritz, K.-D. 1988, *Bull. AAS*, **20**, 1064.
———. 1989, *Astr. Ap.*, in press.
Ghisellini, G., Maraschi, L., and Treves, A. 1985, *Astr. Ap.*, **146**, 204.
Glassgold, A. E., *et al.* 1983, *Ap. J.*, **274**, 101.
Hagen-Thorn, V. A., Marchenko, S. G., and Yacovleva, V. A. 1985, *Astrofizika*, **22**, 5.
Hagen-Thorn, V. A., and Yacovleva, V. A. 1987, in *IAU Symposium 121, Observational Evidence of Activity in Galaxies*, ed. E. Ye. Khachikian, G. Melnick, and K. Fricke (Dordrecht: Reidel), p. 325.
Heavens, A. F., and Meisenheimer, K. 1987, *M.N.R.A.S.*, **225**, 335.
Hughes, P. A., Aller, H. D., and Aller, M. F. 1985, *Ap. J.*, **298**, 301.
———. 1988, in *IAU Symposium 129, The Impact of VLBI on Astrophysics and Geophysics*, ed. M. J. Reid and J. M. Moran (Dordrecht: Reidel), p. 81.
———. 1989*a*, *Ap. J.*, **341**, 54.
———. 1989*b*, *Ap. J.*, **341**, 68.
Impey, C. D., Brand, P. W. J. L., Wolstencroft, R. D., and Williams, P. M. 1984, *M.N.R.A.S.*, **209**, 245.
Impey, C. D., and Neugebauer, G. 1988, *A.J.*, **95**, 309.
Johnston, K. J., *et al.* 1984, *Ap. J. (Letters)*, **277**, L31.
Jones, T. W., O'Dell, S. L., and Stein, W. A. 1974, *Ap. J.*, **188**, 353.
Kondo, Y., *et al.* 1981, *Ap. J.*, **243**, 690.
Landau, R., *et al.* 1986, *Ap. J.*, **308**, 78.
Maraschi, L., Schwartz, D. A., Tanzi, E. G., and Treves, A. 1985, *Ap. J.*, **294**, 615.
Makino, F., *et al.* 1987, *Ap. J.*, **313**, 662.
Marscher, A. P. 1977, *Ap. J.*, **216**, 244.
Marscher, A. P., and Gear, W. K. 1985, *Ap. J.*, **298**, 114.
Miller, J. S., French, H. B., and Hawley, S. A. 1978, *Ap. J. (Letters)*, **219**, L85.
Moore, R. L., *et al.* 1982, *Ap. J.*, **260**, 415.
Moore, R. L., Schmidt, G. D., and West, S. C. 1987, *Ap. J.*, **314**, 176.
Mufson, S. L., *et al.* 1984, *Ap. J.*, **285**, 571.
Mutel, R. L., and Phillips, R. B. 1987, in *Superluminal Radio Sources*, ed. J. A. Zensus and T. J. Pearson (Cambridge: Cambridge University Press), p. 60.
Neugebauer, G., Soifer, B. T., Matthews, K., and Elias, J. H. 1989, *A.J.*, **97**, 957.
Ohashi, T. 1988, in *Physics of Neutron Stars and Black Holes*, ed. Y. Tanaka (Tokyo: Universal Academy Press), p. 301.
Phillips, R. B., and Mutel, R. L. 1982, *Ap. J. (Letters)*, **257**, L19.
———. 1988, in *IAU Symposium 129, The Impact of VLBI on Astrophysics and Geophysics*, ed. M. J. Reid and J. M. Moran (Dordrecht: Reidel), p. 103.
Pica, A. J., Pollock, J. T., Smith, A. G., Leacock, R. J., Edwards, P. L., and Scott, R. L. 1980, *A.J.*, **85**, 1442.
Pollock, J. T. 1983, Ph.D. thesis, University of Florida.
Rieke, G. H., Lebofsky, M. J., and Kinman, T. D. 1979, *Ap. J. (Letters)*, **232**, L151.
Rieke, G. H., Lebofsky, M. J., and Wisniewski, W. Z. 1982, *Ap. J.*, **263**, 73.
Seaton, M. J. 1980, *M.N.R.A.S.*, **187**, 73.
Simonetti, J. H., Cordes, J. M., and Heeschen, D. S. 1985, *Ap. J.*, **296**, 46.
Sitko, M. L., Schmidt, G. D., and Stein, W. A. 1985, *Ap. J. Suppl.*, **59**, 323.
Smith, A. G., Leacock, R. J., and Webb, J. R. 1987, in *Active Galactic Nuclei*, ed. H. Richard Miller and Paul J. Wiita (Berlin: Springer), p. 158.
Smith, P. S. 1986, Ph.D. thesis, University of New Mexico.
Stannard, D., and McIlwrath, B. K. 1982, *Nature*, **298**, 140.
Thaun, T. X., Oke, J. B., and Gunn, J. E. 1975, *Ap. J.*, **201**, 45.
Wardle, J. F. C., and Roberts, D. H. 1988, in *IAU Symposium 124, The Impact of VLBI on Astrophysics and Geophysics*, ed. M. J. Reid and J. M. Moran (Dordrecht: Reidel), p. 143.
Webb, G. M., Drury, L. O'C., and Biermann, P. 1984, *Astr. Ap.*, **137**, 185.
Webb, J. R., Smith, A. G., Leacock, R. J., Fitzgibbons, G. L., Gombola, P. P., and Shepherd, D. W. 1988, *A.J.*, **95**, 374.
Worrall, D. M., *et al.* 1982, *Ap. J.*, **261**, 403.

H. D. ALLER, M. F. ALLER, and P. A. HUGHES: Department of Astronomy, University of Michigan, 953 Physics-Astronomy Building, Ann Arbor, MI 48109

T. J. BALONEK: Department of Physics and Astronomy, Colgate University, Hamilton, NY 11346

P. W. J. L. BRAND: Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, Scotland

J. N. BREGMAN: Department of Astronomy, University of Michigan, Dennison Building, Ann Arbor, MI 48109-1090

W. A. DENT and W. KINSEL: Department of Physics and Astronomy, GR Tower B, University of Massachusetts, Amherst, MA 01003

J. H. ELIAS: Cerro Tololo Inter-American Observatory, P.O. Box 26732, Tucson, AZ 85726-6732

B. GARILLI and D. MACCAGNI: Istituto di Fisica Cosmica, CNR, via Bassini 15, I-20133 Milano, Italy

P. GIOMMI: *EXOSAT* Observatory, ESTEC, Space Science Department, 2200 AG Noordwijk, The Netherlands

A. E. GLASSGOLD and P. J. HUGGINS: Physics Department, New York University, 4 Washington Place, New York, NY 10003

W. H.-M. KU: Columbia Astrophysics Laboratory, Physics Department, Columbia University, New York, NY 10024

R. J. LEACOCK and A. G. SMITH: Department of Astronomy, University of Florida, Gainesville, FL 32611

K. MATTHEWS, G. NEUGEBAUER, and B. T. SOIFER: Division of Physics, Mathematics and Astronomy, Caltech, Pasadena, CA 91125

J. S. MILLER and S. STEPHENS: Lick Observatory, University of California, Santa Cruz, CA 95064

J. T. POLLOCK: Department of Physics and Astronomy, Appalachian State University, Boone, NC 28608

J. R. WEBB: Department of Physics and Astronomy, Stephen F. Austin State University, P.O. Box 13044, SFA Station, Nacogdoches, TX 75962-3044

P. M. WILLIAMS: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

W. Z. WISNIEWSKI: Steward Observatory, University of Arizona, Space Sciences Building, Tucson, AZ 85721

13

IR photometry of flat spectrum radio sources

C. D. Impey & P. W. J. L. Brand

Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 2HJ, UK

IR measurements of optically faint, flat spectrum radio sources have implied the possible existence of a new class of quasar, with IR excess more extreme than that of any extragalactic object. In the present preliminary study, 18 Parkes flat spectrum sources have been observed in the IR without regard to optical morphology, and the red sources are found to be just the tail of the normal distribution of quasar colours. The results imply that a considerable fraction of flat spectrum sources from complete samples may have properties similar to the BL Lac objects.

Rieke *et al.*[1,2] recently detected several objects at 2.2 μm in optically blank fields of flat spectrum radio sources[3,4], and derived such large IR/optical spectral indices ($\alpha = -$d ln flux density/d ln frequency) $\alpha_{KB} \approx 3$ that they proposed a new class of quasar, because other properties of the objects indicated that they were similar to violently variable quasars, or BL Lac objects. They further noted that at the largest (faintest) blue magnitudes, α_{KB} tended to be high, while at $m_B \lesssim 18$, $\alpha_{KB} \approx 1$, a value typical for the optically bright quasars of Neugebauer *et al*[5]. The spectral index α_{KB} was calculated from measurements in the optical B band and the near-IR K band. To test the trend of α with m_B, and to discover whether the red sources represent a new class of quasar, a sample of flat spectrum radio sources was observed in the IR without regard to optical identification. The candidates were selected from the list of 2.7-GHz sources of Condon *et al.*[3] according to limits on spectral index ($\alpha <$ 0.5, 2.7 − 5.0 GHz), flux density ($S_{2.7} > 0.5$ Jy), RA (8 h $< \alpha <$ 15 h) and dec ($-30° < \delta < -4°$).

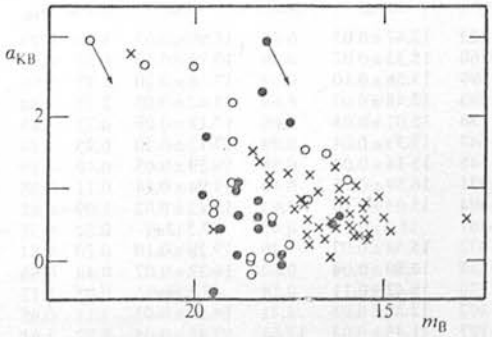


Fig. 1 Optical/IR spectral index versus optical magnitude, showing Rieke *et al.*'s objects[1,2] (○), the optically bright quasars of Neugebauer *et al.*[5] (×) and objects from the present study (●). Trajectories show the effect of 1 mag brightening at B between epoch of Palomar plates and IR measurements.

0028-0836/81/350814—03$01.00

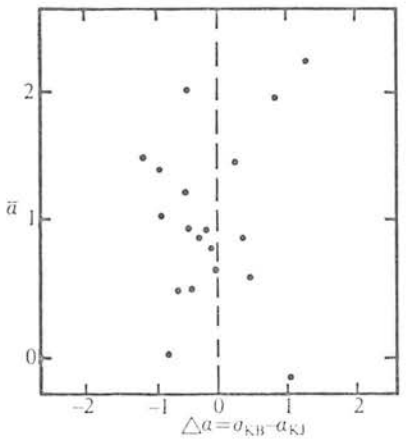


Fig. 2 Plot of mean optical/IR spectral index against the residual $\alpha_{KB}-\alpha_{KJ}$. The deviation from the line $\alpha_{KB}=\alpha_{KJ}$ is a measure of the curvature of the spectrum (either intrinsic or due to variability).

The IR observations were obtained with the Anglo–Australian 3.9-m telescope on 2–3 May 1980 using theInSb Infrared Photometry System. All the observations were made with the broad band K(2.2 μm) and J(1.25 μm) filters. Standard stars from the AAO list were observed to calibrate the photometry. The data have been corrected for near-IR extinction at Siding Spring according to recently determined coefficients (G. Gilmore and I. N. Read, personal communication). The data have also been corrected for the effect of different flux distributions on the broad band flux measurements. Only two of the sources lie at galactic latitude $|b|<30°$, and they were corrected for galactic extinction. The minimum error in the J and K magnitudes was determined from the repeatability of standard observations during a night. In two nights, 21 sources were observed, with 20 detected at K and 18 detected at J.

Table 1 shows the IR data, along with spectral indices calculated between J and K (α_{KJ}) and between the optical B band and K(α_{KB}). There are two major sources of error in these spectral indices; one is the effect of redshifted emission lines passing through the photometric bands. α_{KJ} may be affected, but over the long baseline from 2.2 μm to the visible, α_{KB} is relatively insensitive to this effect[6]. The second is the fact that the optical magnitude is taken from Palomar plates, and variability might lead to errors in α_{KB}. The photographic fluxes have an uncertainty of ±0.5 mag (ref. 3).

Figure 1 shows α_{KB} plotted against m_B, with the cluster of red, optically faint objects showing up clearly in the previous data. The 18 sources in our sample present a distribution of spectral indices covering the full range of α from 0 to 3. Very red sources are relatively common; four out of 18 lie in the top part of the range of spectral indices for BL Lac objects and two are as red as any of those found by Rieke and Lebofsky[1]. The data fill in the distribution of very red objects at brighter optical magnitudes, and shows that the trend of α_{KB} with m_B is a selection effect.

Figure 1 shows that much of the argument rests on the α_{KB} indices of the two reddest objects (1335−127 and 1034−293), which may be affected by variability. However, there are reasons to believe that these points are not spurious. First, direct observations on the AAT TV acquisition system confirmed that neither of the objects had varied by more than a magnitude from the epoch of the Palomar plates. The effect of 1 mag variability in B is shown as a trajectory in Fig. 1 both for 1335−127 from the present study and for the reddest object from refs 1, 2. While anomalous brightness at B would bring the two red objects down towards the bulk of the distribution, the same is true of the original red objects. As both sets of α_{KB} were calculated in the same way, the results are directly comparable. Second, there is no evidence from the cotemporaneous J and K measurements that these objects are abnormal. Figure 2 shows the mean spectral slope $\bar{\alpha}$ plotted against the difference $(\alpha_{KB}-\alpha_{KJ})$, where the steep spectrum sources lie at the top of the graph. Rieke *et al.*[1,2] noted that a roughly constant spectral index satisfied both their J, K and optical fluxes; this is represented by the line $\alpha_{KB}=\alpha_{KJ}$ in Fig. 2. Figure 2 also shows that the objects with large α_{KB} do not have abnormally large residuals $\Delta\alpha$, and are not segregated from the main body of the distribution. While there is not a 1:1 correspondence between steep α_{KB} and steep α_{KJ}, the four reddest objects in α_{KB} are among the eight reddest objects in α_{KJ}. The means and dispersions of both spectral index distributions are similar ($\bar{\alpha}_{JK}=1.05\pm0.72$, $\bar{\alpha}_{KB}=0.92\pm0.79$). There is no evidence from either the optical or IR data that the red (α_{KB}) objects are other than the tail of the normal distribution of quasar colours. Objects which resemble BL Lac objects in their IR properties are relatively common. The four reddest sources are particularly interesting because (unlike the original 'red' sources) they are bright enough for follow-up spectroscopy and polarimetry. The reddest object (1335−127) has recently been confirmed as a BL Lac object. Our 2.2-μm measurements with the UK Infrared Telescope show it to have varied by a magnitude over one year, and radio observations by Aller *et al.*[7] show it to have variable linear polarization at frequencies from 4.8 to 14.5 GHz. In addition, the second reddest object (1034−293) shows evidence of variability ($\Delta m_B>1$ mag) from UK Schmidt plates of the area, another indication that it is a BL Lac-type object.

Even for this relatively small sample it is possible to draw conclusions on the IR properties of flat spectrum radio sources,

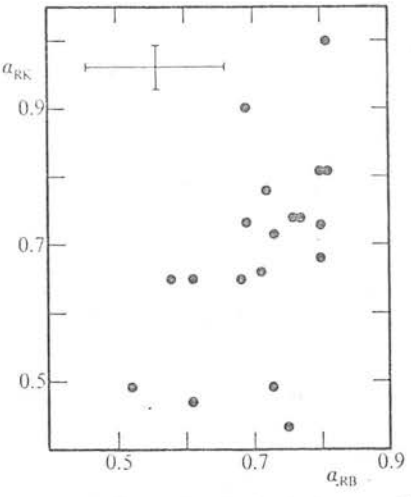


Fig. 3 Plot of radio/IR and radio/optical spectral indices. The dispersions in α_{RK} and α_{RB} are a measure of the relative IR and optical flux variations of the sample. Typical error bars are shown in the top left corner.

Table 1 IR data and spectral indices between the J and K, and B and K bands

Source	K(2.2 μm) (mag)	Flux (mJy)	J(1.25 μm) (mag)	Flux (mJy)	α_{JK}	α_{KB}	M_B
0823−223	12.47±0.03	6.79	14.90±0.03	1.76	2.25	1.5	17.5
0919−260	15.33±0.07	0.49	16.73±0.10	0.33	0.68	1.0	19.0
1032−199	15.58±0.10	0.39	17.16±0.10	0.22	0.96	0.8	19.0
1034−293	12.48±0.03	6.69	14.42±0.03	2.75	1.48	2.4	18.5
1124−186	15.01±0.04	0.66	17.17±0.09	0.22	1.85	0.9	18.5
1128−047	15.33±0.04	0.49	17.13±0.10	0.23	1.29	1.6	20.0
1143−245	15.14±0.04	0.58	16.29±0.05	0.49	0.29	0.8	18.5
1145−071	16.59±0.12	0.16	17.94±0.14	0.11	0.68	0.2	18.5
1156−094	15.03±0.10	0.65	15.42±0.02	1.09	−0.85	0.3	17.5
1203−261	18.3 (2σ)	0.03	19.5 (2σ)	0.02	0.32	−0.5	19.5
1243−072	15.34±0.07	0.49	17.28±0.10	0.20	1.51	1.0	19.0
1244−255	14.80±0.04	0.80	16.32±0.07	0.48	0.86	0.7	18.0
1256−220	16.42±0.11	0.18	18.6 (4σ)	0.05	2.12	0.9	20.0
1302−102	12.88±0.03	4.71	14.26±0.03	3.17	0.66	0.6	16.0
1335−127	11.43±0.03	17.64	13.45±0.03	6.72	1.61	3.0	18.5
1352−104	15.36±0.03	0.49	16.77±0.10	0.31	0.75	0.1	17.5
1354−152	15.23±0.08	0.54	16.86±0.08	0.29	1.03	0.7	18.5
1406−076	16.37±0.11	0.19	18.28±0.16	0.08	1.46	0.6	19.5
1511−100	>18.8	<0.02	—	—	—	—	18.5
1519−273	17.9 (3σ)	0.05	—	—	—	—	18.5

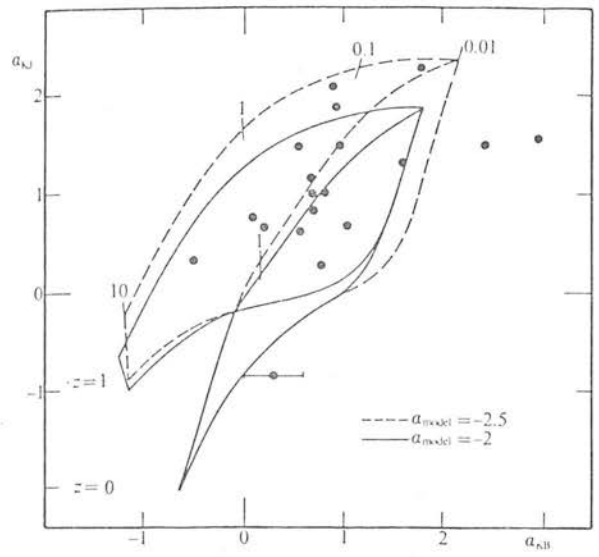


Fig. 4 Plot of near IR and optical/IR spectral indices with permitted zones from a two-component model. The error bars on the lowest point show the effect of a variation of $\pm\frac{1}{2}$ mag in B and typical 5% precision on the J and K flux measurements.

and compare them with the optical properties. Selection effects will not be important because the radio sample was completely identified optically, and with one exception was completely identified at 2.2 μm. Figure 3 shows α_{RK} plotted against α_{RB}, where the spectral index is calculated between the 2.7 GHz measurement and the K and B data points. The range of radio fluxes is less than a factor of 3, while the range of IR fluxes is >500, thus the graph allows a comparison of the range of IR and optical fluxes in the sample. The mean spectral indices are $\bar{\alpha}_{RK}=0.69$ and $\bar{\alpha}_{RB}=0.71$, both very near to the canonical value for optically thin non-thermal radio sources. There is no significant difference between the range of IR and optical fluxes ($\sigma(\alpha_{RK})=0.15$, $\sigma(\alpha_{RB})=0.08$ and taking the relative standard error in the variance of n variates to be 1.4 $(n-1)^{-1/2}$) and the even distribution of the very red objects on the α_{RB} axis is a confirmation of the fact that they do not tend to be optically faint. The objects of Rieke *et al.*[2] do not come from a separate population at better than the 1.5σ level; removing the lower four objects only reduces $\sigma(\alpha_{RK})$ by 0.03 to 0.12.

To see if a simple picture accounts for the range of optical/IR spectra, the energy distribution of a giant elliptical[8] was added to a power law spectrum of slope α which flattened off ($\alpha=0$) at some wavelength λ_n (the knee). Figure 4 shows the data plotted on a two-colour diagram of α_{KB}/α_{KJ}. The extreme red objects lie at large values of α_{KB}. As the knee wavelength travels across the spectrum, the source traces a trajectory in the two-colour plane. As the proportion of non-thermal to thermal components is varied, the trajectory sweeps out an area on the diagram. The right-hand boundary represents the extreme non-thermal case of $S_{GAL}/S_{NT}=0.01$, where S_{GAL} is the peak flux of the elliptical galaxy and S_{NT} is the flux at the point where the power law flattens. As the flux contribution of the elliptical increases, the trajectory sweeps across to smaller values of α_{KB} and α_{KJ}, and the left-hand boundary of the enclosed area shows the case where $S_{GAL}/S_{NT}=10$ and the spectrum is dominated by stellar flux. The redshift distribution of flat spectrum radio samples is expected to peak around $z\sim1$ (ref. 9), and Fig. 4 shows the case for $z=0$ and $z=1$. Given the luminosity of a giant elliptical, the flat spectrum sources would be expected to populate the non-thermal ($S_{GAL}/S_{NT}<1$) portion of the $z=1$ area and the thermal ($S_{GAL}/S_{NT}>1$) part of the $z=0$ area. Finally the areas have been plotted for two values of α; $\alpha=2.0$ and $\alpha=2.5$ are the solid and dashed curves respectively.

Most of the data points lie well within the area described by the model; the $\alpha=2.5$ case covering 16 of the 19 sources. Some of the scatter of the data in α_{KJ} will be caused by emission lines (especially Hα passing through the J band) and some of the scatter in α_{KB} may be caused by variability and the non-simultaneity of the optical and IR flux measurements. However, the agreement is surprisingly good given the simple assumptions. A range of α from 2 to 3 is sufficient to cover even the most extreme sources. The assumption that $\alpha=0$ for $\lambda>\lambda_n$ is not critical. Observations at longer wavelength will help confirm this picture.

We thank the AAO staff for helpful support. C.D.I. acknowledges the receipt of an SRC studentship.

Received 17 March; accepted 30 June 1981.

1. Rieke, G. H. & Lebofsky, M. J. *IAU Symp.* **92**, 263–267 (1980).
2. Rieke, G. H., Lebofsky, M. J. & Kinman, T. D. *Astrophys. J. Lett.* **232**, L151–154 (1979).
3. Condon, J. J., Hicks, P. D. & Jauncey, D. L. *Astr. J.* **82**, 692–700 (1977).
4. Condon, J. J., Jauncey, D. L. & Wright, A. E. *Astr. J.* **83**, 1036–1046 (1978).
5. Neugebauer, G., Oke, J. B., Becklin, E. E. & Matthews, K. *Astrophys. J.* **230**, 79–94 (1979).
6. Hyland, A. R. & Schwartz, M. P. *Proc. astr. Soc. Aust.* **3**, 137–140 (1977).
7. Aller, H. D., Aller, M. F. & Hodge, P. E. *Astr. J.* **86**, 325 (1981).
8. Grasdalen, G. L. *IAU Symp.* **92**, 269–276 (1980).
9. Owen, F. N. & Mufson, S. L. *Astr. J.* **82**, 776–780 (1977).

0028-0836/81/350816—03$01.00

Mon. Not. R. astr. Soc. (1982) 201, 849–866

The calibration of a radio-independent search for BL Lac objects

C. D. Impey* and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

Received 1982 March 24; in original form 1981 July 20

Summary. The first results from a radio-independent search for BL Lac objects using a linear polarizing filter are presented. IIIa-J plates taken on the 1.2-m UK Schmidt telescope were measured with the COSMOS machine and ~ 33 000 objects in a 10 square degree area were searched down to a level of $B \sim 19.5$. The photometry was calibrated with a photoelectric sequence and a photographic extension using the UKST sub-beam prism. The photographic polarimetry was performed differentially and this technique was optimized with a local magnitude transformation between small areas of the plate. The COSMOS image parameters were used to define an algorithm to separate stars and galaxies and to exclude merged and unmatched images. Objects were selected from the plate material according to polarization and variability and a detection limit of $p \sim 15$ per cent was set. In terms of the known polarization properties of BL Lac objects, the search is estimated to be ~ 15 per cent complete. The inferred limits on the surface density of BL Lac objects are $0.9^{+0.3}_{-0.9}$ square degree^{-1} down to $B = 19.0$.

1 Introduction

A linear polarizing filter has been commissioned at the 1.2-m UK Schmidt telescope in Australia in order to search for BL Lac objects. The current sample of BL Lac objects has been discovered almost exclusively from radio source identifications (Angel & Stockman 1980). Techniques involving slitless spectroscopy and searches for UV excesses are not appropriate for BL Lacs, so most have been found from accurate radio positions of flat spectrum sources. At the moment it is uncertain whether BL Lac objects have analogues to the radio-quiet quasars, which greatly outnumber radio-emitting quasars (Murdoch & Crawford 1977; Katgert *et al.* 1973). Despite the selection effects, it seems likely that radio-quiet BL Lac are very rare. The few BL Lac objects which have been selected by optical and X-ray methods are also radio sources: 1Zw 187 (Altschuler & Wardle 1975), Mk 501 (Khachakian & Weedman 1974), 2155 – 304 (Agrawal & Reigler 1979) and 1218 + 304 (Cooke *et al.* 1978). The only highly polarized QSO which is radio-quiet is PHL 5200 and in this object the polarized continuum is unlikely to be synchrotron emission (Stockman,

* Present address: Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA.

Angel & Hier 1980). Current models predict that BL Lac objects are intrinsically strong radio-emitters. Blandford & Rees (1978) have proposed that BL Lac objects are active nuclei in which the continuum radiation is beamed towards the observer. Whether a source is a QSO, optically violently variable quasar, or BL Lac object depends precisely on the line-of-sight with respect to the jet axis. To discriminate between current models of active nuclei it is important to have a radio-independent selection technique for BL Lac objects. The polarization filter was designed to select them by the high linear polarization which is a feature of the continuum flux. Levels of up to 43 per cent have been observed (Impey, Brand & Tapia 1981). This is the first attempt to find BL Lac objects over a large area of sky in a way that is independent of radio flux.

2 Technique

The design and construction of the rotatable polarizing filter has been described by Wolstencroft, Impey & Smith (1981). Laboratory tests on the filter showed that the thickness is 0.68 cm with a tolerance of $< 55\,\mu$m which corresponds to a maximum defocusing across the plate of ~ 0.15 arcsec. The transmission variations were mapped and found to be 4 per cent with a spatial frequency of $0.25–0.3\,\mathrm{cm}^{-1}$. Measurements show that the filter has a polarizance to white light of 2.9×10^4 and a dispersion in the fast axis angle of $< 0.05°$. For the purposes of this project the filter was assumed to be a perfect polarimeter.

Four photographic plates were exposed consecutively, each exposure through a different position angle of the polaroid ($\phi = 0°$, $45°$, $90°$ and $135°$). The alternative of multiple exposures on a single plate was dismissed because of the contamination due to overlapping images. There will be ~ 3 per cent double images from a four exposure plate down to $B \sim 21$ in a medium galactic latitude field (Allen 1973), but the situation will in practice be much worse because of latent images. The continuum of development probabilities for photographic grains means that images considerably fainter than the single exposure limit will be developed in a multiple exposure. Each exposure includes a control region which is an area of plate exposed through a non-rotating piece of polaroid. Images in the control region are used to calibrate out differences in exposure time, emulsion sensitivity and atmospheric transmission between plates. The calibration of the filter using an object of known polarization is discussed by Wolstencroft *et al.* (1981) and the data reduction procedure is covered in Appendix A. A great advantage of the polarimetric method is that objects are detected differentially.

3 Photometric calibration

3.1 PHOTOELECTRIC SEQUENCE

The first field chosen for the photographic survey was centred on $12^{\mathrm{h}}\,58^{\mathrm{m}}$, $-20°\,09'$ (1950.0). The requirement of four consecutive long exposures constrained the field to be in the south, otherwise there would have been trailed images due to atmospheric dispersion on the unfiltered IIIa-J plates. In order to calibrate the photographic material, a photoelectric sequence was established on the 1-m telescope at the South African Astronomical Observatory, Sutherland. Photometry in the Johnson B and V bands was obtained using the St Andrews photometer with a cooled RCA C31034 A pm tube and a 14-arcsec aperture. Standard colour equations and extinction corrections were applied to the data and standard stars from the Cousins 'E' regions were observed (Menzies *et al.* 1980). Thirty-four stars in a one-degree square at the centre of the field were measured, the stars covering a range of

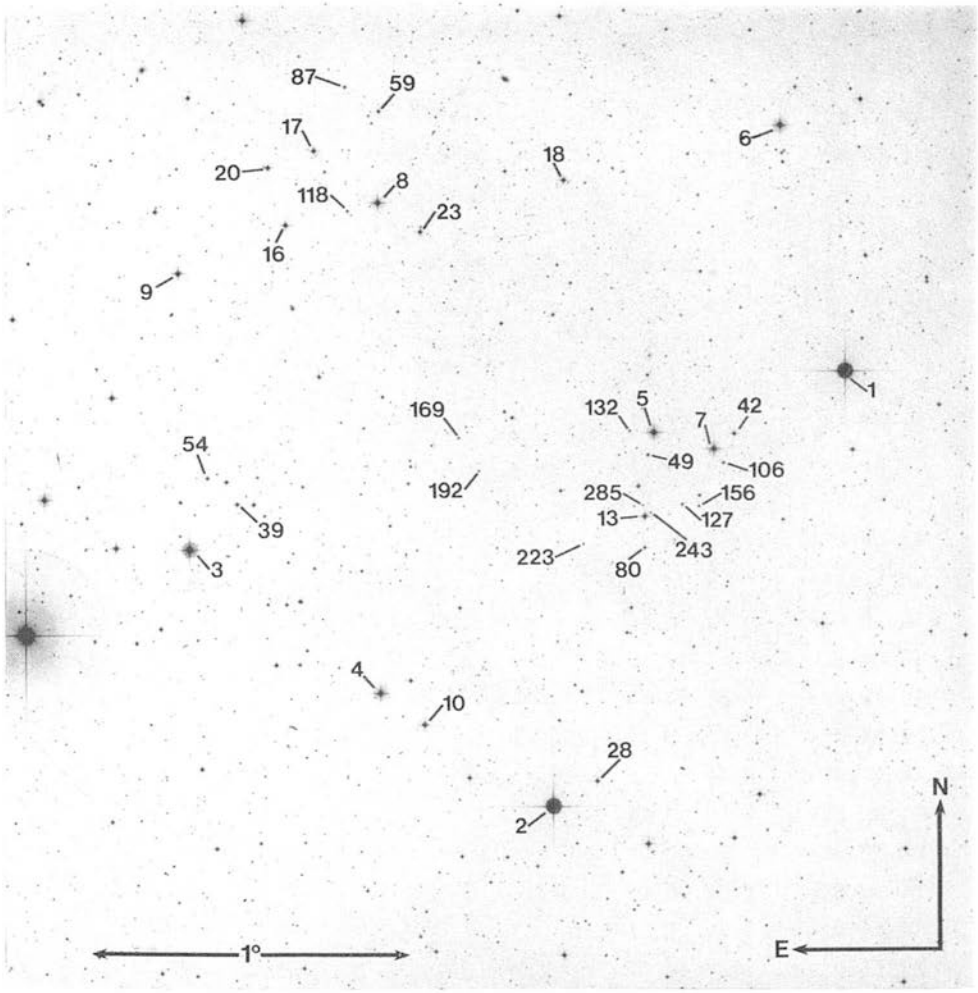


Plate 1. Finding charts for the photoelectric sequence in the field centred on $12^h\ 57^m$, $-\ 20^\circ\ 09'$. The stars are numbered according to their brightness in the 1 degree field; Johnson B and V magnitudes are listed in Table 1. Photograph courtesy of Photolabs, Royal Observatory Edinburgh.

[*facing page* 850]

$7.7 < B < 15.9$ and $0.4 < B-V < 1.7$. Twenty-two stars were observed on two occasions and the faintest was observed on three separate nights. The consistency of each repeated measurement was < 0.02 mag. All photoelectric measurements and positions accurate to one arcsec are listed in Table 1 and the finding chart for the sequence stars is shown in Plate 1.

Table 1. Photoelectric sequence.

No.	V	B-V	Comments	α (1950.0)	δ (1950.0)
2	6.92	1.31	SAO 181319	$12^h58^m33^s.4$	-20°31'31"
1	7.23	0.50	SAO 181297	12 57 08.8	-20 01 19
3	8.65	0.52	SAO 181342	13 00 17.7	-20 13 15
4	9.21	1.08	SAO 181330	12 59 23.0	-20 22 52
6	9.37	0.94	SAO 157645	12 57 27.4	-19 44 47
5	9.73	0.49	SAO 181311	12 58 04.0	-20 05 30
8	9.98	0.54	SAO 157666	12 59 23.5	-19 50 02
10	10.01	1.66		12 59 10.4	-20 25 00
7	10.03	0.42	SAO 181307	12 57 46.5	-20 06 35
8	10.31	0.53	SAO 157679	13 00 20.2	-19 54 43
13	10.31	1.53		12 58 06.5	-20 11 07
17	10.71	1.21		12 59 41.7	-19 46 33
16	10.91	0.72		12 59 49.8	-19 51 32
20	11.24	0.65		12 59 54.8	-19 47 40
18	11.28	0.77		12 58 30.0	-19 48 31
23	11.55	0.63	Possible variable	12 59 36.6	-19 51 47
39	11.64	1.08		13 00 03.8	-20 10 14
42	11.71	0.99		12 57 40.6	-20 05 33
28	11.83	0.56		12 58 20.6	-20 28 49
49	12.23	0.65		12 58 05.7	-20 07 00
54	12.34	0.65		13 00 12.1	-20 08 28
59	12.55	0.74		12 59 23.1	-19 43 55
80	12.94	0.65		12 58 06.6	-20 13 13
87	13.15	0.55		12 59 32.7	-19 42 16
106	13.18	1.00		12 57 43.7	-20 07 31
118	13.45	0.95		12 59 32.2	-19 50 35
127	13.67	0.85		12 57 55.7	-20 10 17
132	13.90	0.87		12 58 10.9	-20 05 22
156	14.10	0.51		12 57 50.8	-20 10 26
169	14.17	0.76		12 59 00.4	-20 05 50
192	14.54	0.72		12 58 54.7	-20 08 02
243	14.72	0.96		12 58 04.9	-20 10 53
223	14.86	0.62	± 0.02	12 58 24.4	-20 12 57
285	15.28	0.67	± 0.02	12 58 07.1	-20 10 22

Table 2. Details of the sub-beam prism plates.

No.	Telescope	Field centre	Exposure	Seeing	Emulsion	Filter	Machine
B 5895	UKST	$12^h\ 58^m$ -20° 09′	20 min.	2″	II aO	GG385	COSMOS/PDS
B 5896	UKST		20	3	II aO	GG385	COSMOS/PDS
B 5897	UKST		20	3	II aO	GG385	COSMOS/PDS

3.2 SUB-BEAM PRISM

Since the Schmidt plate material extends considerably fainter than the photoelectric sequence, a sub-beam prism was employed (Pickering 1891). The UKSTU prism has a diameter of 250 mm and an apex angle of ~ 50 arcsec. A prism constant of $\Delta = 3.3$ mag is quoted in the USKTU Handbook, although as Couch & Newell (1980) point out, the effective prism constant depends on the measuring machine used.

Three consecutive 20-min exposures were taken with the sub-beam prism in seeing of 2–3 arcsec. Table 2 shows the details of the plate material. The three plates were measured on the PDS machine at the Royal Greenwich Observatory. A two-dimensional Gaussian was fitted to each image and the density volume used as a machine magnitude estimator. There is no colour correction because the filter–emulsion combination of the sub-beam prism plates is very close to the Johnson B passband. In Fig. 1, the mean PDS magnitudes for the three plates are plotted against photoelectric B measures for the primary and secondary images of the sequence stars. The photographic extension by this method is essentially an extrapolation, the only verification of coming from the magnitude range where primary and secondary images overlap. From Fig. 1 there are 14 primary and nine secondary images in the region of overlap covering $\Delta B = 2.5$ mag. Straight lines fitted to the overlap region

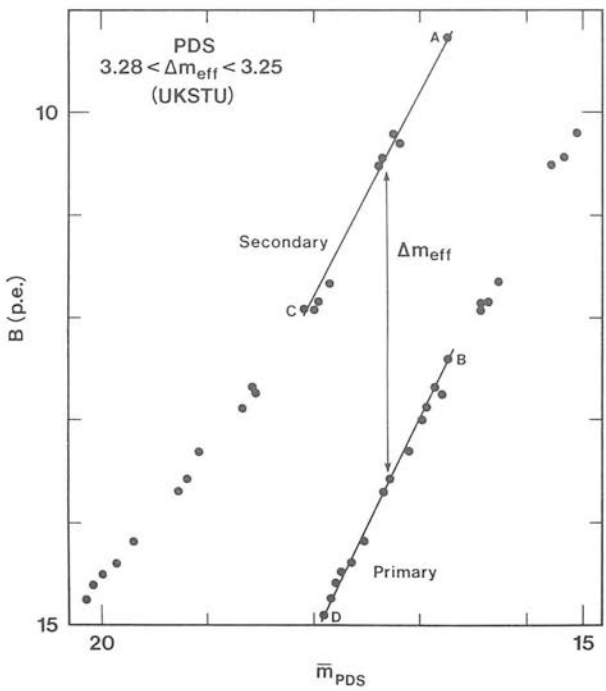


Figure 1. Sub-beam prism calibration (PDS). Johnson B magnitudes are plotted against PDS magnitude for the mean of three plates. A linear regression straight line is fitted to the stars in the overlap region between primary and secondary images.

stars show that Δ changes only from 3.25 at $B = 12.4$ to 3.28 at $B = 15.0$; that is, the prism constant is indeed constant with magnitude within the errors of the fitted lines.

Unfortunately, the overlap region cannot be extended beyond 2.5 mag. At bright end the secondary images become contaminated by the wings of bright saturated primaries and at the faint end the secondaries are close to the plate limit. Furthermore, Fig. 1 shows that the secondary calibration curve bends just at the point where the primaries run out, so that the extrapolation of Δ down to the limit of the secondaries may not be secure. The large dynamic range in density of the PDS ($\Delta D \simeq 4$) ensures that this is not an effect of the measuring machine, but corresponds to the images becoming unsaturated on the plate. Since the PDS measures are relatively insensitive to the degree of image saturation (because the software fits a Gaussian to the *unsaturated wings* of the image and calculates a total volume under the Gaussian), it is assumed that the calibration can be extended to the faintest secondary, i.e., $15.9 + \Delta = 19.2$. The calibration is carried back to the original polarization plate material by measuring several field stars of machine magnitude similar to the faint secondary images and defining a calibration curve on the polarization plates. The polarization plates were all measured on the COSMOS machine (see Section 4). This final transfer sequence is shown in Fig. 2, fit to this extended calibration gives an rms of 0.052 mag and a typical error for the faint end of the calibration is 0.1 mag.

4 Polarimetric calibration

4.1 COSMOS

The plate material taken through the polaroid filter is listed in Table 3. A complete polarization set was taken in 1980 May, with a partial set being taken the year before in 1979 May.

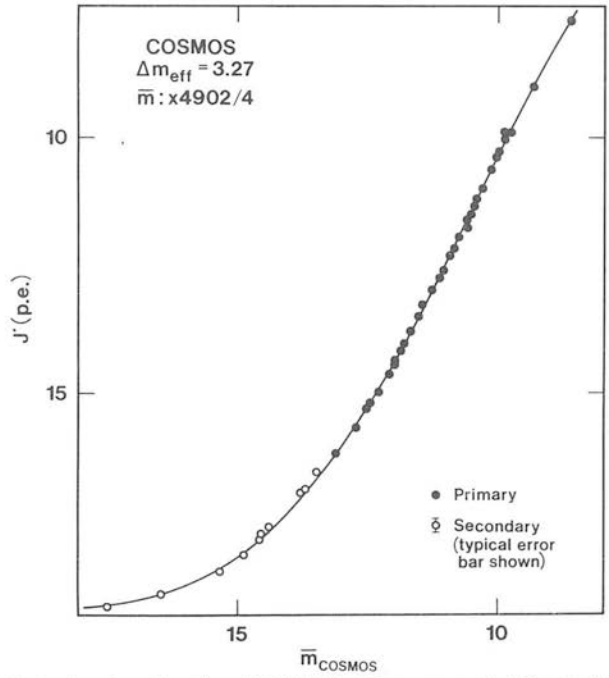


Figure 2. Extended photoelectric calibration (COSMOS). Colour-corrected J' magnitude is plotted against COSMOS magnitude for the mean of plates X 4902 and 4904. Filled circles are photoelectric sequence stars and open circles are the sub-beam prism extension points transferred from plates B5895, 5896 and 5897. A sixth order polynomial fit is shown with rms $\sigma(J') = 0.052$ mag.

Table 3. Details of the Schmidt polarization plates.

No.	Field centre	Position angle	Exposure	Seeing	Min. image size	Machine	Grade
X 4902	$12^h\ 58^m$ -20° 09′	0°	60 min	2″	35 μm	COSMOS	α
X 4903		45	60	2	40	COSMOS	α
X 4904		90	60	2	35	COSMOS	α
X 4905		135	60	2	40 x 55	—	β
X 5742	12 58 -20 09	0	50	1-2	30	COSMOS	α
X 5743		45	50	1-2	30	COSMOS	α
X 5744		90	50	1-2	35	COSMOS	α
X 5745		135	50	1-2	35	COSMOS	α

All plates use unfiltered IIIa-J emulsion taken through HN32 Polaroid. The grading system is the one used by the UK Schmidt Telescope Unit (for details see UKSTU Handbook).

A 10-square degree central region and a 0.7-square degree control region were measured on all these plates using the COSMOS machine at the Royal Observatory, Edinburgh. For references on the design, construction and use of COSMOS, see Pratt (1977), Stobie *et al.* (1979) and Dodd, MacGillivray & Smith (1979). The pixel-integrated intensity is defined as

$$I = \sum_{i=0}^{n} (E_i - E_{sky})/E_{sky},$$

where E is the exposure (intensity × time) and a subscript denotes the polaroid azimuth. The conversion from transmission to exposure *assumes* a IIIa-J contrast (γ) of 4 and it is recognized that COSMOS-derived intensity will in general be a non-linear function of true intensity. The photometric calibration of Section 3 is used to establish the function (Appendix A).

The passband of the IIIa-J/HN 32 combination contains more ultraviolet than the standard Schmidt J passband, because the usual GG 395 filter must be omitted. Fig. 3 shows the transmission curves for the emulsion, the polaroid filter, and the Johnson B and V bands. By convolving these curves at 100 Å resolution, the colour term to polaroid J' was found to be $(B-J') = (0.37 \pm 0.04)\ (B-V)$. The empirical colour term calculated from COSMOS measures of photoelectric sequence stars is $(B-J') = (0.38 \pm 0.03)\ (B-V)$.

4.2 POLARIZATION REDUCTION

The method for calculating polarization involves pairing images on orthogonal plates (0° and 90°, 45° and 135°) which can be done with a completeness of > 98 per cent. The image transformation between plates is accurate to ± 1 μm over an entire Schmidt plate, compared to a typical image size of 40 μm. Only images of more than 100 pixels, corresponding approximately to 1.5 mag above the plate limit, or $J' \sim 19.5$, were considered. Highly polarized images are well displaced from the mean relationship between I_0 and I_{90} [$I_{90} = f_1(I_0)$], I_{45} and I_{135} [$I_{135} = f_2(I_{45})$], where f_1 and f_2 are low-order polynomials fitted to the data. The polarization is calculated from the two residuals of these relationships. The dependence of calculated polarization on magnitude difference (for small polarizations) is

$$p = 0.92\,\Delta m.$$

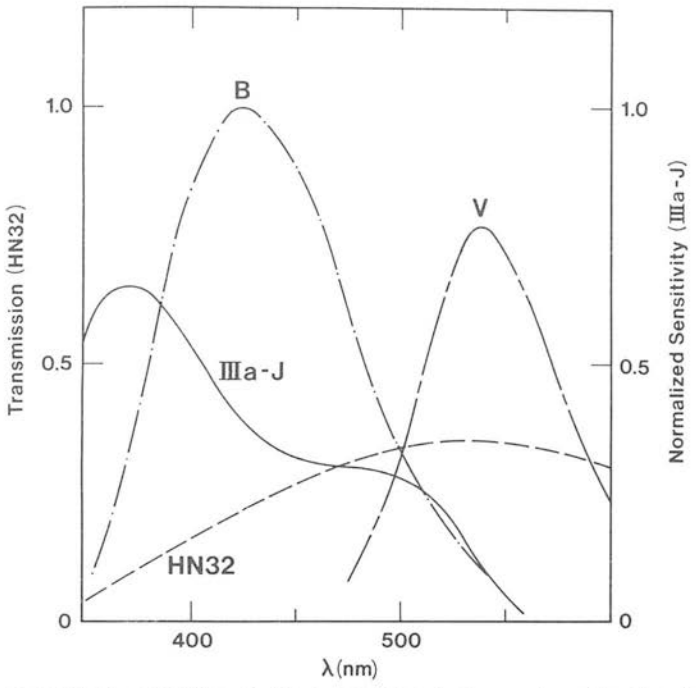


Figure 3. The normalized sensitivity of IIIa-J emulsion is shown, in units of reciprocal exposure ($1\,erg^{-1}\,cm^{-2}$), along with the HN-32 transmission to unpolarized light (data from Kodak Company and Polaroid Corporation). The spectral transmission of the Johnson *B* and *V* filters is also plotted.

In the initial COSMOS reduction, it has been assumed that the photographic contrast is 4 and as a result there is a non-linear relationship between COSMOS magnitude and J'. The true polarization can be written

$$p = p_{\mathrm{COSMOS}}(\Delta J'/\Delta m_{\mathrm{COSMOS}}).$$

Since most objects are not highly polarized, the calculated polarization represents *noise* rather than *signal*, but the distribution of this pseudo-polarization is a good indication of the detection limit. The solid curve in Fig. 4 shows $n(p)$ for 3500 images in a one degree square of the measured area (in the range $15 < J' \lesssim 19.5$). It is clear from the histogram that more than one in six images lie beyond $p = 15$ per cent, so the contamination of any real candidates by the noisy tail of that distribution will be severe. Two other plate pairs were used to corroborate the objects selected from the full polarization set. One consisted of the 0°–90° pair (Stokes parameter Q) taken the year before and the other consisted of the 90° plates which had been taken a year apart. The latter is an important pairing because it diagnoses variability between two epochs. An object must have $p > 15$ per cent from the full polarization set *and* have a large residual on the other two pairings to be a polarization candidate.

4.3 POLARIZATION ZERO POINT

The use of a differential technique has advantages, but it makes the assumption that mean relationships for the field stars, $I_{90} = f_1(I_0)$ and $I_{135} = f_2(I_{45})$, define zero polarization. This is true only if the field stars have no mean intrinsic polarization, *or* if they are polarized but with randomly orientated position angle. There are two reasons for believing the local interstellar polarizations to be low (< 1–2 per cent). First, the field is at $b^{II} = 36°$ and the

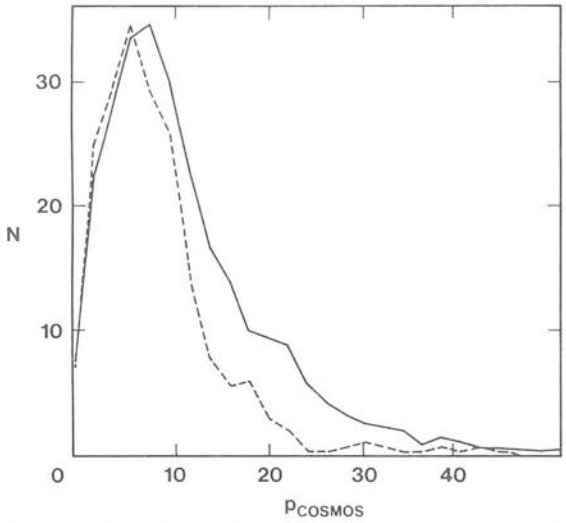


Figure 4. Polarization histogram from COSMOS measures. The number of distribution of machine polarization measurements (p_{COSMOS}) is shown for 3500 images in the range $15 < J \lesssim 19.5$ in a one-square degree field. The local renormalization to compensate for field effects has been applied to all images. The solid curve shows all 3500 images ~ 24 per cent with $p_{COSMOS} > 15$ per cent) and the dashed curve shows only those 140 images with non-circularities less than 1.5 (~ 8 per cent with $p_{COSMOS} > 15$ per cent). By removing non-circular images the number of candidates has been reduced by almost two orders of magnitude.

compilation by Mathewson & Ford (1970) shows the region to be of low polarization. Secondly, there are no differences greater than 2 per cent between the functions $f_1(I_0)$, $f_2(I_{45})$ for program region stars and the same functions for stars in the control region. The control region measurements are not sensitive to polarization.

4.4 SYSTEMATIC ERRORS

A differential technique can in principle remove many of the systematic differences between plates which will degrade the sensitivity of the search. Any differences in sky background, emulsion sensitivity and exposure time will be accounted for in the mean relationship between images. The effects of systematic errors on COSMOS machine magnitudes are discussed more fully in Appendix B. One problem is caused by the thresholding procedure in the COSMOS measurement, whereby only pixels a fixed percentage above sky background are integrated and the wings of an image are lost. If the plates in a polarization set are well matched, then no systematic error is introduced to a differential measurement. Photometry, however, will be compromised. When the sub-beam prism plates described earlier were measured on COSMOS, the prism constant Δ varied substantially with magnitude. This is understandable since the secondary images will be mini-spectra and a larger fraction of the image will slip under the COSMOS threshold than for the primary images. For this same reason, the search technique is only suitable for stellar images; extended images with flatter profiles will have measured intensities which may mimic polarization.

It was discovered that in addition to differences between plates, there were systematic, large amplitude field (~ 0.3 mag) effects across an individual plate. Field effects contribute scatter to the functions $f(I_0)$, $f(I_{45})$; and to diagnose the effect, mean residuals summed within a small area were contour plotted over the entire region. By dividing the region into sub-areas and subtracting the mean residual in a given sub-area from each magnitude within

that area, the field effects were removed. This was checked by contour plotting the residuals *after* correction and seeing that there were no systematic trends. The same procedure was followed for all the plate pairs and typically led to a reduction in the rms of the $f(I)$ relationships by ~ 30 per cent. The correction procedure involves making a local (~ 4 mm^2) transformation between the plates.

4.5 NON-STELLAR IMAGES

To reduce the number of polarization candidates and make a detection limit of $p = 15$ per cent a realistic target, it was necessary to discard non-stellar images, plate flaws and merged images which contaminate the wings of the residual distribution. To reject non-stellar images, those with a non-circularity (ratio of semi-major to semi-minor axes) greater than 1.5 were discarded. A non-circularity cut must be used with care because the shape of an ordinary stellar image containing a small number of pixels increases with fainter magnitudes (i.e. for an N pixel image, the eccentricity is typically $100/\sqrt{N}$ per cent). As a control, a sample of 100 images with non-circularities greater than 1.5 were examined by eye on the original plate material. All 100 were galaxies or double images, putting a limit of 1 per cent on the stellar rejection rate caused by the non-circularity cut. Compact galaxies and face-on ellipticals are not discriminated by this method, but the dotted line in Fig. 5 shows the improvement in detection limit. The number of candidates with $p > 15$ per cent is reduced by an order of magnitude and the total number of images to sort through is reduced by a factor of 100.

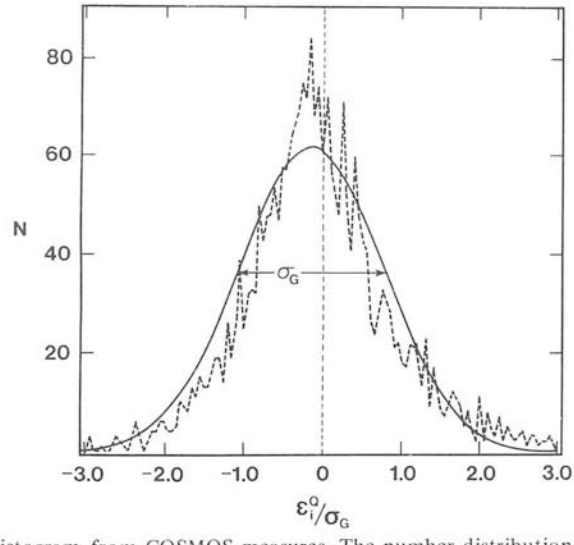


Figure 5. Residual histogram from COSMOS measures. The number distribution of machine magnitude residuals $n(\epsilon_i^Q)$ is shown for 3000 images in the range $15 < J < 19.5$, where $\epsilon_i^Q = [m_{90} - f(m_0)]_i$. The function $f(m_0)$ is derived from the polynomial fit to plates taken through orthogonal angles of the polaroid (i.e. $m_{90} = f(I_0) = a_0 + a_1 m_0 + a_2 m_0^2 + a_3 m_0^3 + \ldots$). An individual residual can be related to the normalized Stokes parameter by

$$\frac{Q}{I} = p \cos 2\theta = [10^{-(\epsilon_i^Q/2.5)} - 1]/[10^{-(\epsilon_i^Q/2.5)} + 1].$$

The localized renormalization to compensate for field effects has been applied to all images and those images with non-circularities greater than 1.5 have been discarded. The distribution of residuals is well approximated by a normal distribution, which is shown as a solid line of rms width σ_G.

The final candidate selection was made from objects appearing with high residuals on lists selected from (i) the complete polarization set, (ii) the pair defining the Stokes parameter Q and (iii) a variability pair separated by one year. The detection of a very small number of polarized objects from 33 000 in the presence of photometric scatter is a severe signal-to-noise problem. To get a reliable estimate of the significance of a polarization candidate, it is necessary to model the error distribution. For each plate pair, the residual distribution was modelled by a Gaussian distribution and candidates assigned significance values in terms of the dispersion. Fig. 5 shows the residual histogram and it appears to be the superposition of a well-behaved normal distribution and a high-σ tail. The data were truncated at $\pm 3\sigma_A$, where σ_A is the rms from the least-squares fit of all the data including compact galaxies and the central part of the distribution fitted by minimum-χ^2 with a Gaussian curve. The width of the fitted curve, σ_G, was then taken to be the spread due to random photometric errors. When the three indicators were combined, there were 33 objects with a significance level $> 3\sigma_G$: a reduction by a factor of 1000 from the original sample. The final selection criterion involved segregating out those galaxies which had escaped the eccentricity cut. If the logarithm of the image area is plotted against machine magnitude, it is seen that stars form a linear relationship and galaxies smear away upwards from the line. Fig. 6 shows this plot and the displacement of objects from the mean relationship is a measure of compactness, with diffuse images displaced upwards. The data for several thousand images are plotted, with the 33 candidates superimposed. From visual inspection only nine are stellar, with the remainder showing up clearly as compact or dwarf galaxies. A line can be drawn which completely separates stars and galaxies on the $\log A$–m plot. The reduction in the number of candidates after the various selection cuts is shown in Table 4.

Therefore, from a 10-square degree area, nine BL Lac candidates have been isolated within the limits $p > 15$ per cent, $B < 19.0$. The BL Lac candidates are listed in Table 5, along with Schmidt J magnitudes, positions at epoch 1950.0 and the significance level of the polarization and variability cuts. The positions were interpolated from a coordinate transformation involving 12 SAO stars around each candidate and each position is accurate to

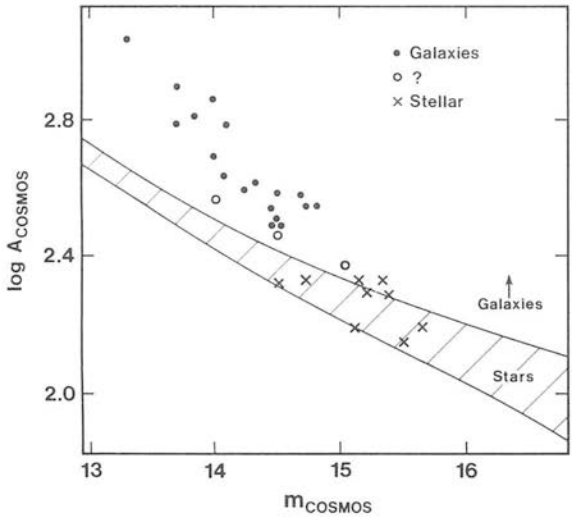


Figure 6. COSMOS image area is plotted against COSMOS magnitude for field stars (shaded area) and polarization candidates (symbols). Filled circles represent galaxies, open circles represent compact images of uncertain morphology and crosses represent stellar images.

Table 4. Polarization candidates (10.2-square degrees).

	All images > 7 percent above sky background	84 420
Reject:	Images fainter than B = 19.5	34 100
Reject:	Images with a/b > 1.50	3 510
Reject:	Images with polarization/variability cut < 3σ	33
Reject:	Images nonstellar on log A–m plot	9

Table 5. Details of polarization candidates.

Name	J magnitude	α(1950.0)	δ(1950.0)	$(\epsilon_i/\sigma_G)_{Pol.}$	$(\epsilon_i/\sigma_G)_{Var.}$
PV 1253-198	18.7	$12^h\ 53^m\ 01\overset{s}{.}3$	-19° 45' 01.7"	4.2	3.0
PV 1254-214	17.3	12 54 08.8	-21 26 19.6	5.7	2.5
PV 1255-213	18.6	12 55 39.0	-21 18 52.8	4.5	4.1
PV 1256-207	17.1	12 56 56.7	-20 34 53.8	3.2	2.2
PV 1259-202	18.9	12 59 34.1	-20 12 57.8	5.6	3.6
PV 1300-212*	16.5	13 00 21.0	-21 13 28.1	5.7	3.6
PV 1301-207	18.7	13 01 06.7	-20 41 23.8	6.6	2.5
PV 1303-206	18.6	13 03 30.8	-20 36 38.8	2.5	2.7
PV 1305-199	18.6	13 05 24.2	-19 51 38.4	3.0	2.5

*Image slightly soft and noncircular on UKSTU IIIa-J plate, but not culled out by COSMOS image eccentricity or log A-m tests.

$\sigma(\alpha) = 1.1$ arcsec, $\sigma(\delta) = 0.5$ arcsec. The significance levels (ϵ_i/σ_G) are calculated from the magnitude residuals for a given pair of plates (A16) and the rms of the Gaussian fit to the residual distribution (A17). Three independent significance levels from three pairs of plates are combined quadratically to produce the polarization number, whereas the variability number represents the significance level from one pair of plates. Further study, including photoelectric polarimetry, spectroscopy and infrared photometry is needed to confirm these candidates as BL Lac objects. The new limit on the space density of BL Lac objects is therefore given as:

$p > 15$ per cent

$V > 0.3$ mag $\qquad n_{BL} = 0.9^{+0.3}_{-0.9}$ (square degree)$^{-1}$,

$J < 18.9$ $(B < 19)$

where the upper limit is statistically defined from the number of candidate objects.

5 Conclusions

From a combined selection for polarization and variability, a list of nine candidate BL Lac objects has been selected. To qualify the implied limit on the surface density of BL Lacs, it is necessary to have an estimate of the completeness of the search. All 57 polarized objects

listed by Angel & Stockman (1980) have *mean* visual magnitudes brighter than the limit of the survey and, accounting for variability, they spend ~ 90 per cent of their time above the brightness limit (Pollock *et al.* 1979; Pica *et al.* 1980). Twenty-four (42 per cent) have maximum optical polarizations greater than 15 per cent. The variable polarization means that each object with $p_{MAX} > 15$ per cent will spend a portion of its time with p under 15 per cent. The most comprehensive monitoring is that of Angel *et al.* (1978). Of the 241 observations with $p_{MAX} > 15$ per cent, 90 (37 per cent) individual measurements were $p > 15$ per cent. The completeness is also affected by the image pairing statistics of matching images on four separate plates: this introduces a factor of 0.96. There is also the proportion of known BL Lacs which would be excluded by the requirement of including only stellar or very compact objects which is ~ 0.86. Finally, the proportion of stellar BL Lac objects, which will typically vary by 0.3 mag in a year, must be taken into account. Taking magnitude levels at V roughly a year apart from the compilation of Pollock *et al.* (1979), approximately 85 per cent of the magnitude differences are greater than 0.3 mag (the typical rms accuracy of the photographic photometry is 0.12 mag, so only ~ 15 per cent of the differences will be greater than 0.3 mag by chance). Combining these factors gives a completeness of ~ 0.11; roughly one in nine known BL Lac objects could be detected photographically. It should be noted that this completeness figure is probably an underestimate, since the current population of BL Lac objects was *not* selected by variability or polarization.

The result of this experiment is of interest if either any or none of the candidates is confirmed as a BL Lac object. There is *no* known example of a radio-quiet BL Lac object; so one confirmation would be evidence for a new class of extragalactic objects. If none of the candidates is confirmed, the result will constrain the beaming models of Blandford & Rees (1978). These models predict that in searches to a given flux density, BL Lac objects will not be rare with respect to quasars, because of the strong selection effect in favour of forward-beamed sources. The surface density of quasars increases by a factor of 4–6 per magnitude interval down to $B = 19$ (Braccesi *et al.* 1980; Usher 1978), so an optical search looks deeper into the luminosity function of BL Lac objects than quasars. For a given luminosity function $[n(L) \alpha L^{-s}]$ and beaming 'open angle', the relative number of quasars to BL Lac objects puts an upper limit on the relativistic bulk motion factor (Γ) of the beamed material. If none of the BL Lacs is confirmed, this implies a severe limit on the allowed parameters of the beam models.

Acknowledgments

We would like to acknowledge the continued support of John Dawe and the UK Schmidt Unit, who have provided excellent plate material for this project. The computational side of this work has drawn heavily on the advice and experience of Ralph Martin and Bernard McNally at ROE; we would like to thank all the members of the COSMOS team for diligently measuring so many plates. Bill Martin of RGO provided assistance with the PDS measurements. Thanks are also due to Steve Heathcote and Paul Hewitt for the use of plotting packages. We benefited from useful discussions with Bob Stobie, Bob Smith, Gerry Gilmore, Mike Hawkins and the Edinburgh UKSTU staff. CDI acknowledges the receipt of an SRC studentship.

References

Agrawal, P. C. & Riegler, G. R., 1979. *Astrophys. J.*, **231**, L25.
Allen, C. W., 1973. *Astrophysical Quantities*, Athlone Press, London.

Altschuler, D. R. & Wardle, J. F. C., 1975. *Nature*, **255**, 306.
Angel, J. R. P., Boroson, T. A., Adams, M. T., Duerr, R. E., Giampapa, M. S., Gresham, M. S., Gural, P. S., Hubbard, E. N., Kopriva, D. A., Moore, R. L., Peterson, B. M., Schmidt, G. D., Turnshek, D. A., Wilkerson, M. S., Zotov, N. V., Maza, J. & Kinman, T. D., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 117, ed. Wolfe, A. M., University of Pittsburgh.
Angel, J. R. P. & Stockman, H. S., 1980. *A. Rev. Astr. Astrophys.*, **18**, 321.
Auer, L. H. & van Altena, W. F., 1978. *Astr. J.*, **83**, 531.
Baker, A. E., 1925. *Proc. R. Soc. Edinburgh*, **45**, 166.
Blandford, R. D. & Rees, M. J., 1978. *Pittsburgh Conference on BL Lac Objects*, p. 328, ed. Wolfe, A. M., University of Pittsburgh.
Braccesi, A., Zitelli, V., Bonoli, F. & Formiggini, L., 1980. *Astr. Astrophys.*, **85**, 80.
Cooke, B. A., Ricketts, M. J., Maccacaro, T., Pye, J. P., Elvis, M., Watson, M. G., Griffiths, R. E., Pounds, K. A., McHardy, I., Maccagni, D., Seward, F. O. & Turner, M. J., 1978. *Mon. Not. R. astr. Soc.*, **182**, 489.
Couch, W. J. & Newell, E. B., 1981. *Proc. astr. Soc. Australia*, **4**, 106.
de Vaucouleurs, G., 1968. *Appl. Opt.*, **7**, 1513.
Dodd, R. J., MacGillivray, H. T. & Smith, G. M., 1979. *Image Processing in Astronomy*, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Trieste.
Impey, C. D., Brand, P. W. J. L. & Tapia, S., 1981. *Mon. Not. R. astr. Soc.*, **198**, 1.
Katgert, P., Katgert-Merkelijn, J. K., Le Poole, R. S. & van der Laan, H., 1973. *Astr. Astrophys.*, **23**, 171.
Khachakian, E. Ye. & Weedman, D. W., 1974. *Astrophys. J.*, **189**, L99.
Mathewson, D. S. & Ford, V. L., 1970. *Mem. R. astr. Soc.*, **74**, 1.
Menzies, J. W., Banfield, R. M. & Liang, J. D., 1980. *SAAO Circ.*, **1**, 5.
Murdoch, H. S. & Crawford, D. F., 1977. *Mon. Not. R. astr. Soc.*, **180**, 41P.
Pica, A. J., Pollock, J. T., Smith, A. G., Leacock, R. J., Edwards, P. L. & Scott, R. L., 1980. *Astr. J.*, **85**, 1442.
Pickering, E. G., 1891. *A. astr. Obs. Harvard*, **26**, 14.
Pollock, J. T., Pica, A. J., Smith, A. G., Leacock, R. J., Edwards, P. L. & Scott, R. L., 1979. *Astr. J.*, **84**, 1658.
Pratt, N. M., 1977. *Vistas Astr.*, **21**, 1.
Stobie, R. S., Smith, A. G., Lutz, R. K. & Martin, R., 1979. *Image Processing in Astronomy*, eds Sedmak, G., Capaccioli, M. & Allen, R. J., Trieste.
Stockman, H. S., Angel, J. R. P. & Hier, R. G., 1981. *Astrophys. J.*, **243**, 404.
Usher, P. D., 1978. *Astrophys. J.*, **220**, 40.
Wolstencroft, R. D., Impey, C. D. & Smith, R. J., 1981. *Mon. Not. R. astr. Soc.*, **194**, 275.

Appendix A

The polarimetry reduction is briefly summarized in this Appendix. When light of intensity I is incident on the filter, the transmitted intensity is

$$T_\phi I = I\left(\frac{K_1 + K_2}{2} + p\,\frac{K_1 - K_2}{2}\cos 2(\phi - \theta)\right), \tag{A1}$$

where ϕ is the filter position angle, K_1 nd K_2 are the principal transmission coefficients of the polaroid and p and θ are the degree of polarization and position angle of the incident radiation. An extra control factor C_ϕ takes into account differences between the plates,

$$I_\phi = C_\phi T_\phi I. \tag{A2}$$

From measured values of I_ϕ and C_ϕ, $T_\phi I = I_\phi / C_\phi$ can be calculated and hence the normalized Stokes parameters

$$\frac{U}{I} = p \sin 2\theta = R\,\frac{T_{45} - T_{135}}{T_{45} + T_{135}},$$
$$\frac{Q}{I} = p \cos 2\theta = R\,\frac{T_0 - T_{90}}{T_0 + T_{90}}. \tag{A3}$$

R is essentially unity ($R - 1 \simeq 2K_2/K_1 < 0.006$) from instrumental tests. For images measured in the control region,

$$I'_\phi = A\, C_\phi I, \tag{A4}$$

where A is a constant. Therefore

$$\frac{C_0}{C_{90}} = \frac{I'_0}{I'_{90}}$$
$$\frac{C_{45}}{C_{135}} = \frac{I'_{45}}{I'_{135}}. \tag{A5}$$

For conciseness write $T_Q = T_0/T_{90}$ and $T_U = T_{45}/T_{135}$, giving

$$p = \left[\left(\frac{T_Q - 1}{T_Q + 1}\right)^2 + \left(\frac{T_U - 1}{T_U + 1}\right)^2\right]^{1/2}, \tag{A6}$$

$$\theta = \frac{1}{2}\arctan\left[\left(\frac{T_U - 1}{T_U + 1}\right)\Big/\left(\frac{T_Q - 1}{T_Q + 1}\right)\right]. \tag{A7}$$

Propagating errors through (A7) and (A8) leads to

$$\sigma^2(p) = \frac{4}{p^2}\left[\left(\frac{T_Q - 1}{T_Q + 1}\right)^2 \frac{\sigma^2(T_Q)}{(T_Q + 1)^4} + \left(\frac{T_U - 1}{T_U + 1}\right)^2 \frac{\sigma^2(T_U)}{(T_U + 1)^4}\right], \tag{A8}$$

$$\sigma^2(\theta) = \frac{1}{p^4}\,\frac{[(T_Q - 1)^2 + (T_U - 1)^2]}{(T_q + 1)^2\,(T_U + 1)^2}, \tag{A9}$$

where

$$\frac{\sigma^2(T)}{T_Q^2} = \sigma^2(T_0) - \frac{\sigma^2(T_{90})}{T_{90}^2}, \tag{A10}$$

and

$$\frac{\sigma^2(T_U)}{T_U^2} = \sigma^2(T_{45}) - \frac{\sigma^2(T_{135})}{T_{135}^2}. \tag{A11}$$

Although equations (A6) and (A7) indicate that p and θ are independent of the incident intensity I, they are in fact dependent on the assumed contrast γ.

Most of the polarization candidates at the faint end of the sample will be on the part of the calibration curve where the contrast is changing (Fig. 2). Therefore, from the relationship between p and p_{COSMOS} given in the text, the *sensitivity* to polarization increases with increasing magnitude. There is also a signal-to-noise effect in the opposite sense, because the *detectability* of a given polarization decreases with increasing magnitude. Both effects must be combined to give a true polarization detection limit. The zero point is fixed at the level where $p_{\mathrm{COSMOS}} = p$; $J' = 16.95$. Fig. A1 shows p/p_{COSMOS}, p_{NOISE}/p and $p_{\mathrm{NOISE}}/p_{\mathrm{COSMOS}}$ plotted against calibrated magnitude J'. With the calibration curve and signal-to-noise taken into account, $p_{\mathrm{NOISE}}/p_{\mathrm{COSMOS}}$ still has a slight magnitude-dependence. The *mean* polarization detection limit obtained by convolving the magnitude distribution with $p_{\mathrm{NOISE}}/p_{\mathrm{COSMOS}} = f(J')$ is $\bar{p} = 11.7$ per cent. A more precise way of stating the limits of

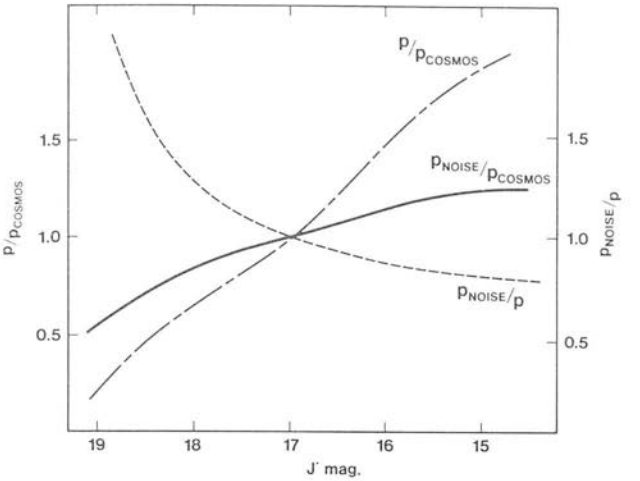


Figure A1. Calibrated J' magnitude is plotted against mean value of p/p_{COSMOS}, p_{NOISE}/p and p_{NOISE}/p_{COSMOS} from a total of 3000 images. p/p_{COSMOS} shows the effect of the changing γ of the emulsion, p_{NOISE}/p shows the effect of signal-to-noise with magnitude and p_{NOISE}/p_{COSMOS} shows how the relative detection limit (normalized to $p = 15$ per cent at $J' = 16.96$) varies with magnitude.

the survey is that it is complete to better than $p = 15$ per cent for images fainter than $J' = 17$, which includes 80 per cent of all images. It is also complete to $p = 15^{+5}_{-8}$ per cent for essentially all images. This discussion does not affect the completeness argument in the conclusions substantially.

Appendix B

Two classes of systematic error which affect COSMOS measures of UK Schmidt plates are discussed. The results are not specific to photographic polarimetry and in view of the increasing use of automated machine measurement should be of general interest.

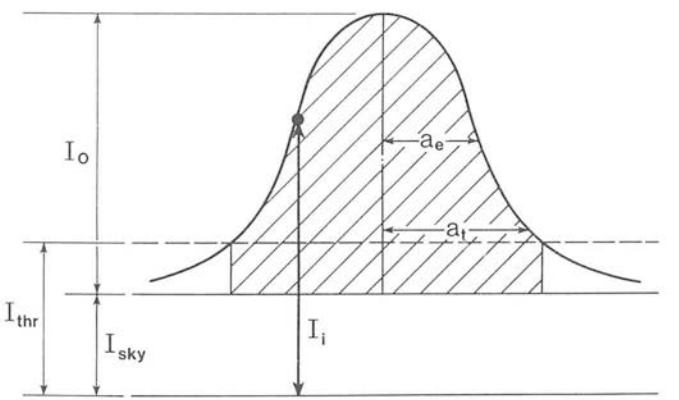


Figure B1. COSMOS parameters for an unsaturated stellar image. I_i is the intensity at a given pixel i. I_{sky}, I_{thr} and $(I_{sky} + I_o)$ are the sky, threshold and peak intensities respectively. The integrated intensity

$$I = \sum_{i=1}^{n} (I_i - I_{sky})$$

is given by the shaded area. a_t is the radius of the threshold image and a_e is the radius where $(I_i - I_{sky}) = I_o/e$ (schematic diagram only).

The principal COSMOS image parameter is the pixel integrated intensity, I, which is summed from all pixels a fixed percentage above the sky or background intensity. This threshold intensity and other COSMOS parameters are shown in Fig. B1 for an unsaturated, stellar image.

1 Emulsion sensitivity and development variations

For any exposure which is not sky limited with a background density of order one, it is convenient to work in Baker densities (Baker 1925; de Vaucouleurs 1968). The Baker density is used to linearize the toe of the characteristic curve

$$D_B = \log \omega = \log \left(\frac{T_c}{T} - 1 \right), \tag{B1}$$

where T_c is the background fog level or clear transmission of the plate. Relating densities to a zero point of T_c produces a linear characteristic curve at low ($D_B < 2.5$) densities

$$I = A \left(\frac{T_c}{T} - 1 \right)^n = A \omega^{1/\gamma'}. \tag{B2}$$

where A is a constant and the contrast $\gamma' = 1/n$ is the gradient of the characteristic curve. The use of Baker density will not produce an exact 1:1 relationship between COSMOS intensity and true intensity. However, γ' for areas of sky and image will be similar enough to allow a simplified treatment of the problem.

Small changes in development time are equivalent to small changes in density and small changes in emulsion sensitivity are equivalent to small changes in exposure, i.e., γ' is constant for dD_B and $d(\log I)$ small (see Kodak Publication P-315).

$$dD_B = \gamma' \, d(\log I) \tag{B3}$$

and

$$\frac{(dD_B)_{\text{sky}}}{(dD_B)_{\text{obj}}} = \frac{\gamma'_{\text{sky}}}{\gamma'_{\text{obj}}} = \alpha. \tag{B4}$$

To move into COSMOS transmission terms,

$$\frac{\log(1 + k_{\text{sky}})}{\log(1 + k_{\text{obj}})} = \alpha, \tag{B5}$$

where k_{sky} and k_{obj} are the fractional changes in transmission for sky and image pixels, respectively. The assumption of linearity over the entire range is a good one. Therefore taking $\alpha = 1$ and if $k_{\text{sky}} = k_{\text{obj}} = k$, the incoming intensity is

$$I_i = A \left[\frac{T_c}{T_i(1 + k)} - 1 \right]^n. \tag{B6}$$

For most pixels $T_c/T \gg 1$ and for small sensitivity changes $k \ll 1$, so approximately

$$I_i = A(1 + k)^{-n} \left[\frac{T_c}{T_i} - 1 \right]^n = A(1 + k)^{-n} \omega_i^n. \tag{B7}$$

The variations in the amplitude of k will always be small over the localized region of one image, so summation over all pixels in an image including the local sky pixels will have the same k value,

$$I_{\rm sky} = A(1+k)^{-n}\left[\frac{T_{\rm c}}{T_{\rm sky}} - 1\right]^n = A(1+k)^{-n}\omega_{\rm sky}^n . \tag{B8}$$

Thus the sensitivity factor can be scaled out by dividing the image intensity by $I_{\rm sky}$ at each pixel,

$$I = \sum_{\rm i=1}^{n} (I_{\rm i} - I_{\rm sky})/I_{\rm sky}. \tag{B9}$$

On one pair of plates, where uniformity of the emulsion batch was in doubt, using the adjusted image parameter reduced the dispersion from one plate against the other by 15 per cent. This correction can be applied to any plate with sensitivity or development non-uniformities, with one important exception. If there are *real* variations in sky intensity across the plate, the unscaled intensity will be a better estimator.

2 Image thresholding

Fig. B1 shows that the COSMOS thresholding procedure loses a small volume in the wings of the image; and for a given threshold cut, the loss is greater for fainter and fainter images. It is assumed that the intensity profile is a Gaussian of half-width $a_{\rm e}$ and maximum height above sky I_0 (Auer & van Altena 1978). $a_{\rm t}$ is the half-width at the threshold intensity $I_{\rm thr}$. Let c be the fractional percentage cut above $I_{\rm sky}$ and s be the ratio of peak object to sky intensity (Fig. B1). Summarizing:

$$\begin{aligned} I_0 &= s I_{\rm sky} \\ I_{\rm thr} &= (1+c) I_{\rm sky} \\ \Delta I &= c I_{\rm sky}. \end{aligned} \tag{B10}$$

The total image intensity can be expressed in terms of I_0, $I_{\rm sky}$, $I_{\rm thr}$ and $a_{\rm t}$ without summing over $(I_{\rm i} - I_{\rm sky})$. The image profile is

$$I(r) = I_0 \exp(-r^2/a_{\rm e}^2). \tag{B11}$$

Integrating over the volume,

$$J = \int_0^\infty I(r)\, 2\pi r\, dr = \pi a_{\rm e}^2 I_0. \tag{B12}$$

However, $\Delta I = I_0 \exp(-a_{\rm t}^2/a_{\rm e}^2)$ and $a_{\rm t} = a_{\rm e}\sqrt{\log s/c}$, since $I_0/\Delta I = s/c$. Therefore,

$$J = \frac{\pi a_{\rm t}^2 I_0}{\log(s/c)}. \tag{B13}$$

The wings of the image are given by

$$J_{\rm w} = \int_{a_{\rm t}}^\infty I_0 \exp(-r^2/a_{\rm e}^2)\, 2\pi r\, dr, \tag{B14}$$

$$J_{\rm w} = \pi a_{\rm e}^2 I_0 \left(\frac{c}{s}\right) = J\left(\frac{c}{s}\right), \tag{B15}$$

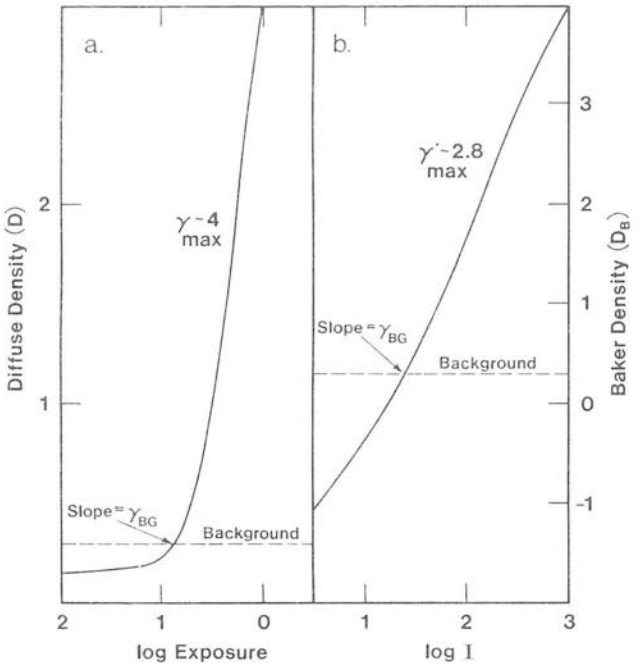


Figure B2. Characteristic curves for IIIa-J emulsion in terms of (a) normal diffuse density and (b) Baker density. The typical background level for the Schmidt polarization plates is indicated (data from Kodak and UKSTU).

therefore,

$$J_w/J = c/s. \tag{B16}$$

This is a very simple relationship between the fraction of the image lost by thresholding, the threshold cut c and the peak signal/sky ratio. Typically, $c = 7$ per cent and if c is much less than 5 per cent, fluctuations due to grain clumps begin to be mistaken for images by COSMOS. Measurements on 20-min exposure, unfiltered IIIa-J plates indicate that a magnitude from the plate limit with $c = 7$ per cent, the photometry is systematically wrong by 5–6 per cent. This number gets dramatically worse for images closer to the plate limit, or for plates where the image profiles are known to vary between faint and bright stars, e.g. sub-beam prism plates.

Mon. Not. R. astr. Soc. (1984) **211**, 485–495

Detection of lines by a calibrated algorithm in a close pair of quasars

A. S. Trew* and P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 1984 June 12. Received 1984 June 12; in original form 1984 December 7

Summary. Comparison of the current quasar line detection techniques shows that the results do not always agree and as a consequence a debate has arisen as the true redshift distribution of the narrow quasar absorption lines. We have developed a method for finding such features and have subjected it to rigorous testing on model spectra. The results are presented here showing that this method is very efficient and has no significant biases. The algorithm has also been applied to a close pair of quasars in a grouping discussed by Arp & Hazard which have been observed at high resolution. We find only marginal evidence for absorption common to both objects and that the redshift distributions of the absorption systems are otherwise normal.

1 Introduction

The increasing use, and number of, digitized quasar spectra over the last decade has led to the development of a number of techniques designed to detect the narrow absorption lines (e.g. Young *et al.* 1979). However, differences between the resultant line distributions produced by these methods, albeit on different observational data, leave open the possibility that existing techniques contain biases. This is most notably typified by the controversy over the putative existence of a peak in the metal-line distribution close to the emission-line redshifts (e.g. Weymann, Carswell & Smith 1981; Young, Sargent & Boksenberg 1982). Moreover, neither side in this debate has published sufficient information, for example based upon the analysis of model spectra, to enable an independent evaluation of their methods to be made. While such a situation exists, it is unlikely that the problem of the redshift distribution of quasar absorption lines will be resolved.

To this end we have developed a technique, based upon that of Young *et al.* (1979) and have subjected it to rigorous testing on large numbers of model quasar spectra. The results of these tests, and the application to a close pair of quasars, are described in this paper.

This close pair of quasars was first discovered in a compact quasar grouping by Arp & Hazard (1980). The five objects (B, C, D, E, G as marked in their fig. 1) are within 5 arcmin on the sky

* Present Address: Edinburgh Regional Computing Centre, University of Edinburgh, King's Buildings, Edinburgh.

and four have redshifts 0.86, 1.01, 1.01 and 1.10, though only two (D and E) are sufficiently bright for high-resolution spectroscopy ($m\sim18.0$). Two reasons for obtaining high-resolution spectra thus present themselves.

First, it is probable, if quasar redshifts are cosmological, that at least two of the quasars are members of a rich cluster of galaxies at $z\sim1$, though at this redshift the galaxies themselves would be unlikely to be observable on the available UK Schmidt plates (*cf.* Bruzual 1981). A simple calculation shows that even without reference to the close redshift agreement between A, B and E (which given the uncertainties in the redshift distribution is difficult to quantify) the probability that this association would arise by chance is ~0.17 per cent (assuming a density of 10 quasars per square degree, typical of the numbers produced by slitless prism surveys which have limiting magnitudes ~19.5), making a random grouping unlikely. As D and E are separated by only 160 arcsec (~0.6–1 Mpc at $z=1$ for $q_0=0-1$) the pair of quasars might therefore be expected to show correlations in their absorption-line spectra indicative of common gas clouds along the line-of-sight, perhaps located within the cluster itself.

The second reason for observing these is that they both belong to one of the quasar alignments which Trew, Clube & Savage (1984) found to be possible physical associations. They are thus prime candidates for having significant non-cosmological redshifts. If the majority of absorption lines are produced in intervening galactic haloes then objects with non-cosmologically redshifts would have very different absorption-line distributions from the norm.

2 The absorption line detection technique

Upon completion of the data reduction the observer is left with a wavelength-calibrated, sky-subtracted and flat-fielded spectrum. If there are no features in these data then they should have a Poisson distribution about the mean, although the noise level will be higher than that expected on the basis of the observed counts because of the sky and detector backgrounds. The presence of narrow spectral lines will also result in raising the fluctuations about the continuum and it is these features which it is hoped to detect by a comparison of the observed and expected noise levels.

While as noted previously the method decribed here is similar to that used by Young *et al.* (1979) we have laid out the equations in detail to clarify the notation and precisely define the means of continuum fitting and line selection.

2.1 CONTINUUM FITTING

Let the final object spectrum consist of a series of counts, where $N(i)$ is the number in the ith channel. These data have been divided into N_s segments of N_c channels width. Within each segment a straight line has been fitted to the continuum by a least-squares method, with all channels being assigned equal weight. This technique is preferable to simply determining a mean continuum level, as practised by Young *et al.* (1979), because it permits sudden changes in slope and rate of change of slope to be accommodated. It is possible to measure the observed variance, σ_0^2, of the channels about the best-fitting line and if the data contain no features then the fluctuations about the mean should have a Poisson distribution, with variance

$$\sigma_T^2=\frac{\sum_{i=1}^{N_c}\sigma^2\,[N(i)]}{N_c}$$

where

$$\sigma^2\,[N(i)]=\left[\frac{1+f(i)}{1-f(i)}\right]N(i)$$

and where $f(i)$ is the ratio of the sky and detector backgrounds to the object +sky and detector counts in the ith channel. Narrow spectral lines or noise spikes, however, will increase the fluctuations about the mean and raise the observed variance, (σ_0), above that expected. In that event the channels within the segment were searched for the one which deviated most from the mean continuum. This point was rejected, the best-fitting line adjusted accordingly and the observed and expected variances redetermined. This process of rejecting the most deviant point was continued until $\sigma_0 \leqslant \sigma_T$ or the number of points removed exceeded $N_c/2$. If the latter proved to be the case then that segment was ignored, otherwise, the continuum level was taken to be that given by the least-squares line at the average wavelength of the channels.

The set of wavelength/continuum level points from all the acceptable segments, excepting the first and last (*cf.* Schoenberg & Whitney 1953) were used as knots for a spline interpolation to give the continuum level, $C(i)$, at each channel. However, as extrapolation with cubic splines is not permitted, the continuum levels in the first and last $N_c/2$ channels were not determined. Consequently, no search for lines was made in these regions.

Experiments with artificial spectra showed that this technique fitted the continuum to within 5 per cent, even in the presence of many strong, narrow lines. In the real data, however, it is possible that the continuum may be affected if there are large numbers of lines below the detection limit.

2.2 DETECTION OF ABSORPTION LINES

Any objective technique for the detection of absorption lines should be constructed in such a way as to select all those lines which have a probability of reality greater than some lower limit. The following procedure determines the equivalent width of every possible feature and discards those for which this is less than S standard deviations of the equivalent width. Thus, in the case of variations in the signal/noise across the spectrum, the minimum equivalent width will not be constant, but will rise as the data become more noisy.

By definition, the centroid of the line is given by

$$\langle\lambda\rangle=\frac{\sum_{i=1}^{M}\lambda(i)\,[C(i)-N(i)]}{\sum_{i=1}^{M}\,[C(i)-N(i)]}$$

where the summations are made across the M channels which form the line and $\lambda(i)$ are the wavelengths of each channel. Also by definition the equivalent width (in units of bins) of an absorption line is

$$W=\sum_{i=1}^{M}[1-N(i)/C(i)].$$

Clearly, because of noise in the data neither the line centroid nor the equivalent width will be determined perfectly, but we have some associated error. It is necessary for this approach that these uncertainties be determined. The variances of these quantities may be obtained by use of the Chain Rule (*cf.* Bevington 1969), which gives for the equivalent width, after some

rearrangement,

$$\sigma_W^2 = \frac{1}{C}\left[M\left(\frac{f+1}{f-1}\right) - W\right]$$

where f and C are the means of $f(i)$ and $C(i)$ across the line respectively. Also for the position of the line centroid we have

$$\sigma_{\langle\lambda\rangle}^2 = \frac{1}{CW^2}\left[\frac{f+1}{f-1}\right]\frac{M^3}{12} - \frac{\sigma_L^2}{CW}$$

where

$$\sigma_L^2 = \frac{1}{CW}\sum_{i=1}^{M}[\langle\lambda\rangle - \lambda(i)]^2\,[C(i) - N(i)].$$

In the case of a Gaussian line profile σ_L^2 takes on a particularly simple form and the above becomes

$$\sigma_{\langle\lambda\rangle}^2 \simeq \frac{\sigma_W^2}{W^2}\frac{M^2}{12}.$$

Since for narrow features the line profile will be determined by the detector rather than by the line-producing region we have used this latter expression for $\sigma_{\langle\lambda\rangle}$ in subsequent work.

Examination of the expressions for $\sigma_{\langle\lambda\rangle}$ and σ_W shows that both increase with M. Some compromise must therefore be made to keep the standard deviations as small as possible without unduly affecting W. The criterion which we have used to set M is to find the value which maximizes W/σ_W. We can do this before analysing the data by noting that the vast majority of narrow quasar absorption lines are unresolved at this resolution, so that the line width is determined by the instrumental profile. On the asumption that this profile is a Gaussian we may rewrite the equation for W as

$$W = \int_{-M/2}^{M/2} \frac{Ig(\lambda)}{C}\,d\lambda$$

where I is the depth of the line and $g(\lambda)$ the Gaussian line profile. Substitution of this into σ_W enables the value of M which maximizes W/σ_W to be found.

This optimum value of M will depend somewhat upon f and the line strength but is $\sim 2.5\times$FWHI of the detector resolution profile for lines with $W \lesssim 3$Å. Clearly, more accurate determinations are of limited value as M must be integral.

While this value of M will give accurate estimates of W for features with a Gaussian profile (only 0.3 per cent of the line intensity lies outside $\pm M/2$) it will be an underestimate if the lines or the instrumental profile have significant wings. For the RGO spectrograph and IPCS combination used here an examination of the arc line profiles revealed no significant wings to the detector profile. Adopting the value for the FWHI of the instrumental profile given in the RGO spectrograph handbook of 2.6–2.8 channels gives $M=7$, a choice corroborated by the arc line widths.

After determination of the continuum level, as described in Section 2.1, the detection of lines was accomplished by stepping a box of M channels width from one end of the spectrum to the other by increments of one channel. At each position the central wavelength, equivalent width and their respective variances were calculated. If the equivalent width of any putative line was found to be greater than S (taken to be 4 in all subsequent work) times σ_W then it was accepted as a true feature.

Such a method may detect features more than once at different relative positions of the selection window and line. In such circumstances it is clear that when only part of the line falls within the selection window (hereafter referred to as subsidiary identifications) the portion detected will have a lower equivalent width than the whole and its centroid will be biased away from the true line centre. We have thus developed a technique to 'merge' all multiple detections.

The first stage in this process is to identify which features detected are actually the same line. Since the maximum positional deviation of any subsidiary identification from the true line centre is $M/2$ channels, we therefore search the list of detections for two consecutive entries which are closer together than $M/2$ channels. If this happens the mean wavelength is calculated, the contribution of the components being weighted by their equivalent widths. We then test whether the next line in the detection list is closer to this mean than $M/2$. If so, the mean wavelength is redetermined to take account of the new member. This procedure of adding components is repeated until there are no more lines which fulfil the condition for inclusion. The mean line centroid is then taken to be the true position of the line, the equivalent width of which is the equivalent width of the strongest component, since this will generally occur when the selection window overlaps the line completely. Tests of this procedure have shown that it gives line positions to within the expected errors.

From the foregoing discussion it is clear that the assumption of Poisson statistics is vital to the correct working of the detection technique and if, for any reason, there are significant wings on the noise distribution then spurious lines may be detected. Some account may be taken of this by ensuring that the algorithm used will detect upward pointing features, which would not be expected, as well as absorption lines. An excess of spikes over the number expected from Poisson statistics would then indicate that the data were unsuitable for this method of analysis.

3 The model spectra and calibration

The purpose of an automatic line detection technique is to select features in an objective manner without missing any and yet being totally insensitive to spurious ones. Obviously in practice these aims cannot be realized and it is therefore necessary to determine the efficiency of line detections and also the proportion of those found which are spurious. It is also necessary to determine whether factors such as local continuum slope affect the rate of detections. In order to test these we have generated an idealized quasar spectrum which mimics as closely as possible the mean continuum and emission-line properties of the quasar population. The continuum is represented by a power law of index 0.6 while the emission lines are a combination of a Gaussian core with wings given by the first Exponential Integral (E1) (Capriotti, Foltz & Byard 1980).

As quasars exhibit a large range in emission line widths we initially constructed a number of models with differing dispersion velocities (*cf.* Capriotti *et al.* loc cit.) in the range 5000–15000 km s^{-1}. When these were subjected to analysis for absorption lines no significant difference was found and so a mean dispersion velocity of 10000 km s^{-1} was adopted thereafter. It must be noted, however, that for very sharp emission lines (which are not common in high-redshift quasars) there may be some loss of efficiency of detection for absorption lines in the emission wings due to difficulties with continuum fitting, but this is difficult to quantify.

On to this spectrum were superimposed a set of approximately 60 absorption lines of varying equivalent widths and with Gaussian profiles matching the detector resolution.

Clearly, as the 'true' properties of this spectrum are known it is possible to examine the rate of spurious detections, etc. For this purpose we have generated 100 spectra, each based upon the computer-derived template but with Poisson noise added. The level of noise was determined by the object magnitude, m, and the sky brightness (taken to be 22 mag per arcsec^2 with a pixel area on the sky of 3 arcsec^2). Each of these spectra were analysed individually for absorption lines and

the resultant list compared with those known to exist. Such a comparison enables a table to be constructed of the number of times a line of a given equivalent width is detected. The ratio of the number of detections to the number of models is then the detection efficiency for a line of that equivalent width. Moreover, features which are detected but which do not correspond to lines in the template are spurious and are either statistical fluctuations or are produced by, for example, bad continuum fitting. We also measured the spurious detection rate, defined as the number of spurious features per 1000 spectral bins.

As a test we have compared the input line centroids and equivalent widths with those determined from the analysis of the model spectra. In all cases the difference between the observed and expected line positions was consistent with the derived value of $\sigma_{\langle\lambda\rangle}$ (*vide* Section 2.2). The situation regarding the equivalent widths of weak features is not so straightforward because of the bias induced by the detection limit. This causes a preferential selection when noise in the spectrum raises the line equivalent width. For stronger features, when this is less pronounced, the equivalent widths from the model spectra were found to have a Poisson distribution about the input equivalent width with standard deviation σ_W.

In order to test for any dependence of detection efficiency and spurious detection rate upon the continuum level, we have generated such sets of models with varying signal/noise (S/N).

A typical efficiency function is illustrated in Fig. 1, from which it is clear that the detection limit is well defined. In order to parameterize these curves the 10 and 90 per cent points, based upon a straight-line fit to the rising portion were measured. As a check on the constancy of these values another set of simulations were carried out which differed only in that they were based upon a new template spectrum. The agreement between both sets was very close and showed that at a resolution of 1.5 Å the efficiency levels may be set to within 0.07 Å. As a guide, the 10 and 90 per cent detection levels were placed at equivalent width=0.25 and 0.28 Å respectively for data with S/N of 40 (e.g. an object of m=16 observed for 10000 s on the AAT with the IPCS) and at equivalent width=1.08 and >1.50 Å for S/N=3 (e.g. a quasar of 19th magnitude with 1000 s

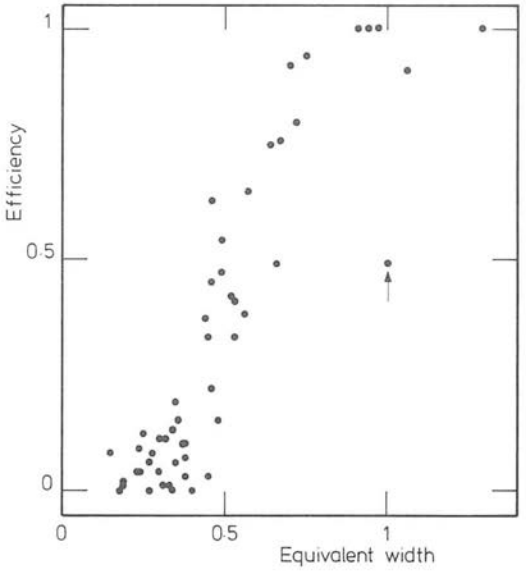


Figure 1. Representative plot of the detection efficiency of absorption lines as a function of equivalent width (Å) (S/N=8). The deviant point (arrowed) is caused by a close pair of lines which were sometimes interpreted by the technique as repeated identifications of the same feature.

integration). For quasars with $m=18.0$–18.5 and integration times of 7000 s as in this study, the appropriate equivalent widths were 0.38–0.48 Å and 0.80–1.06 Å.

The spurious detection rate was found to be constant at 0.35 per 1000 channels for S/N<15. For spectra with better S/N than this, small deviations between the true continuum and that fitted can produce spurious features. In practice, this is not a severe constraint because it may be removed by reducing the number of channels per segment, thus increasing the number of knots. No attempt, however, has been made to optimize this value for each model.

Finally, we constructed a template spectrum in which 20 absorption lines were placed at random in the emission lines and a further 20 were distributed in the continuum. The procedure of constructing a set of models of varying S/N was repeated and analyses carried out as before. Comparison was then made of the detection efficiencies for lines in the emission lines and in the continuum for each value of the S/N. The degree of similarity was determined by submitting the distributions to a Kolmogorov–Smirnov test. In only one of the 35 models was the probability that both efficiency distributions were drawn from the same parent population less than 75 per cent, while in 25 it was >95 per cent. It may, therefore, be concluded that the local continuum shape is not a determining factor in the detection rate.

4 The observations and preliminary reductions

The observations were made on the Anglo-Australian telescope during the nights of 1982 January 28–29 and 27–30 using the 25-cm camera and grating 1200 V, which gave a nominal spectral resolution of 1.5 Å. The IPCS was used in the two-dimensional mode, as the proximity of the two objects enabled both to be observed simultaneously through Dekker 50 (clear aperture). The journal of observations is given in Table 1.

Exposures, typically of 100 s duration, were made of the Cu–Ar arc every 25 min during the night so that any pixel shifts may be allowed for. Each spectral frame was reduced separately using the EDINBURGH SPECTRAL REDUCTION package on the Royal Observatory, Edinburgh node of the SERC STARLINK network. The wavelength scale was determined by a mean of arcs taken before and after each exposure. The mean rms error in calibrating the wavelength scale was only 0.1 Å. All measured wavelengths, and thus redshifts, are as *in vacuo* and have been corrected to heliocentric values. Finally, the wavelength response of the detector was removed by the empirical determination of this function from observations of the white dwarf star L970−30 and application of the appropriate correction. The reduced data from the individual frames were then combined; the resulting spectra are shown in Fig. 2(a–c).

Table 1. Journal of observations.

Object: 1146+11E				
RA	1950 Dec	Date	Wavelength region (Å)	Time (s)
$11^h\ 46^m\ 07^s.83$	11° 03′ 50″.6	1982 Jan 28–29	5010–6120	6000
		1982 Jan 29–30	4020–5050	7000
Object: 1146+11D				
RA	1950 Dec	Date	Wavelength region (Å)	Time (s)
$11^h\ 46^m\ 17^s.12$	11° 05′ 06″.9	1982 Jan 28–29	5010–6120	6000
		1982 Jan 29–30	4020–5050	7000

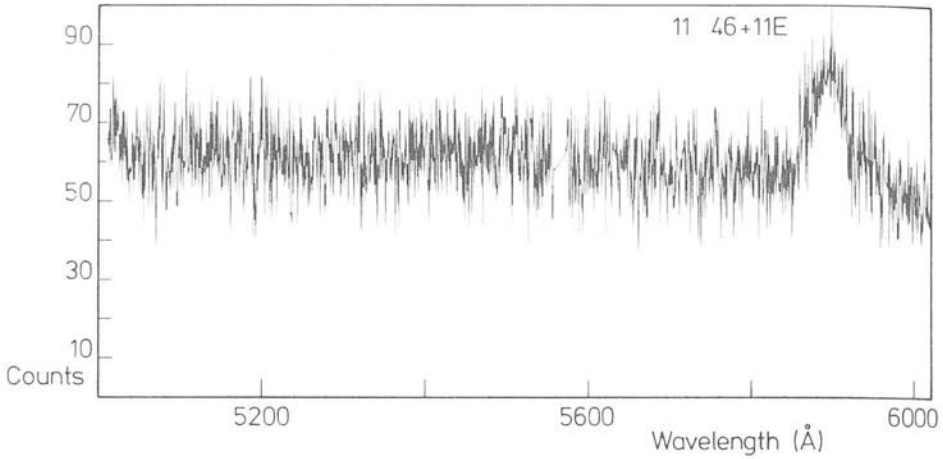


Figure 2. (a) Spectrum of 1146+11E produced from the addition of five integrations covering the range from 5000–6000 Å. No absorption lines were detected.

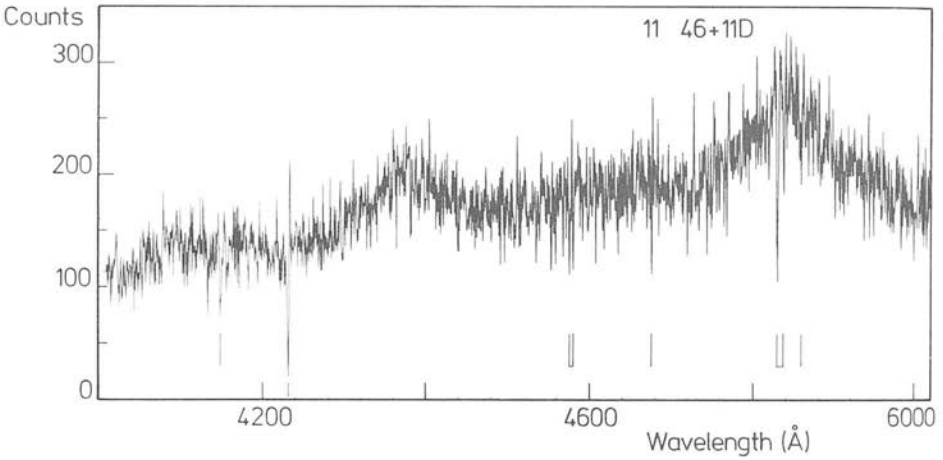


Figure 2. (b) Spectrum of 1146+11D produced from the addition of five integrations covering the range from 4000–5000 Å. The absorption lines found are marked.

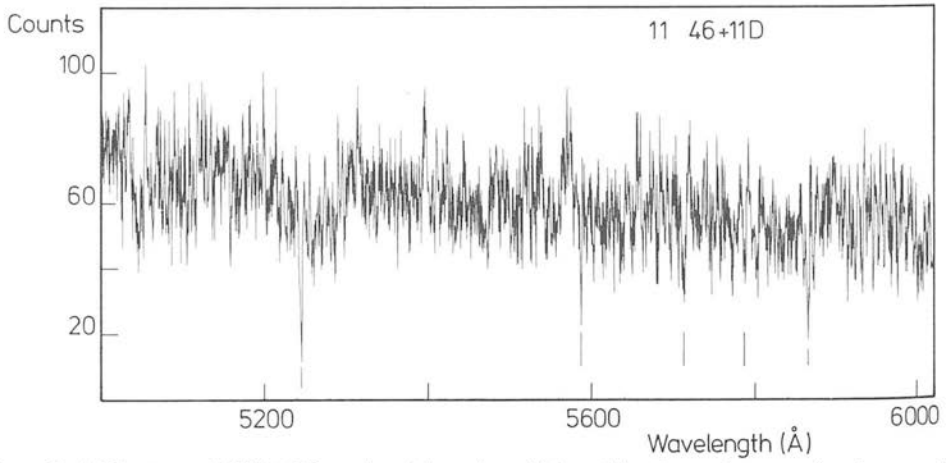


Figure 2. (c) Spectrum of 1146+11D produced from the addition of five integrations covering the range from 5000–6000 Å. The absorption lines found are marked.

5 The emission lines

The centroids of the emission lines were determined in two ways: first, by finding the centre of gravity of the line and secondly using a method initially developed by Schneider & Young (1980). This involved the convolution of the line profile with a function which is of the same width as the line and zero at the centre thus giving most weight to the sides of the line. The central wavelengths so determined are given in Table 2. Both methods give centroid positions which agree to within the formal errors, indicating that the emission lines in both quasars have relatively symmetric profiles.

In 1146+11D the identification of the emission lines is clear, the ratio of wavelengths being indicative of Si IV+O IV] and C IV. At this redshift it might be expected that the C III] line should be visible at 5960 Å and, although common in high-redshift quasars, an examination of Fig. 2b shows that no obvious feature is visible at this wavelength.

For the other object, 1146+11E, the existence of only one line means that the identification of the redshift must rely on the low-resolution data of Arp & Hazard (loc. cit.). Comparison of the spectra presented here with theirs allows us to identify the emission as Mg II at a redshift of 1.107.

6 Absorption redshift systems

The method used to identify the absorption lines was essentially that of Bahcall (1968). A set of standard lines commonly observed in quasar absorption systems was compared with the observed lines; a matrix of redshifts, in which each line was compared with every one in the standard package, was formed. Any absorption systems could then be seen as a peak in the redshift distribution.

The standard line set was taken from the list of Aaronson, McKee & Weisheit (1975), though only those transitions with rest wavelengths >1000 Å and which were considered to be strong lines by Aaronson *et al.* (loc. cit.) were included. Although some selection techniques (e.g. Aaronson *et al.*, loc. cit.) impose extra conditions based upon the relative strengths of the lines and ionizational considerations these were not applied here because of the relatively low signal/noise and paucity of the data. The wavelengths of the standard lines give a range of redshifts from 0.4 to 4.9, though in practice the upper redshift limit was set to be only somewhat greater than the emission redshift.

Not all peaks in such redshift distributions need represent true absorption systems, especially when the number of lines is small. Spurious features may be produced by the repeated identification of a single line with both members of a close doublet, e.g. C IV. In order to determine which peaks are real, the proposed identifications within each were checked. Any spurious systems can thus be removed.

Table 2. Emission lines.

Object: 1146+11E			
λ_{vac} (Å)	ID (Å)	Z	EW (Å) (observed)
5895	Mg II 2798	1.107	33±5
Object: 1146+11D			
λ_{vac} (Å)	ID (Å)	Z	EW (Å) (observed)
4376	Si IV 1397 +O VI 1402	2.121	45±6
4829	C IV 1549	2.118	51±6

7 Results

When the technique described in Section 2 was applied to the close pair of quasars, we found that there were no cases of a continuum segment having to be discarded (*cf.* Section 2.1). In practice, the rejection of segments is rare but may occur in regions where the continuum slope changes very rapidly or if the absorption line density is high.

Unfortunately, we also discovered that the data in the wavelength range 4000–5000 Å for quasar 1146+11E had numerous spikes of comparable significance to any putative absorption lines. We must therefore conclude that in this case the noise is not Poissonian and have thus not presented these data.

The discovery of non-Poisson noise is most disturbing as this invalidates the most basic assumption of the method described here. However, the other spectra of 1146+11D which were taken simultaneously with those of 1146+11E do not show this effect, nor does an analysis of the white dwarf L970–30. We must therefore conclude that, though inexplicable, the problem with these data is confined to only one object and that the checks built into the method are capable of detecting such problems.

For the remaining spectra the lines detected are listed in Table 3 and marked on Fig. 2(a–c). As may be seen from the table, only 30 per cent of the absorption lines have been identified in 1146+11D, while no lines were discovered in 1146+11E. In 1146+11D there is one identified redshift system – $z_{abs}=2.1203$. This system contains only three lines, a C IV doublet which is located close to the emission line peak and the Al III 1854 Å line. The implied expulsion velocity is only $\sim 100\ km\ s^{-1}$ and as such is a member of the peak in the C IV redshift distribution discussed by Weymann *et al.* (1981).

8 Conclusions

(1) Empirical testing of our method for finding absorption lines shows that those lines detected are true features at the preset significance level. Consequently, restriction of the sample to those lines above the 90 per cent efficiency point should provide a complete data set without contamination. Moreover, it has also been shown that the efficiency of detection for such lines is not dependent upon the local continuum slope. As such it is, therefore, of considerable importance for future objective studies of absorption lines in quasars.

(2) Application of this technique to the analysis of the spectra of a close pair of quasars has resulted in the discovery of a number of absorption lines. No absorption lines were discovered in the smaller data set from the lower redshift object (1146+11E) and as such there was no evidence for common absorption.

Table 3. Absorption lines in 1146+11D.

Wavelength	W (Å)	$\sigma_{(\lambda)}$(Å)	σ_W(Å)	ID λ(ID) (Å) z(ID)
4149.191	0.72	0.25	0.15	
4231.269	1.85	0.22	0.15	
4576.179	1.01	0.20	0.15	
4580.720	0.54	0.14	0.15	
4676.048	0.59	0.23	0.15	
4830.580	1.30	0.22	0.15	C IV 1548.20 2.1201
4838.333	0.74	0.15	0.15	C IV 1550.77 2.1200
4858.896	0.46	0.10	0.14	
5245.477	1.96	0.23	0.22	
5587.760	1.05	0.16	0.21	
5713.852	1.00	0.12	0.20	
5788.063	0.90	0.11	0.20	Al III 1854.72 2.1207
5868.018	1.42	0.10	0.20	

Although the number of absorption systems in 1146+11D is small, the distribution of redshifts is consistent with that found in studies of other objects (e.g. Weyman *et al.* 1981). Therefore, in those objects in which there is the greatest probability of a significant non-cosmological redshift component, there is no evidence for an abnormal relative velocity distribution.

Acknowledgments

We should like to thank Dr R. F. Carswell for his help and guidance at the time of observation. AST was in receipt of a postgraduate studentship for DENI. Our thanks also go to the Panel for the Allocation of Telescope Time for their allocation of time to Clube, Brand and Trew.

References

Aaronson, M., McKee, C. F. & Weisheit, J. C. 1975. *Astrophys. J.*, **198,** 13.
Arp, H. C. & Hazard, C. 1980. *Astrophys. J.*, **240,** 726.
Bahcall, J. N. 1968. *Astrophys. J.*, **153,** 879.
Bevington, P. R. 1969. *Data Reduction and Error Analysis for the Physical Sciences,* McGraw-Hill, New York.
Bruzual, G., 1981. *PhD thesis,* University of California, Berkeley.
Capriotti, E., Foltz, C & Byard, P. 1980. *Astrophys. J.*, **241,** 903.
Schneider, D. P. & Young, P. J. 1980. *Astrophys. J.*, **238,** 946.
Schoenberg, I. J. & Whitney, A. 1953. *Trans. Am. Math. Soc.*, **74,** 246.
Trew, A. S., Clube, S. V. M. & Savage, A. 1984. *Astr. Astrophys.*, submitted.
Weymann, R. J., Carswell, R. F. & Smith, M. G. 1981. *Ann. Rev. Astr. Astrophys.*, **19,** 41.
Young, P. J., Sargent, W. L. W. & Boksenberg, A. 1982. *Astrophys Suppl.*, **48,** 455.
Young, P. J., Sargent, W. L. W. & Boksenberg, A., Carswell, R. F. & Whelan, J. A., 1979. *Astrophys Suppl.*, **229,** 891.

Mon. Not. R. astr. Soc. (1988) **231**, 617–634

Shocked molecular hydrogen in the supernova remnant IC 443

M. G. Burton[1], T. R. Geballe[2,3], P. W. J. L. Brand[1] and A. S. Webster[4]

[1] *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

[2] *United Kingdom Infrared Telescope Unit, 665 Komohana Street, Hilo, Hawaii 96720, USA*

[3] *Foundation for Astronomical Research in the Netherlands (ASTRON)*

[4] *James Clerk Maxwell Telescope Unit, 665 Komohana Street, Hilo, Hawaii 96720, USA*

Accepted 1987 August 26. Received 1987 August 24; in original form 1986 June 11

Summary. Emission from the $v=1-0\ S(1)$ line of molecular hydrogen has been mapped over a section of the supernova remnant IC 443. The emission originates in a sinuous ridge where the expanding shell of the SNR is interacting with a molecular cloud. The relative intensities of the 1–0 $S(1)$, 1–0 $S(0)$ and 2–1 $S(1)$ lines at 2.1–2.2 μm were measured and found to be characteristic of shock-excitation of the gas. The ridge shows bright spots which are possibly density enhancements in the molecular cloud. The total luminosity of the molecular hydrogen lines in the mapped region is estimated to be about 1000 $L_{\odot}$, making IC 443 one of the most luminous galactic molecular hydrogen sources yet detected. The morphology of the $S(1)$ line emission is similar to that of the millimetre emission lines from CO, HCO^+ and HCN and to that of the 21-cm emission line from atomic hydrogen. There is evidence for partial dissociation of molecular hydrogen caused by the shock interaction, but the absence of Brγ line emission puts a low upper limit on the degree of ionization. The energetics of the SNR are consistent with most of the mechanical energy incident on the molecular cloud being radiated away through H_2 lines and/or being used to dissociate the H_2, except in the densest clumps (where shocks are either slow or non-existent), and where far-infrared emission from collisionally heated grains may provide the dominant cooling mechanism. A model for IC 443 is presented involving the expansion of the SN shock-wave inside the remains of a molecular disc left over from the process of formation of the star which exploded.

1 Introduction

On red sky-survey plates the supernova remnant (SNR) IC 443 appears as a bubble nearly a degree across, displaying a delicate filamentary structure. The interior of the bubble contains

tenuous X-ray emitting gas at a temperature of $\sim 2\times10^7$ K (Petre *et al.* 1983; Watson *et al.* 1983). A molecular cloud is known to be associated with the SNR because CO line emission has been detected from a region running NW–SE across the optical bubble (Cornett, Chin & Knapp 1977). Near the SE edge, a section of the shell of the SNR appears to be expanding into the molecular cloud, causing shock-excitation of the gas there. Evidence for this interaction was first presented by DeNoyer (1978), who found high-velocity neutral hydrogen (H I) line emission. This was followed by the discovery of shock-excited line emission from OH (DeNoyer 1979a), CO (DeNoyer 1979b) and molecular hydrogen (H_2) (Treffers 1979). More recently Braun & Strom (1986a) have made extensive maps of the H I 21-cm line emission and found that it takes the shape of an incomplete semi-circular arc containing small-scale structure. White *et al.* (1987) has partially mapped the CO, HCO^+ and HCN $J=1$–0 lines and finds them to originate in clumps with similar spatial distributions.

Shock-excitation is indicated by the broad widths (30–40 km s^{-1}, or more) observed in the H I, OH, CO, HCN and HCO^+ lines. In addition, there are distinct dynamical differences between neighbouring clumps, and the molecular abundances in the shocked gas are unusual (White *et al.* 1987), both of which may be taken as further indication of the action of a shock. We have mapped the distribution of the $v=1$–0 $S(1)$ line of molecular hydrogen (2.122 μm) in a section of the SNR and find the line emission to originate in a sinuous ridge. A number of emission peaks are located along the ridge. Where shock-excited lines have been observed in this region by others, their spatial distribution is found to be similar to that of the $S(1)$ line emission. Spectra were also obtained which demonstrate that the H_2 is indeed excited by shocks and which allow estimates of the extinction to, and the luminosity of, the shocked gas. It is clear from these results that IC 443 is one of the most luminous galactic H_2 emission line sources yet detected. The observations are described in Section 2 and the results in Section 3. Section 4 contains a discussion of the origin of the shocked gas, Section 5 presents a model for the SNR, and Section 6 discusses the cooling of the gas.

The generally accepted distance to IC 443 is 1500 pc (e.g. Sharpless 1965), although distances in the literature range from 500 pc (Malina, Lampton & Bowyer 1976) to 3000 pc (Duin & van der Laan 1975). The positional coincidence of IC 443 with the H II region S249 and the possible evidence for their association (Donati-Falchi & Tofani 1984; Fesen 1984) suggest that the distance to the star thought to be exciting S249, $d=1500$ pc, should also be taken for the distance to IC 443 (see e.g. Georgelin 1975). Optical studies of the filaments in S249 by Fesen (1984) also suggest that IC 443 is at a distance of 1500–2000 pc. In this paper we adopt 1500 pc as the distance to the complex.

2 Observations

The observations described in this paper were obtained with the 3.8-m United Kingdom Infrared Telescope (UKIRT), on Mauna Kea, Hawaii, during 1984 December and 1985 January. A map of the distribution of the H_2 $v=1$–0 $S(1)$ line was obtained by placing a Fabry–Perot spectrometer (FP) with $\sim$130 km s^{-1} velocity resolution in front of a cooled circular variable filter (CVF) tuned to transmit at 2.122 μm. An InSb detector was employed with an aperture diameter of 19.6 arcsec FWHM.

The map of the $S(1)$ line was obtained using the technique of frequency switching the FP. A reference frequency, near that of the $S(1)$ line, was chosen in an uncontaminated part of the atmospheric emission spectrum and the FP was switched between the frequencies of the $S(1)$ line and that of the reference. In this way contributions to the 2.122-μm flux from the sky and any (weak) continuum from the source were removed so that spatial chopping was not necessary. The problem of contamination due to H_2 line emission in the reference beam, often encountered when

observing diffuse extended sources, was thereby avoided. Changes in the sky emission spectrum with time, and drifts in the FP mirror spacing, necessitated periodic checks of the sky background level and of a standard spectrum. For the former, a sky reference position at (0′, −83′) [coordinates relative to a (0, 0) map position at $6^h\,14^m\,43\overset{s}{.}0$, +22° 23′ 00″ (1950.0)] was chosen; for the latter the peak $S(1)$ line flux, at (−20″, −20″), was checked frequently.

The integration time at each point was 16 s, of which 8 s were spent on the line frequency and 8 s on the 'off-line' frequency (in two cycles, each of 4 s on the line frequency and 4 s off). The grid had a 20-arcsec spacing between adjacent pixels so the resulting map is under-sampled. It is possible that localized bright spots have been missed because of the gaps in the spatial coverage. The FP 'on-line' frequency was set at that of the maximum emission from the line peak at (−20″, −20″). Subsequent observations (Burton 1987) show that velocity shifts are less than 40 km s^{-1} in the $S(1)$ line emission velocities at various positions along the emission ridge and that the linewidths are $\ll$130 km s^{-1}, the velocity resolution of the FP. We therefore believe that the line intensity distribution derived is accurate.

The resulting $S(1)$ line emission map, smoothed by a Gaussian filter of twice the beam size, is presented in Fig. 1. It contains about 800 points and has an effective resolution of 40 arcsec. Wavelength and flux calibrations were made on the peak of the 1–0 $S(1)$ line in OMC−1, which is $(3.5\pm0.5)\times10^{-11}$ erg s^{-1} cm^{-2} through the same aperture (Garden 1985, private communication). Absolute intensities are estimated to be reliable to about 20 per cent and the relative intensities of different pixels are accurate to about one unit (6.6×10^{-14} erg s^{-1} cm^{-2}) on the flux scale used on the map.

Low-resolution, fully-sampled spectra from 2.1 to 2.4 μm were obtained at three of the brighter positions. This wavelength range includes the H_2 1–0 $S(1)$, $S(0)$, and several Q-branch lines, and the 2–1 $S(1)$ line, as well as the H I Brγ line. These observations utilized a CVF (~2500 km s^{-1} resolution), a 19.6-arcsec aperture and the chopping secondary (120 arcsec E–W throw). Total integration times were about 30 min per position (including time spent chopping on to the sky). The intensities were calibrated by observing HD 84800 and HR 2241, assumed to have K magnitudes of 7.53 and 4.00, respectively.

A 2×2 arcmin^2 area, centred on (+1160″, +840″), was also searched for $S(1)$ line emission. This area covers part of the region containing optical filaments NE of the spherical shell of the SNR and is also situated on the southern side of the H II region S249. Fesen (1984) has suggested that these filaments, which display shock-excited optical emission lines, are physically related to the SNR. To the sensitivity limit of 5×10^{-14} erg s^{-1} cm^{-2} in a 19.6-arcsec beam, there is no $S(1)$ line emission in this region.

3 Results

3.1 MORPHOLOGY

Fig. 1 shows that the $S(1)$ line emission from the mapped region comes from an S-shaped ridge of low surface brightness containing several emission peaks. The ridge is about 2.5 arcmin wide and at least 22 arcmin long, which, at the assumed distance of 1500 pc, corresponds to ~1 pc by at least 10 pc. There are 11 peaks resolved along the ridge, each typically 1 arcmin in diameter (~0.5 pc). Table 1 lists the position of these clumps and their $S(1)$ line fluxes. The greatest flux, 3.1×10^{-12} erg s^{-1} cm^{-2} through a 19.6-arcsec aperture, was measured at $6^h\,14^m\,41\overset{s}{.}7$, +22° 22′ 40″ (1950.0). For zero extinction, this corresponds to a column density in the $v=1$, $J=3$ level of the hydrogen molecule of 1.7×10^{16} molecule cm^{-2}. The location, at (−20″, −20″) relative to the map zero point, is 5 arcsec west and 50 arcsec south of the position observed by Treffers (1979). Allowing for the larger beam used by Treffers, the flux we measured at his position is roughly consistent with his measurement.

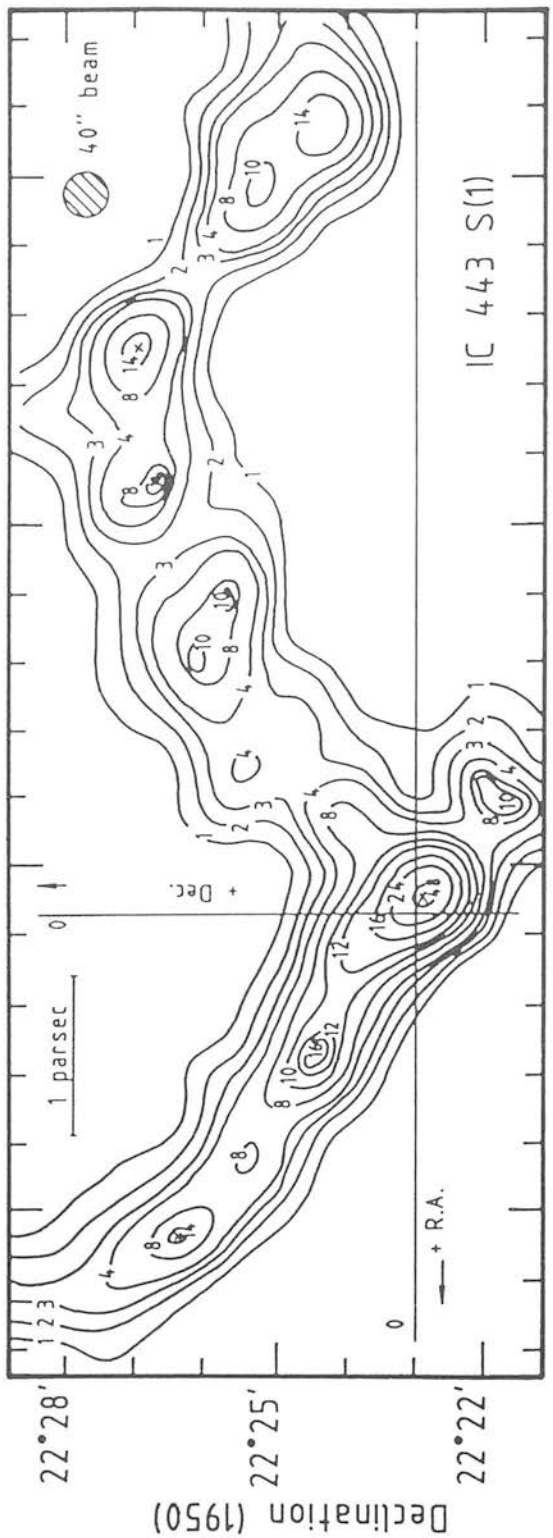


Figure 1. Contour map of the $v = 1\text{–}0\ S(1)$ emission line intensity of molecular hydrogen at 2.122 μm in IC 443. The velocity range observed is 130 km s^{-1} (FWHM), centred on the velocity of the emission peak at (−20″, −20″). The map resolution is 40 arcsec. The (0, 0) position is marked and crosses denote positions where the CVF spectra (Fig. 3) were observed. The distance scale bar assumes that the source distance is 1500 pc. The intensity levels are marked in multiples of the outermost contour; 1 unit corresponds to 6.6×10^{-14} erg s^{-1} cm^{-2} through the 19.6-arcsec beam.

Table 1. H_2 $v=1-0$ $S(1)$ line fluxes for bright peaks in IC 443.[a]

RA offset[b] (arcsec)	Dec offset[b] (arcsec)	Flux[c] (10^{-13} erg s^{-1} cm^{-2})
−20	−20	32
+280	+200	9.2
+220	+140	5.3
+120	+80	11
−100	−80	7.9
−220	+180	6.6
−280	+140	6.6
−380	+210	6.6
−500	+230	9.2
−620	+130	6.6
−690	+70	9.2

[a]The line was measured through the CVF+FP and a 19.6-arcsec aperture.
[b]Offsets refer to a (0, 0) position at 6^h 14^m 43^s, +22° 23′ 00″ (1950.0).
[c]Absolute fluxes are thought to be accurate to ±20 per cent. Relative fluxes are accurate to ±10 per cent.

The three locations of shocked CO line emission reported by DeNoyer (1979b) (and labelled as positions A, B and C by her – see Plate 1a) are linked by the H_2 line emission ridge. Position C coincides with the peak region of the H_2 line emission at the southern bend of the ridge, position B is at the northern bend of the ridge, and position A is the westernmost peak found. White *et al.* (1987) has mapped the CO, HCO^+ and HCN line emission along that section of the ridge which lies NE of the H_2 line peak. Fig. 2 shows his map of CO line emission overlaid on the H_2 line emission map around the peak. The shocked CO lines are emitted from clumps situated along the ridge of H_2 line emission, but the CO data, which were taken with higher spatial and velocity resolution than the H_2 data, resolve the peak region into more clumps than are apparent from the H_2 line map. The CO line map with the closest morphological similarities to the H_2 map is that for the velocity range $V_{LSR}=-20$ to -10 km s^{-1}. The CO spectrum taken at the position where the CO line emission peaks, at (−25″, −20″), has a broad velocity maximum between about −20 and

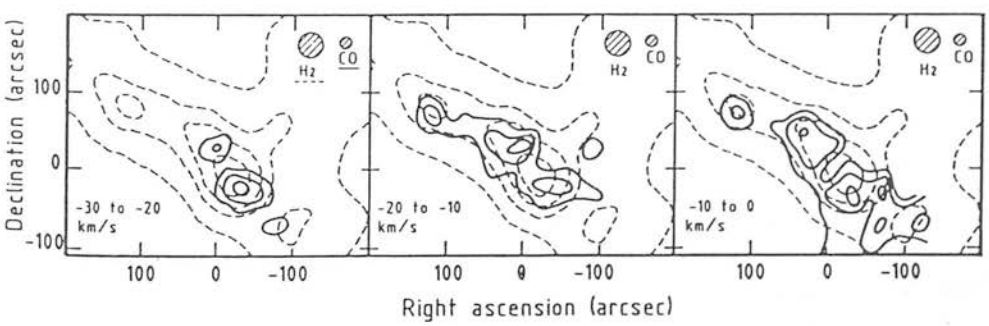


Figure 2. Maps of the accelerated CO line ($J=1-0$) (bold lines) observed by White *et al.* (1987) overlaid on the H_2 line map (dashed lines) of the peak emission region. The CO line maps are shown in three velocity bins; 0 to −10 km s^{-1}, −10 to −20 km s^{-1} and −20 to −30 km s^{-1}. The beamsize for the CO data is 15 arcsec and for the H_2 data 40 arcsec. Offsets are referenced to the (0, 0) position of the H_2 line map (Fig. 1), at 6^h 14^m 43^s.0, +22° 23′ 00″ (1950.0). The spatial coverage of the CO line emission is less than that of the H_2 line emission, being approximately the area enclosed by the second lowest H_2 contour in the maps.

$-46\,\mathrm{km\,s^{-1}}$, peaking at $V_{LSR}=-35\,\mathrm{km\,s^{-1}}$. This is similar to Treffers' value of $V_{LSR}=-33\pm5\,\mathrm{km\,s^{-1}}$ for the maximum intensity in the spectrum of the H_2 line emission. The total velocity-width of the CO line emission is greater than $70\,\mathrm{km\,s^{-1}}$.

The S-shaped ridge of emission also is clearly seen in the H I line maps by Braun & Strom (1986a) in the velocity range -43 to $-18\,\mathrm{km\,s^{-1}}$ V_{LSR}; H I line emission from the peak position can be detected out to $V_{LSR}=-150\,\mathrm{km\,s^{-1}}$, but high background levels prevent the comparisons from being made to velocities more positive than $-18\,\mathrm{km\,s^{-1}}$. The velocity of the bulk of the neutral gas in the vicinity of IC 443 is $-3.5\,\mathrm{km\,s^{-1}}$, by comparison (Cornett *et al.* 1977). Like the $S(1)$ line emission, the 21-cm line emission mapped by Braun & Strom comes from several peaks distributed along the ridge of emission. Their H I line map covers a larger region than the H_2 line map, and high-velocity H I line emission is detected outside the region of the H_2 line map. In the region of overlap of the two maps, the H I and H_2 line emissions are morphologically similar. The similarities, and detailed differences, between the two maps are discussed further in Section 4.4.

Due to the different beamsizes used for mapping the various emission lines, and the different spatial coverages of the various maps, a precise comparison of the intensity distributions of the lines is not possible. The data for the high-velocity species present are, however, consistent with their origin in the same spatial locations, to within 0.3 pc (corresponding to the 40-arcsec resolution of the H_2 map).

3.2 EXCITATION CONDITIONS

The excitation of the $S(1)$ line emission has been investigated by measuring several different H_2 lines. From the CVF spectra shown in Fig. 3, H_2 line intensities were derived (by Gaussian decomposition) and are listed in Table 2; important line-ratios are given in Table 3. The figures measured at the peak of OMC−1 on the same night are given for comparison. The 1–0 $S(1)$ line fluxes agree well with the fluxes measured through the CVF with the FP. The relative intensities of the 1–0 $S(1)$, 1–0 $S(0)$ and 2–1 $S(1)$ lines are 13.5:2.9:1.0 on the peak and are similar at the

Table 2. H_2 line fluxes and column densities at selected positions.

RA (offsets)	Dec		1–0 $S(1)$ 2.122 μm	1–0 $S(1)$ 2.223 μm	2–1 $S(1)$ 2.248 μm	Q-branch 2.41–2.42 μm
−20	−20	Flux	31	6.7	2.4	67
		Error	1	0.5	0.5	1
		N_{upper}	17	5.3	0.97	
+280	+200		9.3	1.9	0.8	20
			0.3	0.2	0.2	1
			5.1	1.5	0.3	
−500	+230		8.2	1.9	0.7	19
			0.3	0.2	0.2	1
			4.5	1.5	0.3	
0	0		18	3.9	1.5	39
			9.6	3.0	0.61	
OMC−1	Peak 1		342	84	34	780
			5	5	5	10
			190	66	14	

First row for each position lists the flux ($10^{-13}\,\mathrm{erg\,s^{-1}\,cm^{-2}}$), measured through a 19.6-arcsec aperture. The second row lists the statistical $\pm1\sigma$ error (in the same units). The third row lists the upper-level column density/$10^{+15}\,\mathrm{cm^{-2}}$, uncorrected for extinction. Only one measurement was made at position (0,0) and so the statistical error for that position cannot be calculated.

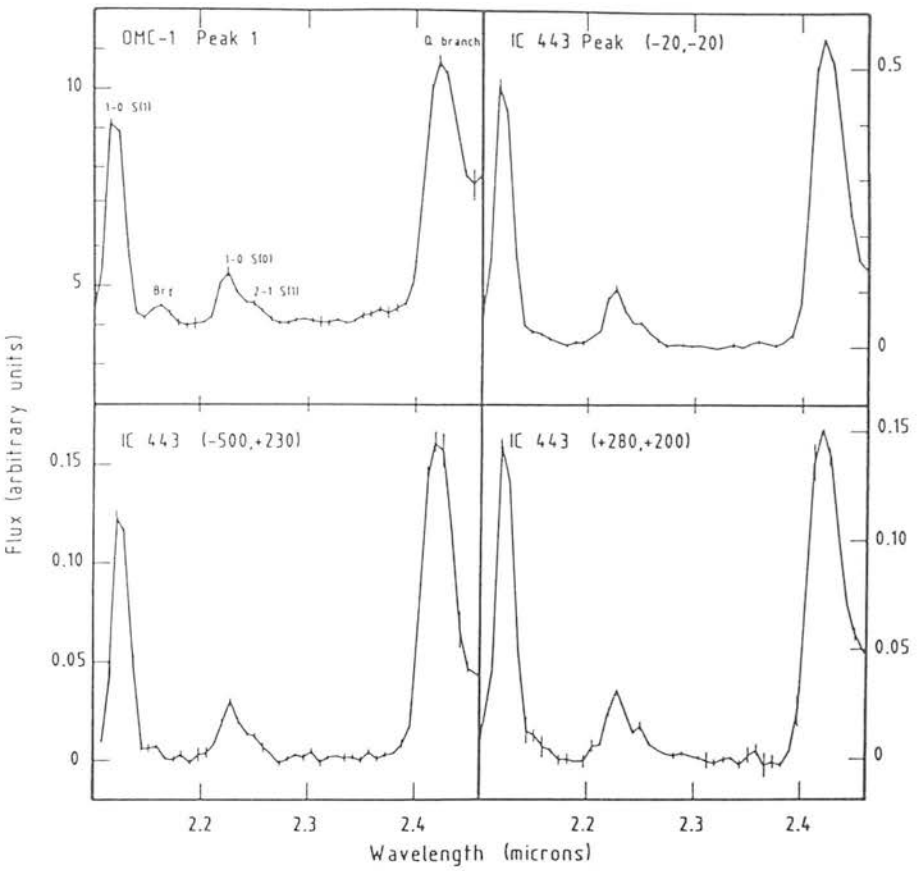


Figure 3. CVF spectra of the H_2 emission lines from 2.1–2.4 μm in OMC−1 and in IC443. The aperture size is 19.6 arcsec and spectral resolution 0.85 per cent. Important lines are labelled. The fluxes are in arbitrary units, with absolute fluxes listed in Table 2. Spectra are shown for three different positions in IC443, with positions given in arcsec from a (0,0) position at $6^h14^m43^s.0$, $+22°23'00''$ (1950.0). Error bars denote the $\pm 1\sigma$ error level.

Table 3. H_2 line intensity ratios at selected positions.

RA	Dec	1–0/2–1 $S(1)$	1–0 $S(0)$/2–1 $S(1)$	1–0 $S(1)/S(0)$	$Q/S(1)$
−20	−20	13.5	2.9	4.7	2.16
		[10.3–16.8]	[2.1–3.8]	[4.2–5.2]	[2.1–2.3]
+280	+200	12.4	2.5	4.9	2.15
		[9.0–16.0]	[1.7–3.5]	[4.3–5.6]	[2.0–2.3]
−500	+230	12.7	2.9	4.4	2.3
		[8.8–17.0]	[1.9–4.2]	[3.8–5.0]	[2.1–2.5]
0	0	12.0	2.6	4.6	2.2
OMC−1	Peak 1	10.3	2.5	4.1	2.28
		[8.6–12.0]	[2.0–3.1]	[3.8–4.4]	[2.21–2.34]

The limits on the ratios are calculated from the 1σ errors of the line fluxes.

other positions. They are also similar to the ratios observed in OMC−1 (10.3:2.5:1.0) and several other sources (e.g. NGC 2071, CRL 618, HH 7 – see Burton *et al.* 1988) where shock-excitation of the gas is believed to be responsible for the H_2 line emission. Such line ratios are predicted by theoretical models for typical shock velocities of 10–20 km s^{-1} and pre-shock densities greater

than about $10^5\,cm^{-3}$ (e.g. London, McCray & Chu 1977; Shull & Hollenbach 1978). Models using other mechanisms for the excitation tend to produce substantially different line ratios. Fluorescent excitation models, for instance (e.g. Black & Dalgarno 1976), give a ratio of 1.8 for 1–0 $S(1)/2$–1 $S(1)$, a factor of 7 smaller than observed in IC 443; line ratios of this order have been observed in the Orion Nebula Bar, for example (Hayashi *et al.* 1985). Our spectra, together with the positional coincidence of the H_2 line emission with the high-velocity H I, CO, HCN and HCO^+ species, indicate clearly that shock-excitation is responsible for the H_2 line emission.

The H I Brγ line, at 2.166 μm, is not detected in the CVF spectra at any of the four positions observed. Adopting a conservative 3σ upper limit, the Brγ line flux is $<3\times10^{-14}\,erg\,s^{-1}\,cm^{-2}$. This implies a low level of ionization, and will be discussed in Section 4.2.

3.3 MOLECULAR HYDROGEN LINE FLUXES

The total $v=1$–0 $S(1)$ line flux measured from the section of IC 443 mapped is $1.5\times10^{-10}\,erg\,s^{-1}\,cm^{-2}$. This corresponds to a line luminosity of

$$L_{S(1)}\sim11\left(\frac{d}{1500\,pc}\right)^2 L_\odot \tag{1}$$

and a mass of shocked gas in the $v=1$, $J=3$ level of

$$m_{1,3}\sim2.1\times10^{-4}\left(\frac{d}{1500\,pc}\right)^2 M_\odot \tag{2}$$

where d is the distance to the source. This calculation is based on the H_2 decay rates from Turner, Kirby-Docken & Dalgarno (1977), and on the H_2 energy levels of Dabrowski & Herzberg (1984). To convert these values into a total H_2 line luminosity, L_{total}, and total mass of shocked gas, m_{total}, the following quantities must be known: the fraction of the total emission through the 1–0 $S(1)$ line, the fraction of molecules in the $v=1$, $J=3$ level, and the extinction to the source. We now discuss these in turn.

3.3.1 Fraction of emission in the 1–0 S(1) line

If it is supposed that the 1–0 $S(1)$ line emission comes from gas at 2000 K (a common assumption, based on an excitation temperature derived from the 1–0 $S(1)/2$–1 $S(1)$ line ratio in the source OMC–1 – e.g. Scoville *et al.* 1982), then one-tenth of the total H_2 line emission comes from the $S(1)$ line. In practice, significant H_2 line emission takes place when the kinetic temperature lies between about 500 and 4000 K. Above 4000 K the dominant cooling process is the dissociation of hydrogen molecules, whereas below 500 K the upper vibrational–rotational levels of H_2 are not populated. If molecular hydrogen is the dominant coolant in the temperature range 500–4000 K (see discussion in Burton *et al.* 1987), then for every degree drop in temperature the same amount of energy is radiated away through the H_2 lines. If we suppose that the gas is approximately in local thermodynamic equilibrium as it cools over the above temperature range, then the fraction of the emission in the 1–0 $S(1)$ line may be calculated to be about one-fifteenth of the total line intensity. This gives a total H_2 line luminosity, not corrected for extinction, of 160 $L_\odot$.

3.3.2 Fraction of molecules in v=1, J=3 level

Using molecular constants from Huber & Herzberg (1979) the partition function for the hydrogen molecule at an excitation temperature of 2000 K is calculated to be ~48. From this, the fraction of

molecules in the $J=1$, $v=3$ level at 2000 K is determined to be

$$N_{1,3}/N \sim 1.36\times10^{-2}, \tag{3}$$

where $N_{1,3}$ is the column density of level (1, 3) and N the total column density of all levels. This gives a total mass of hot H_2, not corrected for extinction, of $1.5\times10^{-2}\,M_\odot$.

3.3.3 Extinction

The extinction to the H_2 line-emitting gas can be crudely estimated from the ratio of the $v=1$–0 Q-branch lines (2.4 μm) to the 1–0 $S(1)$ line (2.12 μm). The Q-branch lines [$v=1$–0 Q(1)–Q(7)] lie on the edge of the atmospheric window, and it is recognized that measurements of their intensity can be unreliable (e.g. Scoville *et al.* 1982). Experience at Mauna Kea shows, however, that, at low airmass, measurements near the Q(3) line are reasonably accurate, because the contributing H_2 lines [and especially the dominant Q(1) and Q(3) lines] are not seriously attenuated by telluric absorption (e.g. see Traub & Stier 1976).

The Q-branch and $S(1)$ lines were measured in the source OMC−1 and at several points in the SNR. In OMC−1 the ratio of the Q-branch lines to the $S(1)$ line is 2.3(±0.1) and in IC 443 it is 2.2(±0.2) at the four positions observed (see Table 3). We conclude that the extinctions at 2.2 μm are similar. In OMC−1, A_K is believed to be about 2 mag (e.g. Beckwith *et al.* 1983; Scoville *et al.* 1982) and so this value is adopted for IC 443.

Applying this extinction, the total luminosity and mass of the emitting gas are

$$L_{\rm total} \sim 1.0\times10^{3}\left(\frac{d}{1500\,\rm pc}\right)^{2} L_\odot, \tag{4}$$

$$m_{\rm total} \sim 9.7\times10^{-2}\left(\frac{d}{1500\,\rm pc}\right)^{2} M_\odot. \tag{5}$$

If $d=1500$ pc, this estimate for the total flux is five times the value for OMC−1 (Beckwith *et al.* 1983). It is in fact only a lower limit to the total flux from the SNR because we have not mapped the whole remnant. IC 443 is therefore undoubtedly one of the most luminous H_2 emission line sources yet observed in the Galaxy.

At the $S(1)$ line emission peak, (−20″, −20″), the column density of shock-heated H_2, corrected for extinction, is 7.7×10^{18} molecule cm^{-2}. This is one-eleventh of the greatest column density observed in OMC−1 on Peak 1. Table 4 lists the column densities of shocked H_2 at selected positions, allowing for 2 mag of extinction.

The pre-shock H_2 number density, n_0, can be estimated from

$$N_{\rm total} \approx \frac{n_0 V_s \tau_{\rm cool}}{\cos^2\theta}, \tag{6}$$

Table 4. Shocked H_2 column densities at selected positions.

RA	Dec	Column density (cm^{-2})
−20	−20	7.7×10^{18}
+280	+200	2.4×10^{18}
−500	+230	2.1×10^{18}
0	0	4.5×10^{18}
OMC−1	Peak 1	8.8×10^{19}

Column densities are derived from the intensity of the 1–0 $S(1)$ line and allow for 2 mag of extinction.

where N_{total} is the observed column density of the shock-heated H_2 gas, V_s is the shock speed (taken to be 30 km s^{-1}, the peak emission velocity of the shocked CO with respect to the quiescent molecular cloud), τ_{cool} is the cooling time for the gas (~0.3 yr) and θ is the angle between the shock velocity and the line-of-sight. At the $S(1)$ line emission peak we obtain $n_0 \sim 2.6 \times 10^5 / \cos^2\theta$ cm^{-3}. If dissociation is occurring, this is a lower limit.

For the peak emission region (the region between RA offsets −140″ to +140″), the integrated $S(1)$ line flux is $\sim 6 \times 10^{-11}$ erg s^{-1} cm^{-2}. This is equivalent to an H_2 luminosity of $\sim 400\ L_\odot$ under the assumptions made above. This luminosity will be compared with the SNR energy budget in Section 6.

4 The origin of the shocked molecular and atomic gas

The present data, coupled with the previous observations of other molecular and atomic lines, constitute clear evidence for the presence of a shock, driven by the expanding gas of a supernova remnant, within a molecular cloud. Taken together, these data relating to the distribution, dynamics and composition of the shocked gas provide an opportunity to determine the physics of this shock interaction in particular, and of such interactions in general.

The clumped appearance of the H_2, H I and CO line emissions suggests that there were pre-existing density inhomogeneities in the molecular cloud, and that the shock wave is simply illuminating this structure within the medium. The CO data show that the emission velocities of neighbouring clumps vary by 10–20 km s^{-1} (White *et al.* 1987; see also Fig. 2), which may be due to the clumps having been accelerated to different speeds by the shock. Alternatively, the variation may be a projection effect, the projection of the emission velocity along the line-of-sight changing with the position of the clump in the cloud. In addition, the density ϱ of the clumps will affect the shock speed V_s. Assuming that the dynamical pressure $P_{dyn} \sim \varrho V_s^2$ is approximately constant, then $V_s \propto \varrho^{-1/2}$, and the shocks in the high-density clumps will have low velocities, and vice versa.

The available data allow some conclusions to be drawn regarding the history of the line-emitting gas. In principle, the molecular gas might be shocked molecular material, or it might be gas which was dissociated by the shock and subsequently reformed on grains downstream. For the atomic gas there are three possibilities; it might be dissociated molecular gas, recombined H II in the swept-up shell of the remnant, or shocked ambient H I. In the following subsections we consider evidence for the following alternatives:

(i) the shock fully dissociates the molecules, which subsequently reform behind it;

(ii) the shock ionizes some of the molecular cloud material;

(iii) the shock occurs in a cloud containing significant fractions of both molecular and atomic gas;

(iv) the shock partially dissociates the molecules, yielding both shocked molecular gas and high-velocity atomic gas in a swept-up shell.

4.1 MOLECULE DISSOCIATION AND SUBSEQUENT REFORMATION BEHIND A FAST SHOCK

Behind a sufficiently fast shock, the hydrogen molecules will be collisionally dissociated. Molecular dissociation occurs when the post-shock temperature is high enough to drive collisional dissociation of the gas, but it is unclear what is the minimum shock speed required to achieve this. In a hydrodynamic shock, collisional dissociation of the H_2 molecules occurs for shock speeds $\gtrsim$25 km s^{-1} (e.g. Kwan 1977), whereas models which take into account the effects of a magnetic field raise this speed to $\gtrsim$45 km s^{-1} (e.g. Draine, Roberge & Dalgarno 1983).

After the dissociative shock has passed, the capture of the H atoms by grains can lead to H_2 formation. Subsequent ejection of the molecules results in the reformation of the molecular gas. The formation spectrum of the H_2 molecules is uncertain; Duley & Williams (1986) suggest that the H_2 will form in states with $v \leq 3$, $J \leq 1$ and kinetic energy <0.2 eV. The molecules are then redistributed via collisions to states with $v=0$ or 1 and $J<6$, and finally radiate the low-lying rotational–vibrational lines. The time-scales involved for the various processes between the shock front and the formation zone are sufficiently short ($=10^{17}/n$ s) that the different regions cannot be spatially resolved and the appearance would be given of molecular hydrogen surviving a fast shock and being detected in the 1–0 $S(1)$ emission line. In Duley & Williams' model, each H_2 molecule acquires about 2 eV of internal energy on formation. Taking the observed 1–0 $S(1)$ line luminosity (Section 3.3), the reformation rate required of molecular hydrogen is $\sim 4\times10^{-3} F_1/F_2\, M_\odot\,\mathrm{yr}^{-1}$, where F_1 is the fraction of the total $S(1)$ line emission which originates in reformed molecules, and F_2 is the fraction of total reformation energy released that is emitted in the 1–0 $S(1)$ line.

The observed H_2 line emission spectrum in IC 443 does not, however, correspond to the predictions of Duley & Williams' model but is typical of shock-excited sources. In particular, lines from the $v=2$ [2–1 $S(1)$] and $J=15$ [0–0 $S(13)$] levels are seen (Burton *et al.* 1988), which have high excitation energies and are therefore hard to account for in the reformation model as described. Moreover, if the shock were strong enough to have completely dissociated the molecular gas, it would be likely that an appreciable fraction would have been ionized as well, but the upper limit on the flux in the Brγ line (see next section) puts a low upper limit on the degree of ionization. We therefore believe it unlikely that the $S(1)$ line emission observed is due to this process.

4.2 THE IONIZATION OF THE GAS

If a shock is sufficiently strong, collisional ionization of the gas will occur behind the shock front. For a hydrodynamic shock, the post-shock temperature (for an adiabatic index $\gamma=\frac{5}{3}$ and mean particle mass $\mu=0.6\,m_{\mathrm{H}}$, m_{H} being the mass of the hydrogen atom) is given by

$$T\sim 1.4\times10^5 \left(\frac{V_s}{100\,\mathrm{km\,s^{-1}}}\right)^2 \mathrm{K} \tag{7}$$

where V_s is the shock speed. On the assumption of balance between collisional ionization and radiative recombinations, the fraction of H nuclei, X, that are ionized is given by

$$X \simeq \frac{1}{1+\Gamma/\alpha}, \tag{8}$$

where Γ is the rate coefficient for ionization and α the recombination coefficient back to the ground state (Spitzer 1978). Using rate coefficients for $H+H\rightarrow H^+ + e^- + H$ from Draine *et al.* (1983) and recombination coefficients from Spitzer (1978), we obtain $T\sim10^5$ K for $X=\frac{1}{2}$. This corresponds to a minimum shock speed for ionization of $V_s\sim85\,\mathrm{km\,s^{-1}}$. In a magnetically-mediated shock, the post-shock temperature is less than that of a hydrodynamic shock, so the shock speed required for ionization is still higher.

Recombination downstream of the shock will produce Brγ line emission at 2.166 μm. For an H II region which is optically thin in Brγ and for case B, the Brγ line flux is

$$F_{\mathrm{Br}\gamma}\sim 1.57\times10^{-4}\,\frac{n_e M(\mathrm{H\,II})}{d^2}\ \mathrm{erg\ s^{-1}\ cm^{-2}}, \tag{9}$$

where n_e is the electron volume density in cm^{-3}, $M(H\,II)$ the mass of the H II region in g, and d is the distance to the SNR in cm (Osterbrock 1974; Giles 1977).

The mass of the H II region may be estimated from

$$M(H\,II) \sim t_{rec}(\tfrac{3}{4}V_s) n_0 m_{H_2} d^2 \Delta\Omega \tag{10}$$

where t_{rec} is the recombination time-scale for the gas ($\sim 3\times10^{12}/n_e$ s), $\frac{3}{4}V_s$ the post-shock velocity, n_0 the pre-shock H_2 number density, m_{H_2} the mass of the hydrogen molecule, and $\Delta\Omega$ the beam size used (19.6-arcsec diameter), giving

$$F_{Br\gamma} \sim 4\times10^{-12}\left(\frac{n_0}{10^5\,cm^{-3}}\right)\left(\frac{V_s}{100\,km\,s^{-1}}\right) erg\ s^{-1}\ cm^{-2}. \tag{11}$$

Even allowing for 2 mag of extinction, this predicted flux exceeds our observational upper limit for Brγ line emission (Section 3.2). The lack of Brγ line emission is therefore inconsistent with the complete ionization of the gas behind a fast shock. The upper limit to the degree of ionization is approximately 10 per cent of the gas.

4.3 SHOCKED ATOMIC GAS

The H_2 line emission from IC 443 probably originates largely from dense shocked molecular clumps. If the H I line emission originates in coexisting interclump gas, the emission line velocities of shocked H I and H_2 should not be the same. The shock speed in the molecular gas would be expected to be slower than in the atomic material, due to the greater density there. The velocity data for the molecular emission (using the shocked CO as a tracer for the shocked H_2) are still too limited to allow a detailed comparison with the atomic gas; nevertheless, to the resolution available ($\sim 10\,km\,s^{-1}$), the CO and H I channel maps are similar. Clearly an important observation will be to obtain velocity channel maps of the H_2 line emission with high spatial resolution.

In addition, if 21-cm line emission does originate from atomic hydrogen in the ambient, unshocked cloud, then 21-cm line emission should also be radiated at the rest velocity of the cloud. This emission would, unfortunately, be confused with the high background levels of galactic H I line emission at this velocity. There is, however, an absence of 21-cm line emission at the molecular cloud rest-velocity, in the direction of the ridge of accelerated H I line emission.

In conclusion, there is at present no evidence to support the existence of large regions of atomic gas in the pre-shock medium.

4.4 THE PARTIAL DISSOCIATION OF MOLECULAR GAS BY A SHOCK

The emission regions in which the H_2 and H I lines originate appear to be remarkably similar. The emission ridges are coincident to within the resolution of the data (40 arcsec, corresponding to a spatial scale of $\sim$0.3 pc) and the locations of the peaks along the ridge are also coincident at this resolution (there is possibly a slight morphological difference between the H_2 and H I peaks at DeNoyer's position B), although the relative brightness of the H_2 and H I peaks varies with location. In particular, the H I line fluxes from positions A and B are similar to that from position C, whereas the H_2 line fluxes from positions A and B are about one-third of the flux from position C. The data therefore suggest that the radiating H_2 and H I species occupy the same regions of space when averaged over spatial scales $\geq$0.3 pc. This would be expected if the shock were strong enough to partially dissociate the molecular gas, leaving both shocked H_2 and accelerated H I gas behind it. The bulk of the shocked gas is known to be moving at velocities of 15–40 $km\,s^{-1}$ with respect to the quiescent molecular cloud (White *et al.* 1987), and the H I channel maps suggest that

there is material present with velocities up to 100 km s^{-1}. These velocities are large enough for the shock to produce dissociation.

A direct comparison of the H I column densities with the H_2 column densities has little meaning, because the H_2 lines are only emitted from a thin layer where the shock has just excited the gas, whereas the H I lines are emitted wherever atomic hydrogen exists along the line-of-sight. In addition, even if dissociation is occurring, there will be at least a small component of the atomic gas that is swept-up material from the interior of the SNR shell.

The degree of dissociation can be estimated by comparing the column densities of shocked CO and H I. At her position B, DeNoyer (1978, 1979b) obtains a column density of shocked H I of 6.9×10^{20} cm^{-2}, and for shocked CO of $\sim10^{16}$ cm^{-2}. The CO estimate is a lower limit, but is consistent with the value of 4×10^{16} cm^{-2} derived for the quiescent CO cloud (Cornett *et al.* 1977). On the assumptions that the pre-shock ratio $N_{CO}/N_H\sim7\times10^{-5}$ (Draine & Roberge 1984), and that this ratio remains constant in the post-shock gas, then $N_{H_2}\sim7\times10^{19}$ cm^{-2} in the post-shock gas. If it is assumed that all of the 21-cm H I line emission results from dissociated molecular gas, then these two estimates give a formal value of 80 per cent of the pre-shock molecular hydrogen having been dissociated into atomic hydrogen by the passage of the shock wave. This is obviously a very rough estimate, but we believe dissociation provides the simplest explanation for the high-velocity atomic gas, and that the degree of dissociation probably lies between 0.1 and 0.99.

The similar distributions of the H_2, H I and CO line emissions and the similar velocities of the shocked CO and H I line emissions support the hypothesis that partial dissociation takes place. The degree of dissociation might vary considerably within the molecular cloud. It is expected to depend on the local shock velocity at each position which, assuming an isobaric driving pressure behind the shock, is proportional to $\varrho_0^{-1/2}$. Thus the greatest degree of dissociation is expected to occur in the clumps of lowest density and the variations in the relative intensities of the H_2 and H I line emissions between the various peaks will reflect the degree of dissociation. One might infer, therefore, that at positions A and B the degree of dissociation is greater than at position C.

By assuming that dissociation is occurring, the lifetime of atomic hydrogen, τ_{HI}, behind the shock can be estimated from

$$\frac{N_{\text{line-emitting H}_2}}{N_{\text{shocked HI}}}\sim\frac{\tau_{\text{cool}}}{2\tau_{HI}}\cdot\frac{1-f}{f} \tag{12}$$

where τ_{cool} is the cooling time for shocked H_2 ($\sim$0.3 yr) and f the fraction of hydrogen molecules that are dissociated. At position B, $N_{\text{line-emitting H}_2}\sim2.1\times10^{18}$ cm^{-2}. This gives a value of $\tau_{HI}\approx5$–500 yr for a range of $f=0.1$–0.9. This time represents either the reformation time for molecular gas from the atomic gas or, when multiplied by the shock speed, the thickness of the compressed shell between the shock-front and an ionization-front (perhaps from the hot X-ray-emitting gas filling the interior of the SNR).

5 A model for IC 443

IC 443 is part of an extended cloud complex which includes the H II region S249 and members of the Gem OB1 association and contains several massive O and B stars. Braun & Strom (1986b) have shown that the complex contains three interconnected, spherical subshells with different radii and centres. The shocked molecules which have been observed lie on the southern rim of one of the subshells (subshell A in their notation). This subshell is delineated by the warm dust found by *IRAS*, and also by the optical filaments defining the NW boundary of the SNR, and by the accelerated H I and shocked H_2 emission lines defining the southern boundary.

A self-consistent picture of the SNR IC 443 can be inferred from the infrared data presented here, together with the X-ray, optical and radio data in the literature. Plate 1(a) shows the

relationship of some of the major emission features to each other. A schematic diagram showing where the emission lines originate in the proposed model is shown in Plate 1(b). The following features of the SNR evolution are assumed for IC 443:

(i) the SN blast wave has propagated outwards forming a bubble:

(ii) the hot X-ray-emitting gas filling the interior of the bubble (Petre *et al.* 1983; Watson *et al.* 1983) keeps it approximately in pressure equilibrium so that the force driving the blast wave is approximately isotropic;

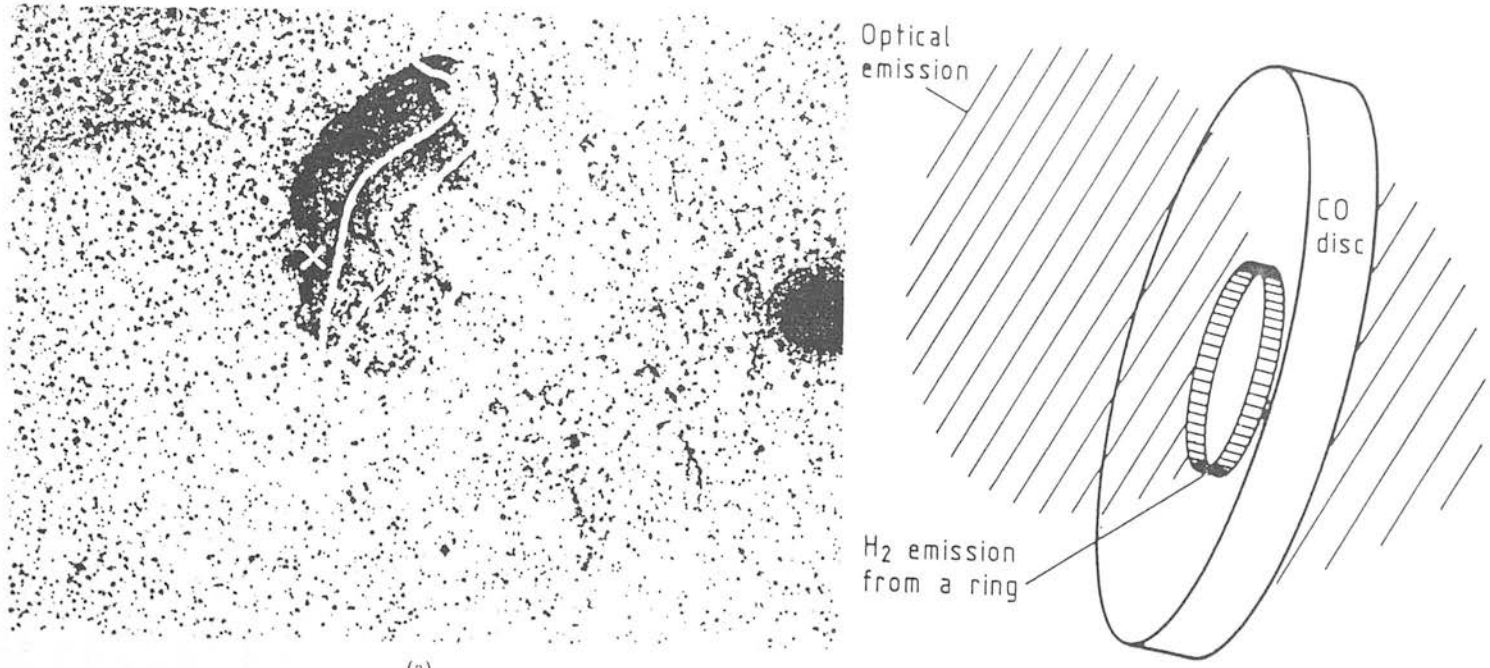
(iii) the SNR is now in a radiative phase, because the expansion velocities are less than 200 km s^{-1} (Woltjer 1972);

(iv) the quiescent molecular cloud running NW–SE across the face of the SNR (Cornett *et al.* 1977), and roughly bisecting the optical shell, actually contains much of the SNR, and the optical filaments in the NE and SW are located where the blast wave has broken out of the dense molecular cloud, shocking ambient atomic gas or dissociating and ionizing diffuse molecular gas.

A possible history of this region is that a molecular cloud collapsed preferentially along one axis, forming a planar, disc-like structure, and that a massive star formed within the disc. The star evolved rapidly, creating a stellar-wind-blown cavity around it, and then exploded as a supernova (probably a type II SN) before the original molecular cloud had been completely dispersed. The expanding shock wave compressed the gas in the disc, forming a shell on its inside surface, and then broke through the side of the cloud. The shocked molecular lines are being emitted from this shell, and are seen as a thin ring (see Plate 1b). The brightest emission observed is along that line-of-sight through the plane of the disc which takes in the greatest portion of the ring. If the ring is complete, shocked H_2 lines will also be observed in the NW portion of the remnant, diametrically opposite the peak of the shocked line emission in Fig. 1. The inner edge of the shell is bounded by an ionization front from the hot X-ray-emitting gas filling the interior of the SNR. The optical filaments are largely Hα line emission from recombining gas, and result from the expansion of the blast wave above and below the disc into lower density neutral gas. The optical emission is brighter on the NE-side than the SW-side, suggesting that the disc is tilted from the line-of-sight, obscuring emission from the SW-side. The quiescent CO cloud overlaps some of the filaments in the SW, supporting this assertion, but the filaments extend beyond the CO cloud and define a larger hemisphere in the SW than they do in the NE. As conjectured by Cornett *et al.* (1977), this may be due to the expansion in the SW being less impaired than in the NE because the density of atomic gas is lower in the SW. A higher density for the shocked gas in the NE than the SW may also explain why the X-ray and radio emission shows limb-brightening there.

This model for the SNR is different from previous models (e.g. Cornett *et al.* 1977; Braun & Strom 1986b). Cornett *et al.* propose that a sheet-like molecular cloud lies across the face of the SNR, but ~10 pc in front of where the pre-SN star was located. In their model the shell of the SNR expanded isotropically until it reached the molecular cloud, where the optical filaments in the NE are now located, and has heated the gas sufficiently to dissociate the molecular gas and ionize the neutral hydrogen. The southwestern part of the remnant expands more or less unimpaired into lower density material. The molecular line observations, however, show that the SNR is interacting with molecular gas in the SE portion of the remnant. The model by Cornett *et al.* seems to have two difficulties with it. First, it cannot explain why the H_2 line emission is confined to a narrow ridge. Secondly, if the shock were strong enough to ionize gas in the NE, it would likely be strong enough to ionize gas in the SE too, but our data do not show the presence of Brγ recombination line radiation coincident with the shocked gas. We therefore do not believe that their model can explain the observations.

In Braun & Strom's (1986b) model the subshells existing in the IC 443 complex were formed prior to the SN explosion by bubbles driven by stellar winds from members of the Gem OB1



(a) (b)

Plate 1. (a) Map of the ^{12}CO line emission from Cornett *et al.* (1977) overlaid on the red *Palomar Sky Survey* print of IC 443 (© National Geographic Society). The optical emission from the shell is mainly Hα line emission. The ridge of shocked H_2 line emission mapped is shown. Also indicated are the three locations (A, B and C) where DeNoyer (1979b) detected accelerated CO and the positions of the three bright 100-μm sources measured by *IRAS*. The H II region S249 is labelled. (b) Schematic diagram of a model for IC 443. This is a side view, looking at a plane through the molecular disc and optical shell. The SN has exploded within the remnant of a molecular disc left over from the formation of the pre-SN star. The H_2 line emission originates in a narrow ring where the blast wave is interacting with the clumpy molecular disc. Behind the shock lies accelerated molecular and atomic gas, the atomic gas being dissociated H_2. The optical (mainly Hα) results from the expansion of the blast wave above and below the disc into lower-density neutral gas. A cavity filled with hot low-density material is left behind and emits X-rays. This material keeps the expansion in approximate pressure equilibrium.

association. The SN exploded within one of the subshells and the blast wave propagated rapidly through its interior before encountering high-density molecular gas at the walls of the cavity. The shell of shocked molecular line emission arises from this gas. The model by Braun & Strom is not inconsistent with the model presented in this paper, if it is assumed that a cavity was created within the molecular disc prior to the SN explosion, presumably by wind from the star which exploded.

6 The cooling of the gas

6.1 LINE EMISSION AND DISSOCIATION

Nearly all of the emission from IC 443 at optical and near-infrared wavelengths is line emission. The optical filaments seen on the Palomar *R* plate are predominantly Hα line emission because they are barely detectable on the *I* plate. Molecular hydrogen lines dominate the emission in the near-IR in the SE portion of the SNR since the CVF spectra of Fig. 3 contain negligible contributions from a continuum or from other lines.

The data can be used to investigate the importance of H_2 line radiation as a coolant in the shocked molecular gas. The mechanical energy of the blast wave is converted to internal energy in the molecules and then radiated in lines; it also provides energy to dissociate some of the molecules. The total mechanical luminosity, L, is given by

$$L=\int \tfrac{1}{2}\varrho_0 V_s^3 \, dA \tag{13}$$

and the pressure behind a strong shock is

$$P\sim \frac{2}{\gamma+1}\varrho_0 V_s^2. \tag{14}$$

The uniformity of the X-ray emission within the SNR suggests the expansion is approximately in pressure equilibrium all over the remnant. Assuming that $P\sim P_{\text{X-ray}}$, the mechanical luminosity per unit area is

$$\frac{L}{\Delta A}\sim \frac{\gamma+1}{4} P_{\text{X-ray}}\langle V_s\rangle \tag{15}$$

where ΔA is the area of the shock front. Taking $\langle V_s\rangle\sim 30\,\text{km s}^{-1}$, as before, and $P_{\text{X-ray}}\sim 1\times 10^{-8}$ dyne cm^{-2} (derived from an X-ray temperature of $\sim 2\times 10^7$ K (Watson *et al.* 1983) and an ambient density of $\sim 0.5\,\text{cm}^{-3}$ (Malina *et al.* 1976), then the mechanical luminosity delivered per unit area of shock front is $\sim 2.0\times 10^{-2}\,\text{erg s}^{-1}\,\text{cm}^{-2}$.

The total H_2 line luminosity from the peak emission region is $\sim 400\,L_\odot$ (Section 3.3). For the ring geometry of the previous section, the shocked area, ΔA, is estimated to be $\sim 10\,\text{pc}^2$. We do not expect that this estimate will be greatly different for other geometries. The minimum area for the shock front, which occurs if we are viewing it face on, is $\sim 4\,\text{pc}^2$.

In the present model, the H_2 line luminosity per unit area in the peak emission region is $\sim 1.6\times 10^{-2}\,\text{erg s}^{-1}\,\text{cm}^{-2}$, and is comparable to the estimate of the mechanical luminosity per unit area delivered to it. These estimates are subject to several uncertainties, in particular the assumed geometry for emission, the X-ray pressure and the H_2 line luminosity (through the distance and extinction to the source). The implication is, nevertheless, that molecular hydrogen is an important coolant of the gas. In other words, a significant fraction of the SN blast energy driving the shock into the molecular cloud has been converted into internal energy of the hydrogen molecules and then radiated away in the H_2 emission lines. Since the degree of molecular dissociation

appears to be high, a significant fraction of the remaining mechanical energy also must go into dissociating the molecules.

In the above discussion, it is assumed that the efficiency of conversion of mechanical energy to internal energy is unity. This is equivalent to assuming that the gas remains in LTE behind the shock front. It can be shown (e.g. Fischer *et al.* 1985) that, for an isothermal shock, the rate at which energy is radiated is equal to the rate at which kinetic energy is deposited in the gas.

6.2 FAR-INFRARED CONTINUUM EMISSION – THE *IRAS* SOURCES

Three bright 100-μm *IRAS* sources appear to be associated with IC 443 (Huang, Dickman & Snell 1986). The 100-μm sources have fluxes ~40 Jy, colour temperatures ~35 K and luminosities ~100 $L_\odot$. The sources show good positional coincidence with clumps of highly perturbed gas. Huang *et al.* speculate that the sources may be cool protostars, whose collapse was induced by the SN shock wave. We shall show below, however, that it is possible to account for the 100-μm luminosity through collisional heating of the dust grains behind the shock, with the grains reradiating the energy in the far-infrared.

The total heating rate of the grains is given by

$$H \sim \varepsilon n_{\rm gas} n_{\rm grain} \langle \sigma v \rangle \langle \tfrac{1}{2} m v^2 \rangle V \tag{16}$$

where $n_{\rm gas}$ and $n_{\rm grain}$ are the gas and grain number density, $\langle \sigma v \rangle$ the collision rate coefficient, $\langle \frac{1}{2} m v^2 \rangle$ the mean kinetic energy per gas particle, V the volume containing the grains, and ε an efficiency factor. The total grain mass is $M_{\rm grain} = n_{\rm grain} m_{\rm grain} V$, where $m_{\rm grain}$ is the mass of an individual grain.

Following Campbell, Hoffman & Thronson (1981), if a grain emits as a blackbody, then

$$M_{\rm grain} \sim \frac{F_\nu}{B_\nu(T)} \frac{d^2}{\varkappa} \sim 2.5 \times 10^{-3}\, M_\odot, \tag{17}$$

where F_ν is the 100-μm flux at distance d, $B_\nu(T)$ the Planck function and $\varkappa \sim 250\,{\rm cm^2\,g^{-1}}$ the mass absorption coefficient (Campbell *et al.* 1981). Taking the density of a grain to be $1\,{\rm g\,cm^{-3}}$ and the grain radius $a \sim 0.1\,\mu$m, the heating rate is

$$H \sim \varepsilon \left(\frac{n_{\rm gas}}{10^6\,{\rm cm^{-3}}} \right) \left(\frac{v}{10\,{\rm km\,s^{-1}}} \right)^3 80\, L_\odot. \tag{18}$$

The 100-μm luminosity (~100 $L_\odot$) can thus be accounted for by gas–grain collisions if the kinetic energy conversion efficiency is high. Since only three 100-μm sources are seen, this argument cannot apply to all the clumps. Collisional heating of dust should be dominant in the densest clumps ($>10^6\,{\rm cm^{-3}}$) where the shock velocity is low (and the H_2 barely excited), whereas collisional heating of hydrogen molecules should be more important in the less-dense clumps (e.g. $\sim 10^4\,{\rm cm^{-3}}$) (with correspondingly faster shock speeds).

Thus a consistent picture of the energetics and cooling in the shocked gas is that the bulk of the mechanical energy incident on the molecular cloud is radiated away through molecular hydrogen lines or goes into dissociating the molecules, except in the densest clumps, where far-infrared emission from collisionally heated grains may provide the dominant cooling mechanism.

7 Conclusions

(i) The H_2 1–0 $S(1)$ line has been mapped over a portion of the SNR IC 443, and found to come from a sinuous ridge where the supernova shock wave interacts with a molecular cloud. The ridge is about 1 pc across and 10 pc long or more, and contains at least 11 bright emission peaks.

(ii) Allowing for an extinction of 2 mag and for emission from all other levels, the total H_2 line luminosity over the mapped region is $\sim 1000\, L_\odot$. IC 443 is among the most luminous galactic H_2 emission line sources yet detected.

(iii) The spatial distributions of the line emission from high-velocity CO, HCO^+ and HCN molecules are remarkably similar to that of the H_2 $S(1)$ line. The distribution of 21-cm line emission from high-velocity atomic hydrogen is also similar to that of the $S(1)$ line. This observation, together with measurements of column densities of the H I, CO and excited H_2 and the similar velocities of the H I and CO gas, provide support for the conjecture that partial dissociation of molecular material to atomic gas is taking place.

(iv) The energetics of the SN explosion and the shock-heated molecular gas are consistent with the idea that line emission from and dissociation of the hydrogen molecules are the major cooling mechanisms of the gas.

Acknowledgments

We are grateful to the UKIRT staff for their support and encouragement in obtaining the observations and to PATT for allocating telescope time. MGB acknowledges the support of the SERC through a SERC studentship.

References

Beckwith, S., Evans, N. J. II, Gatley, I., Gull, G. & Russell, R. W., 1983. *Astrophys. J.*, **264**, 152.
Black, J. H. & Dalgarno, A., 1976. *Astrophys. J.*, **203**, 132.
Braun, R. & Strom, R. G., 1986a. *Astr. Astrophys. Suppl.*, **63**, 345.
Braun, R. & Strom, R. G., 1986b. *Astr. Astrophys.*, **164**, 193.
Burton, M. G., 1987. *Ql. J. R. astr. Soc.*, **28**, 269.
Burton, M. G., Brand, P. W. J. L., Geballe, T. R. & Webster, A. S., 1988. *Mon. Not. R. astr. Soc.*, submitted.
Campbell, M. F., Hoffmann, W. F. & Thronson, M. A., 1981. *Astrophys. J.*, **247**, 530.
Cornett, R. H., Chin, G. & Knapp, G. R., 1977. *Astr. Astrophys.*, **54**, 889.
Dabrowski, I. & Herzberg, G., 1984. *Can. J. Phys.*, **62**, 1639.
DeNoyer, L. K., 1978. *Mon. Not. R. astr. Soc.*, **183**, 187.
DeNoyer, L. K., 1979a. *Astrophys. J.*, **228**, L41.
DeNoyer, L. K., 1979b. *Astrophys. J.*, **232**, L165.
Donati-Falchi, A. & Tofani, G., 1984. *Astr. Astrophys.*, **140**, 395.
Draine, B. T. & Roberge, W. G., 1984. *Astrophys. J.*, **282**, 491.
Draine, B. T., Roberge, W. G. & Dalgarno, A., 1983. *Astrophys. J.*, **264**, 485.
Duin, R. M. & van der Laan, H., 1975. *Astr. Astrophys.*, **40**, 111.
Duley, W. W. & Williams, D. A., 1986. *Mon. Not. R. astr. Soc.*, **223**, 177.
Fesen, R. A., 1984. *Astrophys. J.*, **281**, 658.
Fischer, J., Sanders, D. B., Simon, M. & Solomon, P. M., 1985. *Astrophys. J.*, **293**, 508.
Georgelin, Y. M., 1975. *PhD dissertation*, Université de Provence, Marseille, France.
Giles, K., 1977. *Mon. Not. R. astr. Soc.*, **180**, 57P.
Hayashi, M., Hasegawa, T., Gatley, I., Garden, G. & Kaifu, N., 1985. *Mon. Not. R. astr. Soc.*, **215**, 31P.
Huang, Y.-L., Dickman, R. L. & Snell, R. L., 1986. *Astrophys. J.*, **302**, L63.
Huber, K. P. & Herzberg, G., 1979. *Constants of Diatomic Molecules*, Van Nostrand, New York.
Kwan, J., 1977. *Astrophys. J.*, **216**, 713.
London, R., McCray, R. & Chu, S.-I., 1977. *Astrophys. J.*, **217**, 442.
Malina, R., Lampton, M. & Bowyer, S., 1976. *Astrophys. J.*, **207**, 894.
Osterbrock, D. E., 1974. *Astrophysics of Gaseous Nebulae*, W. H. Freeman, San Francisco.
Petre, R., Canizares, C. R., Winkler, P. F., Seward, F. D., Willingale, R., Rolf, D. & Woods, N., 1983. In: *Supernova Remnants and their X-ray Emission*, p. 289, eds Danziger, J. & Gorenstein, P., Reidel, Dordrecht, Holland.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G. & Ridgway, S. T., 1982. *Astrophys. J.*, **253**, 136.

Sharpless, S., 1965. In: *Stars and Stellar Systems*, Vol. 5, *Galactic Structure*, p. 136, eds Blaauw, A. & Schmidt, M., University of Chicago Press.

Shull, J. M. & Hollenbach, D. J., 1978. *Astrophys. J.*, **220**, 525.

Spitzer, L., 1978. *Physical Processes in the Interstellar Medium*, John Wiley, New York.

Traub, W. A. & Stier, M. T., 1976. *Appl. Opt.*, **15**, 364.

Treffers, R. R., 1979. *Astrophys. J.*, **233**, L17.

Turner, J., Kirby-Docken, K. & Dalgarno, A., 1977. *Astrophys. J. Suppl.*, **35**, 281.

Watson, M. G., Willingale, R., Pye, J. P., Rolf, D. P., Wood, N., Thomas, N. & Seward, F. D., 1983. In: *Supernova Remnants and their X-ray Emission*, p. 273, eds Danziger, J. & Gorenstein, P., Reidel, Dordrecht, Holland.

White, G. J., Rainey, R., Hayashi, S. S. & Kaifu, N., 1987. *Astr. Astrophys.*, **173**, 337.

Woltjer, L., 1972. *Ann. Rev. Astr. Astrophys.*, **10**, 129.

THE ASTROPHYSICAL JOURNAL, 334:L103–L106, 1988 November 15

RATIOS OF MOLECULAR HYDROGEN LINE INTENSITIES IN SHOCKED GAS: EVIDENCE FOR COOLING ZONES

P. W. J. L. BRAND,[1] A. MOORHOUSE,[1] M. G. BURTON,[2] T. R. GEBALLE,[3,4] M. BIRD,[1] AND R. WADE[5]

Received 1988 February 3; accepted 1988 August 19

ABSTRACT

Column densities of molecular hydrogen have been calculated from 19 infrared vibration-rotation and pure rotational line intensities measured at peak 1 of the Orion molecular outflow. The run of column density with energy level is similar to a simple cooling zone model of the line-emitting region, but is not well fitted by predictions of *C*-shock models current in the literature.

Subject headings: infrared: spectra — interstellar: molecules — molecular processes — shock waves

I. INTRODUCTION

The discovery of line emission from shocked molecular hydrogen in the Orion molecular cloud OMC-1 (Gautier *et al.* 1976) stimulated a series of theoretical studies (Kwan 1977; London, McCray, and Chu 1977; Hollenbach and Shull 1977) of the structure of molecular shocks. Observations of an extremely supersonic range of velocities in the H_2 1–0 *S*(1) line profile (Nadeau and Geballe 1979) prompted several workers (Draine 1980; Chernoff, Hollenbach, and McKee 1982; Draine, Roberge, and Dalgarno 1983) to develop magnetically moderated *C*-shock models.

In this *Letter* we present data from the brightest part of the OMC-1 outflow, peak 1 ($5^h32^m46^s$, $-5°24'02''$ [1950] Beckwith *et al.* 1978). The data consist of a set of molecular hydrogen line intensities observed between 2 μm and 4 μm, which cover a wide range of upper level energies (6500–25,500 K).

II. OBSERVATIONS

All of the observations of peak 1 were made at the United Kingdom 3.8 m Infrared Telescope on Mauna Kea. Most of them utilized the facility cooled grating spectrometer. The beam diameter of this instrument was set to 5″; the resolving power was typically 500. Standard chopping and nodding (60″ EW) practices were employed. The stars BS 1552 and BS 1713 were used for flux calibration. All of the 3 μm lines were measured in 1985 November; most of the 2 μm lines were observed in 1987 January and February. In addition, during the latter period a number of the strong lines in the 2 and 3 μm bands were measured in a single scan so that their relative intensities could be determined. The reduced 2 μm spectrum from 1987 January is shown in Figure 1. The 3 μm spectrum has been published elsewhere (Geballe 1986).

Several 2 μm H_2 lines, including the 4–3 *S*(3) and 3–2 *S*(2) (which had not previously been observed), were measured at peak 1 in 1988 January. These data were obtained using the facility CVF spectrometer in series with an ambient temperature Fabry-Perot interferometer. The beam diameter was 12″, and the velocity resolution was ~ 120 km s^{-1}. Chopping and nodding practices were as above. The 1–0 *S*(1) line was measured with this instrument, so that the former two line intensities could be scaled to previously measured H_2 lines and thus be included in the analysis. The 4–3 *S*(3) line is the highest excitation H_2 line yet observed in the 2 μm band; its spectrum is shown inset in Figure 1 together with that of the 3–2 *S*(2) line.

III. ANALYSIS AND RESULTS

The intensities of lines measured by the cooled grating spectrometer were determined by least-squares fitting Gaussian profiles to the observed lines, together with polynomials to the continua. Because of pointing and beam size differences between the different observations sets, the spectra were scaled to one another using the previously described spectrum that contained lines common to the various 2 and 3 μm spectra. The resulting line intensities are presented in Table 1. We note that the relative intensities of the 2 μm lines measured by us are consistent with those observed by Oliva and Moorwood (1988).

Of the approximately 30 H_2 lines detected, only those 19 lines whose intensities were believed to be free from significant uncertainties (e.g., due to attenuation by telluric absorption lines) were considered for further analysis. Observed column densities, N_0 (assuming no extinction), were determined from the observed intensities using transition probabilities from Turner, Kirby-Docken, and Dalgarno (1977). The extinction may be estimated by comparing intensities of lines which arise from a common upper energy level. In the present data set the 1–0 *S*(1), 1–0 *Q*(3), and 1–0 *O*(5) lines may be used. Assuming $A_\lambda \propto \lambda^{-1.5}$ we determine that $A_K = 0.8 \pm 0.3$. The dereddened column densities in the last column of Table 1 assume this form of the extinction law. If it is assumed that $A_\lambda \propto \lambda^{-1.0}$, the derived extinction is $A_K = 1.0 \pm 0.3$. However, the derived ratios of dereddened column densities are not changed significantly from those in Table 1.

Figure 2 is a plot of the dereddened column density in each level divided by that predicted for a slab of gas at 2000 K. The data are presented in this way in order to show the deviation from such a constant temperature environment, which has often been considered to be characteristic of the H_2 line-emitting region. A Boltzmann distribution at any other temperature is a straight line in this diagram. The best-fit single excitation temperature for the data is ~ 2200 K, which is consistent with most previous observations of H_2 lines in OMC-1 (e.g., Knacke and Young 1981). However, it is clear that no straight line will be a satisfactory fit to all of the data points.

[1] Department of Astronomy, University of Edinburgh.
[2] NASA/Ames Research Center.
[3] Joint Astronomy Centre, Hilo, Hawaii.
[4] Foundation for Astronomical Research in The Netherlands—ASTRON.
[5] Royal Observatory, Blackford Hill, Edinburgh.

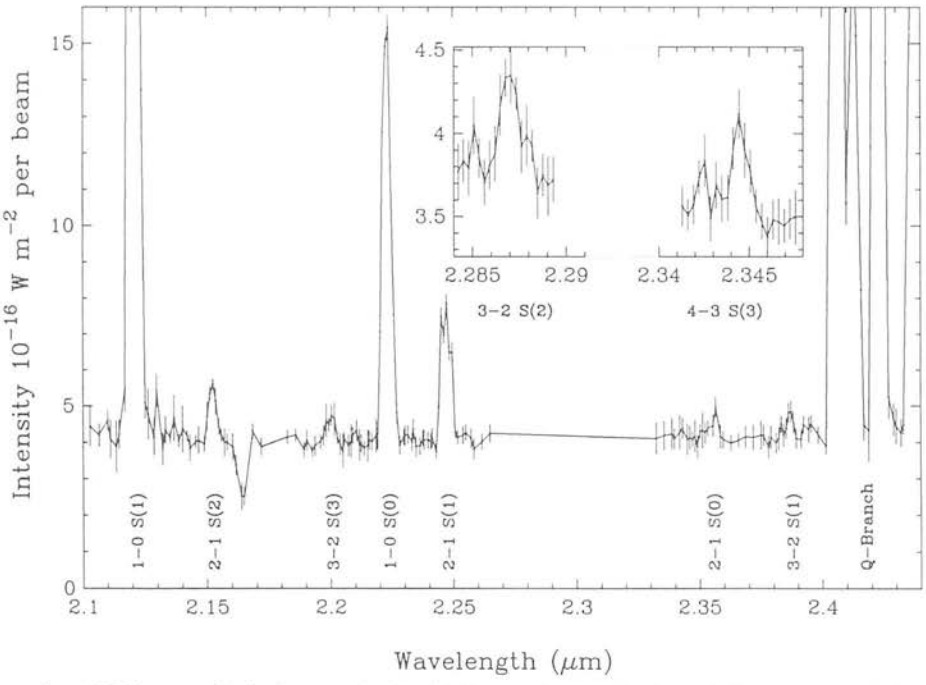


FIG. 1.—Raw 2 μm spectrum from 1987 January, obtained at a resolution of 0.004 μm. Lines used in the analysis are labeled. The feature at 2.166 μm is due to Brγ in the offset beam. The inset is the spectrum of the 3–2 $S(2)$ and 4–3 $S(3)$ lines, at a resolution of 0.0009 μm, and with the same flux scale as the main diagram.

TABLE 1

PARAMETERS AND INTENSITIES OF OBSERVED H_2 LINES

Line	Wavelength[a] (μm)	Upper Level Energy[a] (K)	Flux Density[b] (10^{-16} W m^{-2})	Dereddened Column Density[c]/g_J (10^{18} m^{-2})
1–0 $S(7)$[d]	1.7480	12818	4.2 ± 0.4	...
1–0 $S(1)$	2.1218	6956	50.1 ± 0.5	326.3 ± 2.9
2–1 $S(2)$	2.1542	13150	1.6 ± 0.1	14.8 ± 1.2
3–2 $S(3)$	2.2014	19086	0.8 ± 0.1	1.9 ± 0.2
1–0 $S(0)$	2.2235	6471	12.2 ± 0.2	467.5 ± 11.1
2–1 $S(1)$	2.2477	12550	4.3 ± 0.1	18.8 ± 0.8
3–2 $S(2)$	2.2870	18386	0.8 ± 0.2	2.5 ± 0.6
4–3 $S(3)$	2.3445	23955	0.5 ± 0.1	0.6 ± 0.1
2–1 $S(0)$	2.3556	12095	0.9 ± 0.1	22.5 ± 4.5
3–2 $S(1)$	2.3846	17818	0.8 ± 0.1	3.1 ± 0.8
1–0 $Q(1)$[d]	2.4066	6149	40.0 ± 5.0	...
1–0 $Q(2)$[d]	2.4134	6471	16.5 ± 2.0	...
1–0 $Q(3)$	2.4237	6956	45.1 ± 5.0	367.0 ± 41.0
1–0 $Q(4)$[d]	2.4375	7585	16.0 ± 2.0	...
1–0 $O(4)$[d]	3.0039	6471	13.4 ± 2.0	...
1–0 $O(5)$	3.2350	6956	25.7 ± 2.1	299.0 ± 25.0
2–1 $O(5)$[e]	3.4378	12550	2.1 ± 0.3	16.9 ± 2.4
0–0 $S(17)$	3.4857	25541	1.5 ± 0.5	0.21 ± 0.07
1–0 $O(6)$	3.5007	7584	4.3 ± 0.6	167.0 ± 22.0
0–0 $S(16)$	3.5475	23461	0.8 ± 0.4	0.38 ± 0.18
0–0 $S(15)$	3.6261	21413	3.5 ± 0.6	0.74 ± 0.13
2–1 $O(6)$[d]	3.7236	13150	2.2 ± 0.3	...
0–0 $S(14)$[d]	3.7244	19405		...
1–0 $O(7)$	3.8075	8365	9.2 ± 0.8	145.0 ± 12.0
0–0 $S(13)$	3.8461	17445	8.4 ± 0.9	3.11 ± 0.03
0–0 $S(12)$	3.9960	15542	3.8 ± 0.4	5.8 ± 0.6
2–1 $O(7)$[d]	4.0540	13891	1.4 ± 0.4	...

[a] Obtained from Dabrowski 1984.

[b] In a 5″ aperture at peak 1, except for the 3–2 $S(2)$ and 4–3 $S(3)$ lines, which were in a 12″ aperture. These latter two fluxes need to be divided by 2.9 to scale with those in a 5″ aperture [based on a measurement of the 1–0 $S(1)$ line through the 12″ aperture]. The 2 μm line fluxes combine two sets of measurements. The 3 μm line fluxes have been multiplied by a factor of 1.07 from the observed values based on a composite 2 and 3 μm spectrum.

[c] Assumes $A_\lambda \propto \lambda^{-1.5}$ with $A_K = 0.8$. All values apply to a 5″ aperture.

[d] Line detected, but either contaminated by telluric absorption lines, blended with other lines, or observed with incomplete spectral coverage. Intensities not reliable enough to be included in the analysis.

[e] Blended with 0–0 $S(18)$, but we estimate the contamination is $< 5\%$.

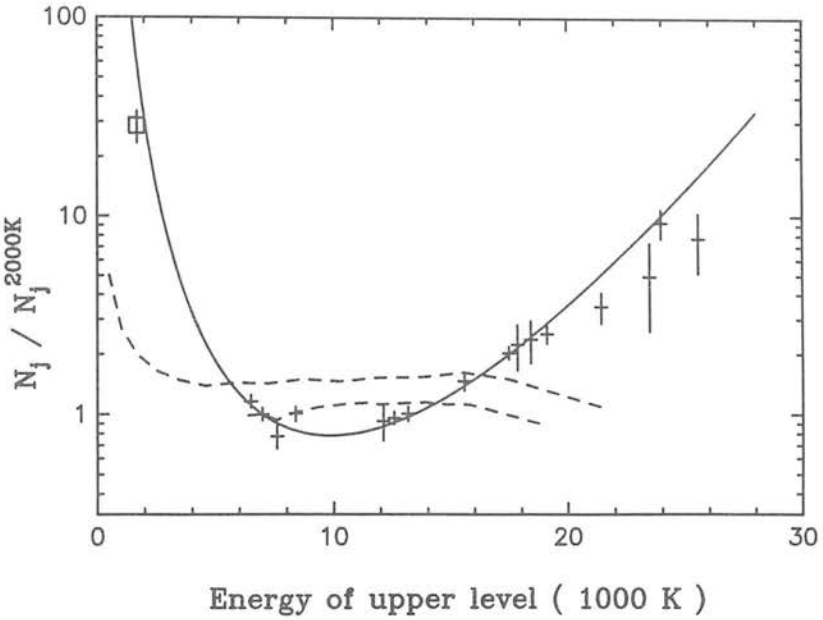


FIG. 2.—Plot of the ratios of observed, dereddened H_2 column densities to those from a Boltzmann distribution at 2000 K [normalized so that the 1–0 $S(1)$ ratio is unity] vs. upper energy level. Error bars are $\pm 1\ \sigma$. The continuous line is from the cooling flow calculations described in the text, drawn through the 1–0 $S(1)$ point. The dashed lines are the predictions of the C-shock models for OMC-1 by Draine and Roberge (1982) and by Chernoff, Hollenbach, and McKee (1982). The upper line is for pure rotational lines. The box represents the value of the 0–0 $S(2)$ line observed by Beck *et al.* (1979).

The weak, high-excitation lines are considerably more intense than would be expected from a straight line fit based on the strong, low-excitation lines. Since weak and strong lines come from both the 2 and the 3 μm spectra and correspond to both ortho- and para-H_2, neither errors in scaling, errors in dereddening, nor any particular ortho-para ratio can be the cause of the curvature in the locus of data points in Figure 2.

A significant contribution to the high-v lines from fluorescence is highly unlikely at peak 1, since the intensity of the 3–2 $S(2)$ line decreases with that of the 1–0 $S(1)$ line just off-source, but within the ionized nebula. It is clear on theoretical grounds that intensities of the weak, high-J pure rotational lines at peak 1 cannot be significantly enhanced by fluorescent emission.

IV. DISCUSSION

The solid curve in Figure 2 is the prediction from a "toy" calculation (which nevertheless contains all the major features of the full calculation) of column density in the cooling zone behind a hydrodynamic shock. The cooling rate is taken to be $\Lambda = \Lambda_0 T^s$ W (H_2 molecule)$^{-1}$, with s set to 4.7 to match roughly the calculations (Hollenbach and McKee 1979; Burton 1986) of cooling by thermalized H_2. If the shock is strong enough, the H_2 column density per state at level j with energy T_j degrees Kelvin is approximately

$$N_j/g_j \propto \int e^{-T_j/T}(Q\Lambda)^{-1}\,dT \propto T_j^{-s} - (T_j + T_v)^{-s}, \quad (1)$$

where the partition function Q is approximated by $AT(1 - e^{-T_v/T})$ with $T_v = 6000$ K.

A proper calculation of column density in a J-shock cooling zone, and a corresponding investigation of C-shocks, is being prepared for publication.

The cooling flow calculation described above provides a surprisingly good fit to the observed data. The apparent range of temperatures is naturally explained by the shape of the cooling function, independently of local conditions such as the shock velocity. This is consistent with other evidence (Brand *et al.* 1988) which shows that the ratio of a pair of H_2 lines with upper energy levels at 8365 K and 17,458 K respectively, is constant throughout the outflow.

In a calculation with the detailed shape of the cooling function properly treated, a better fit will result. Current C-shock models provide a poor fit to the wide range of H_2 data now available, and it appears that a superposition of several C-shocks may be required. The dashed curves are from the C-shock calculations by Draine and Roberge (1982) and Chernoff, Hollenbach, and McKee (1982).

We wish to thank the staff of the United Kingdom Infrared Telescope for friendly and able assistance during the several observing runs from which this *Letter* resulted. A. M. and M. B. are supported by SERC studentships. This work was done, in part, while M. G. B. held a National Research Council NASA Research Associateship at Ames Research Center.

REFERENCES

Beck, S. C., Lacy, J. H., and Geballe, T. R. 1979, *Ap. J. (Letters)*, **234**, L213.
Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978, *Ap. J.*, **223**, 464.
Brand, P. W. J. L., Toner, M. P., Webster, A. S., Geballe, T. R., and Williams, P. M. 1988, *M.N.R.A.S.*, submitted.
Burton, M. G. 1986, Ph.D. thesis, University of Edinburgh.
Chernoff, D. F., Hollenbach, D. J., and McKee, C. F. 1982, *Ap. J. (Letters)*, **259**, L97.
Dabrowski, I. 1984, *Canadian J. Phys.*, **62**, 1639.
Draine, B. T. 1980, *Ap. J.*, **241**, 1021.
Draine, B. T., and Roberge, W. G. 1982, *Ap. J. (Letters)*, **259**, L91.
Draine, B. T., Roberge, W. G., and Dalgarno, A. 1983, *Ap. J.*, **264**, 485.
Gautier, T. N., III, Fink, U., Treffers, R. R., and Larson, H. P. 1976, *Ap. J. (Letters)*, **207**, L29.
Geballe, T. R. 1986, in *Summer School on Interstellar Processes*, ed. D. J. Hollenbach and H. A. Thronson (NASA TM 88342), p. 129.

Hollenbach, D. J., and McKee, C. F. 1979, *Ap. J. Suppl.*, **41**, 553.
Hollenbach, D. J., and Shull, J. M. 1977, *Ap. J.*, **216**, 419.
Knacke, R. F., and Young, E. T. 1981, *Ap. J. (Letters)*, **249**, L65.
Kwan, J. 1977, *Ap. J.*, **216**, 713.
London, R., McCray, R., and Chu, S-I. 1977, *Ap. J.*, **217**, 442.
Nadeau, D., and Geballe, T. R. 1979, *Ap. J. (Letters)*, **230**, L169.
Oliva, E., and Moorwood, A. F. M. 1988, *Astr. Ap.*, **197**, 261.
Roberge, W. G., and Dalgarno, A. 1982, *Ap. J.*, **255**, 176.
Turner, J., Kirby-Docken, K., and Dalgarno, A. 1977, *Ap. J. Suppl.*, **35**, 281.

M. BIRD, P. W. J. L. BRAND, and A. MOORHOUSE: University of Edinburgh, Department of Astronomy, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

M. G. BURTON: NASA/Ames Research Center, MS 245-6, Moffett Field, CA 94035

T. R. GEBALLE: Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720

R. WADE: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

Mon. Not. R. astr. Soc. (1989) **236**, 929–934

The constancy of the ratio of the molecular hydrogen lines at 3.8 μm in Orion

P. W. J. L. Brand,[1] M. P. Toner,[1] T. R. Geballe,[2] A. S. Webster,[2] P. M. Williams[3] and M. G. Burton[4]

[1] *Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ*

[2] *Joint Astronomy Centre, 665 Komohana Street, Hilo HI 96720, USA*

[3] *Royal Observatory, Edinburgh EH9 3HJ*

[4] *NASA Ames Research Center, Space Sciences Division, MS: 245-6, Moffett Field CA 94035, USA*

Accepted 1988 September 1. Received 1988 August 15; in original form 1988 May 23

Summary. The 1–0 O(7) and 0–0 S(13) lines of H_2, at 3.807 and 3.846 μm, have been mapped over the region of the Orion molecular outflow. The intensity ratio of these lines is found to be independent of position in the outflow. From this it is inferred that the structure of the shocks and their cooling flows in Orion may be more akin to hydrodynamic shocks than the low-temperature C-shocks that are currently favoured.

1 Introduction

The shocked molecular hydrogen in Orion was first fully mapped by Beckwith *et al.* (1978), who employed the 2.122 μm $v=1$–0 S(1) line. We have mapped the same region in the $v=1$–0 O(7) and $v=0$–0 S(13) lines of H_2. These lines occur at 3.807 and 3.846 μm, respectively, and are sufficiently close together in wavelength to be little affected by even large variations in the extinction due to dust. However, their upper energy levels (17458 and 8365 K, respectively) differ by an amount that makes their ratio strongly temperature dependent at temperatures thought to be typical of shocked gas. Thus they can serve to test current models of shock cooling, which make different predictions of the temperature behaviour of the shocked gas.

2 Observations and data reduction

The observations were made in 1983 November at the 3.8-m United Kingdom Infrared Telescope, using the facility 7-channel cooled grating spectrometer. The aperture diameter was 5 arcsec and the chopper throw was 90 arcsec, large enough for the reference beam to be well away from the outflow. The resolving power was ~ 500. The dispersion of the spectrometer is such that both the O(7) and S(13) lines can be observed in a single grating setting, so that the intensity ratio is determined from simultaneous measurements of the two lines.

A total of 291 fully sampled spectra covering the 3.79–3.86 μm interval were obtained on a spatial grid of points separated by 3 arcsec and referenced to the position of the Becklin-Neugebauer (BN) object. Integration times ranged from 2 to 10 min, depending on the strength of the line emission. The star BS 1552 was used as an intensity calibrator. The typical statistical uncertainty in the line flux at each position is 0.9×10^{-16} W m^{-2} in the 5-arcsec beam. Examination of the spectra reveal in some locations a weak Humphreys $\kappa(16-6)$ line of atomic hydrogen, from the foreground nebula, at 3.819 μm, in the long wavelength wing of the O(7) line. This small contribution to the apparent O(7) intensity was subtracted. The data reduction technique is described in greater detail in Toner (1986). The summed spectrum is shown in Fig. 1.

A map of the summed O(7)+S(13) line intensities is shown in Fig. 2(a). The summed intensity is displayed because, as is shown below, the ratio of the two lines is approximately unity and is independent of position in the flow. The map has been smoothed by a Gaussian profile of 3 arcsec FWHM and then contoured at intervals of 1.5×10^{-16} W m^{-2} beam^{-1}

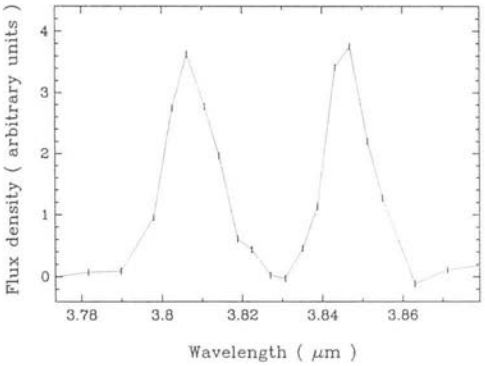


Figure 1. The spectrum produced by coadding all the spectra obtained in OMC-1 weighted by their squared signal-to-noise ratios. The lines are produced by the H_2 transitions $v=1-0$ O(7) on the left, and $v=0-0$ S(13) on the right. The symbols represent the size of the formal error bars after coadding.

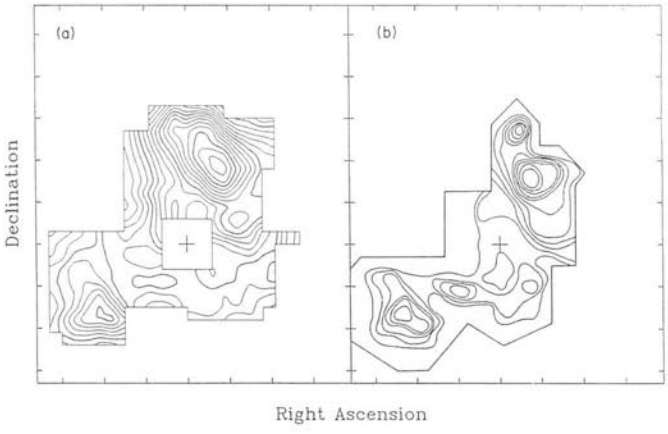


Figure 2. (a) Map of the sum of O(7) and S(13) intensities. The tick marks are at intervals of 10 arcsec. The cross is at the position (see Section 2) of the Becklin-Neugebauer object, around which data are not plotted, because of high continuum. The contours are at intervals of 1.5×10^{-16} W m^{-2} beam^{-1}, and the beam diameter was 5 arcsec. Data were taken at intervals of 3 arcsec in RA and declination. (b) A map of the 1-0 S(1) line intensity from Beckwith *et al.* (1978), at the same scale.

(2.8×10^{-7} W m^{-2} sr^{-1}), which corresponds approximately to a 1.5 σ detection per beam. The BN object is marked with a cross and is the origin of the map. This position (Beckwith *et al.* 1978) is at RA = 5^h 32^m 46?7, Dec = $-5°$ 24′ 17″ (1950). Near the BN object the signal-to-noise ratios on the lines were poor due to large spatial gradients in the thermal continuum; therefore these points are not plotted.

Fig. 2(b) shows several contours drawn to the same scale from the 5-arcsec resolution map by Beckwith *et al.* (1978). It is clear that the morphologies of the two maps are very similar, implying that the 2.1 and 3.8 μm lines come from the same regions and that the extinction does not vary strongly across the map. Since line intensities vary rapidly with position and the two sets of data were obtained in very different fashions, it was not regarded as sensible to draw further conclusions from comparisons of these data. In addition to the 3.8 μm measurements, a limited set of measurements of the 1–0 S(1) line was made at UKIRT, using the same instrument and observing technique. The region measured lies just to the west of Peak 1 in the Beckwith *et al.* map. The mean intensity ratio 1–0 O(7)/1–0 S(1) in this region is 0.23 ± 0.05.

3 Analysis

In the brightest regions of the map an accurate measurement of the S(13)/O(7) ratio can be obtained at each individual position. However, in the low-intensity regions, which are of equal interest, the signal-to-noise ratios of the lines are poor and ratios derived from individual positions are meaningless. In order to determine the line ratio in these regions, a scheme for coadding weak spectra from different locations was devised. First, the mapped region was divided into five subregions: Peak 1 and northwards; Peak 1 and southwards; the BN region; BN to Peak 2; and south of Peak 2. (the positions of Peaks 1 and 2 are given in Beckwith *et al.* 1978). In each of these regions the brightest pixel was selected. If the signal-to-noise ratio in each line was greater than 10, that line pair was stored and the next pair examined. Otherwise, pixels ranked in order of decreasing intensity were coadded until the signal-to-noise ratio of each line in the combined spectrum was greater than 10. Thus the line ratio as a function of mean intensity per pixel (of the summed line intensities) was determined. No significant differences between the results for the various regions could be detected, so the entire map was re-analysed as a single region.

The result of this analysis is displayed in Fig. 3. Two conclusions may be drawn from this figure. First, there is no significant difference in the behaviour of the line ratio from region to region. Secondly, the ratio does not vary with pixel brightness. Within 1 σ the ratio is 1.0.

The excitation temperature of this line pair, calculated assuming Boltzmann equilibrium of the two upper levels and using standard formulae (with molecular constants from Turner, Kirby-Docken & Dalgarno 1977, and Dabrowski 1984), is 2400 K over the entire outflow region. A 10 per cent change in the ratio, which is equivalent to the 1 σ uncertainty in a line flux, corresponds to a change in temperature of 60 K. The excitation temperature for the O(7)–S(13) pair is somewhat higher than that derived by Beckwith *et al.* (1983) and most other observers from the 2 μm lines.

4 Discussion and conclusions

The temperature behaviour of the shocked molecular gas responsible for the H_2 line emission is determined by the initial density of the shocked gas, the shock velocity, the magnetic field strength, etc. These may vary within the outflow region by large factors and therefore cause observable differences in different parts of the outflow region, e.g., in relative line strengths. The ratio of lines selected specifically to discern this behaviour, instead shows remarkable

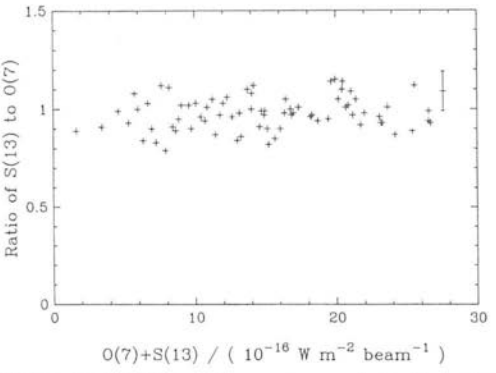


Figure 3. The ratio of 0–0 S(13) to 1–0 O(7) intensity versus mean intensity, as described in the text. The error bar per point is shown.

constancy. A possible explanation is briefly outlined here, and will be discussed in detail in Brand (1988).

First of all, it should be emphasized that in this context, a problem with C-shock (Draine 1980; Draine & Roberge 1982) models is that heating in such a shock is by drag between ions and neutrals and therefore depends on velocity, while cooling does not. Therefore the peak neutral temperature, which determines the observed H_2 infrared (IR) line ratios, depends on shock velocity, which might be expected to vary significantly throughout the Orion flow. The 0–0 S(13)/1–0 O(7) line ratio in particular, with its Boltzmann factor of $\exp(-9093/T)$, should as a consequence vary. This behaviour will be fully analysed in Brand (1988), but is taken here as a reason for re-examining the classical J-shock, where heating is by collision in a hydrodynamic jump, in a time short compared to the cooling time for the gas, and the observed radiation comes from a cooling zone behind the shock. It will be shown that this structure gives rise to velocity-independent line ratios.

The density is supposed high enough such that the rotational and vibrational states are thermalized. Then the line cooling rate of H_2 is readily calculated (Hollenbach & McKee 1979; Burton 1986)

$$\Lambda(T) = \Sigma g_j e^{-T_j/T} A_j h\nu_j / Q(T), \tag{1}$$

where $Q(T)$ is the partition function and the other symbols have their usual meanings. A good fit to this function is (Burton 1986):

$$\begin{aligned} \Lambda(T) &\propto T^{4.7}, && 200 < (T/\mathrm{K}) < 2000 \\ &\propto \exp(-7700/T), && 2000 < (T/\mathrm{K}) < 20000. \end{aligned} \tag{2}$$

Dissociational cooling proportional to $\exp(-52800/T)$ (Roberge & Dalgarno 1982) is important at the higher temperatures, above 3500 K. It is straight forward to show that the column density per state of H_2 in energy level j through the cooling zone is

$$N_j/g_j = k n_0 V_s \int_0^{T_{\max}} \phi(T) e^{-T_j/T} (Q(T)\Lambda(T))^{-1} dT, \tag{3}$$

where n_0 is the number density of H_2 before the shock, V_s is the shock speed, $\phi(T)$ is a slowly varying factor, of order 3, to account for dissociation and the varying ratio of enthalpy to temperature in the molecular gas, and T_j is the level energy in Kelvin. The integrand is dominated by the two factors $(Q\Lambda)^{-1}$ and $\exp(-T_j/T)$, which together create a peak analogous to the Gamow peak in thermonuclear reaction rate calculations.

For simplicity, assume initially that $\Lambda \propto T^s$, $Q(T) \propto T$, that $\phi(T)$ is constant, and that T_{max}, the maximum temperature behind the shock, is high enough that it can be set to infinity in the integral. Then the peak of the integrand is at $T = T_j/(s+1)$. The major contribution to the integral comes from the vicinity of this temperature. Thus, provided that the exponent $s = d\log\Lambda/d\log T$ varies slowly in this region (and does not decrease too fast at high T), the analysis can be applied to a more general function, represented locally as a power law of exponent s. The cooling function of H_2 in thermal equilibrium has these properties. Referring still to the simple case of $\Lambda \propto T^s$, $Q \propto T$, the result of the integration is

$$N_j/g_j \propto T_j^{-s}. \tag{4}$$

That is, the inferred column density ratio depends not on any temperature in the gas, but on level energy, and on the behaviour, through s, of the cooling function.

The temperature derived from the S(13) and O(7) lines is

$$\begin{aligned} T(\text{derived}) &= (T_{S(13)} - T_{O(7)})/s\log(T_{S(13)}/T_{O(7)}) \\ &= (12398/s)\text{K}. \end{aligned} \tag{5}$$

Thus a derived temperature of 2400 K, which the observations imply, gives $s \simeq 5$. The emission peak for O(7) is at a temperature of 1700 K, which with equation (2) implies $s = 4.7$, while that for S(13) is at 3500 K, and the contribution by dissociation to the cooling rate will cause s to be higher. It is therefore to be expected that the value of s obtained from equation (5) should be somewhat greater than 4.7. Thus the value deduced above is entirely consistent with H_2 cooling. A proper treatment of the cooling zone is given in Brand (1988), but differs only in detail from the simple result above.

The crucial point is that this derived temperature is independent of local conditions, and in particular is independent of velocity.

If shocks in which heating is fast and cooling is slow predominate in the Orion molecular outflow, then there is a clear reason for the observed constancy of the measured line ratio. Brand *et al.* (1988) present evidence based on many line ratios from the brightest part of the Orion outflow, which also points toward this interpretation of the emission region.

We conclude, therefore, that the intensity ratio of the H_2 lines $v = 0$–0 S(13) and $v = 1$–0 O(7), a number which is highly sensitive to post-shock temperatures, is constant to within ± 10 per cent throughout the shocked gas of the Orion molecular outflow. We suggest that throughout the flow, shock heating of the gas occurs promptly in a region preceding and far narrower than the cooling zone (whose structure is independent of the exact means of heating), and that this may be the reason for the constancy of the line ratio. This suggested structure is like that of a high-density jump shock.

Acknowledgments

We thank the staff at the UK Infrared Telescope for their able and friendly assistance. MPT was supported by a SERC studentship during this work.

References

Beckwith, S., Evans, N. J., Gatley, I., Gull, G. & Russell, R. W., 1983. *Astrophys. J.*, **264,** 152.
Beckwith, S., Persson, S. E., Neugebauer, G. & Becklin, E. E., 1978. *Astrophys. J.*, **223,** 464.
Brand, P. W J. L., Burton, M. G., Geballe, T. R., Moorhouse, A., Bird, M. & Wade, R. D., 1988. *Astrophys. J.*, **334,** L103.
Burton, M. G., 1986. *PhD thesis,* University of Edinburgh.
Dabrowski, I., 1984. *Can. J. Phys.*, **62,** 1639.

Draine, B. T., 1980. *Astrophys. J.*, **241**, 1021.
Draine, B. T. & Roberge, W. G., 1982. *Astrophys. J.*, **259**, L91.
Hollenbach, D. J. & McKee, C. F., 1979. *Astrophys. J. Suppl.*, **41**, 647.
Roberge, W. G. & Dalgarno, A., 1982. *Astrophys. J.*, **255**, L176.
Toner, M. P., 1986. *MPhil thesis*, University of Edinburgh.
Turner, J., Kirby-Docken, K. & Dalgarno, A., 1977. *Astrophys. J. Suppl.*, **35**, 281.

Mon. Not. R. astr. Soc. (1989) **236**, 409–423

Molecular hydrogen line ratios in four regions of shock-excited gas

M. G. Burton,[1,2] P. W. J. L. Brand,[1] T. R. Geballe[3,4] and A. S. Webster[3]

[1] *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[2] *NASA Ames Research Center, Space Science Division, MS:245-6, Moffett Field, CA 94035, USA*
[3] *Joint Astronomy Centre, 665 Komohana St, Hilo, HI 96720, USA*
[4] *Foundation for Astronomical Research in The Netherlands - ASTRON*

Accepted 1988 August 9. Received 1988 July 15

Summary. Five emission lines of molecular hydrogen, with wavelengths in the ranges of 2.10–2.25 and 3.80–3.85 μm, have been observed in four objects of different type in which the line emission is believed to be excited by shocks. The relative intensities of the lines $1-0\,S(1):1-0\,S(0):2-1\,S(1)$ are approximately 10.5:2.5:1.0 in all four objects. The $0-0\,S(13):1-0\,O(7)$ line ratio, however, varies from 1.05 in OMC-1 to about 2.3 in the Herbig–Haro object HH 7. The excitation temperature derived from the $S(13)$ and $O(7)$ lines is higher than that derived from the $1-0$ and $2-1\,S(1)$ lines in all four objects, so the shocked gas in these objects cannot be characterized by a single temperature. The constancy of the $1-0/2-1\,S(1)$ line ratio between sources suggests that the post-shock gas is 'thermalized' in each source. This requires a post-shock gas density (in purely molecular gas) of $\geq 10^7$ cm^{-3}. The $S(13)/O(7)$ ratio is particularly sensitive to the density and temperature conditions in the gas. We discuss the conditions necessary to excite these lines. It is puzzling that this ratio varies from source to source, while the $1-0/2-1\,S(1)$ ratio remains constant.

1 Introduction

Line emission from molecular hydrogen is known to be commonplace in a wide variety of astronomical objects. The fundamental vibration–rotation and pure rotational transitions of the molecule, which occur in the near-infrared, have been observed in regions of massive star formation such as OMC-1 (e.g. Beckwith *et al.* 1978; Scoville *et al.* 1982; Nadeau, Geballe & Neugebauer 1982; Beckwith *et al.* 1983) and NGC 2071 (e.g. Lane & Bally 1986), in supernova remnants (e.g. Treffers 1979; Burton *et al.* 1988), in proto-planetary nebulae (e.g. Beckwith, Beck & Gatley 1984) and planetary nebulae (e.g. Treffers *et al.* 1976), in

Herbig–Haro objects (e.g. Zealey, Williams & Sandell 1984), in reflection nebulae (e.g. Gatley *et al.* 1987) and in extra-galactic sources (e.g. Fischer *et al.* 1987).

There are two mechanisms that have been invoked to account for the H_2 line emission: fluorescence induced by ultraviolet pumping, and radiative cooling following collisional excitation by shock waves. In shock-excitation of a molecular cloud, collisions populate the rotation–vibration levels of the ground electronic state. Lines of shocked H_2 characteristically have widths > 20 km s^{-1} (widths of fluorescent H_2 lines are expected to be considerably narrower) and have rotational and vibrational excitation temperatures of ~ 2000 K for the 2 μm lines. Although it has become a routine matter to estimate the temperature of shocked gas from ratios of line intensities, in some models the H_2 line emission occurs over a wide range of temperatures (e.g. Brand *et al.* 1988b), so it is clear that it will not be possible to distinguish between them unless careful measurements are made of more than the few standard 2 μm lines.

We have measured the intensities of five H_2 lines, covering a wide range of energy levels, in four galactic sources of different type: a dense molecular cloud impacted by a massive protostellar molecular outflow (OMC-1), a Herbig–Haro object (HH 7), a proto-planetary nebula (CRL 618) and a supernova remnant (IC 443). In each of these sources the excitation mechanism is thought to be violent collisions induced by molecular shock waves. The preshock conditions and shock speeds can reasonably be expected to vary from source to source, so by observing the line ratios one might be able to discern differences in the physical conditions in the shocked molecular gas of these objects.

2 Observations

The observations were made during 1984 December, 1986 January and 1987 January at the 3.8-m UK Infrared Telescope (UKIRT) on Mauna Kea, Hawaii. The five H_2 emission lines [$1-0\,S(1)$, $1-0\,S(0)$ and $2-1\,S(1)$ lines near 2.1–2.25 μm and the $1-0\,O(7)$ and $0-0\,S(13)$ near 3.8 μm] were measured in each source. The log of the observations is shown in Table 1 and the molecular parameters of the five H_2 transitions are listed in Table 2.

The seven-channel cooled grating spectrometer (CGS) was used for most of the measurements. The 300 lines mm^{-1} grating provided a resolution of 260 at 2.2 μm and 460 at 3.8 μm. The aperture diameter was 5 arcsec. The 2 μm lines of IC 443, however, were observed through a circular variable filter (CVF) of resolving power ~ 120. Aperture diameters of 12 and 19 arcsec were employed. During all of the measurements standard spatial chopping and nodding of the telescope were performed and telescope drift was eliminated by offset guiding on visible stars. Standard stars, observed at air masses similar to those of the sources, were used for the flux calibration and are noted in Table 1.

The $S(13)$ and $O(7)$ lines are particularly favourable for observation with the grating spectrometer. They are sufficiently close together in wavelength to be observable simultaneously with the seven-element array. Thus, their intensity ratio is insensitive to pointing errors, to inaccuracies and drifts in calibration, to variations in the transparency of the sky and to uncertainties in the extinction.

3 Results

3.1 LINE RATIOS

The spectra of the four sources are shown in Figs 1 and 2, and the line fluxes listed in Table 3. Line fluxes were determined by fitting the instrumental profile, taken to be a Gaussian whose

Table 1. Observing log.

Object	RA 1950.0	DEC 1950.0	Instrument & Aperture	Wavelength (μm)	Date Observed	Standard & Magnitude	Integration Time (min)
HH 7	3 26 3.0	31 5 10	CGS 5"	2.2	13/12/84	HR1552 (K=4.140)	100
			CGS 5"	3.8	16/12/84	HR622 (L'=2.60)	260
			CGS 5"	3.8	08/01/86	HR1552 (L'=4.170)	145
			CGS 5"	3.8	07/01/87	HR622 (L'=2.60)	100
CRL 618	4 39 33.8	36 1 15	CGS 5"	2.2	13/12/84	HR1552 (K=4.140)	40
			CGS 5"	3.8	16/12/84	HR1203 (L'=2.63)	25
OMC-1 Peak 1	5 32 46.3	-5 24 2	CGS 5"	2.2	14/12/84	HR2241 (K=4.004)	20
			CGS 5"	3.8	08/01/86	HR1552 (L'=4.170)	10
IC 443	6 14 43 -20"	22 23 00† -20"	CVF 12"	2.2	13/12/84	HR2484 (K=2.30)	15
			CGS 5"	3.8	16/12/84	HR1791 (L'=2.00)	100
			CGS 5"	3.8	06/01/87	HR2821 (L'=1.35)	65
	-844"	+635"	CVF 19"	2.2	07/01/86	HR3188 (K=2.32)	25
			CGS 5"	3.8	07/01/87	HR2821 (L'=1.35)	50

†The coordinates listed for IC 443 refer to a nominal (0,0) position.

Table 2. Lines observed.

Line	Wavelength (μm)	Upper Energy Level (K)	Decay Rate $\times 10^{-7}\,s^{-1}$
1-0 S(1)	2.1218	6952	3.47
1-0 S(0)	2.2233	6474	2.53
2-1 S(1)	2.2477	12556	4.98
1-0 O(7)	3.807	8366	1.06
0-0 S(13)	3.846	17450	16.2

width was determined from spectra of high signal-to-noise ratio in OMC-1, to the measurements. The intensities for the 3.8 μm lines in HH 7 combine three separate sets of observations. Two sets were combined for IC 443 (-20 arcsec, -20 arcsec) (see Table 1 for notation). In all cases the 1 σ errors were obtained from the standard deviation of the mean of the data points and do not include any systematic effects such as inaccuracies in the flux calibration.

The line ratios are given in Table 4, where it will be seen that the four sources show only slight variations in the ratios of their 2.1–2.25 μm lines, but considerably larger variation in the ratios of their 3.8 μm lines. At OMC-1 Pk1 the 2 μm lines are in the ratio $1-0\,S(1):1-0\,S(0):2-1\,S(1)=10.5:2.5:1.0$ and the 3.8 μm lines are in the ratio $0-0\,S(13)/1-0\,O(7)=1.05$. These values are found throughout the OMC-1 shocked region (Toner 1986; Brand *et al.* 1988a). Among all four sources observed here, the range is

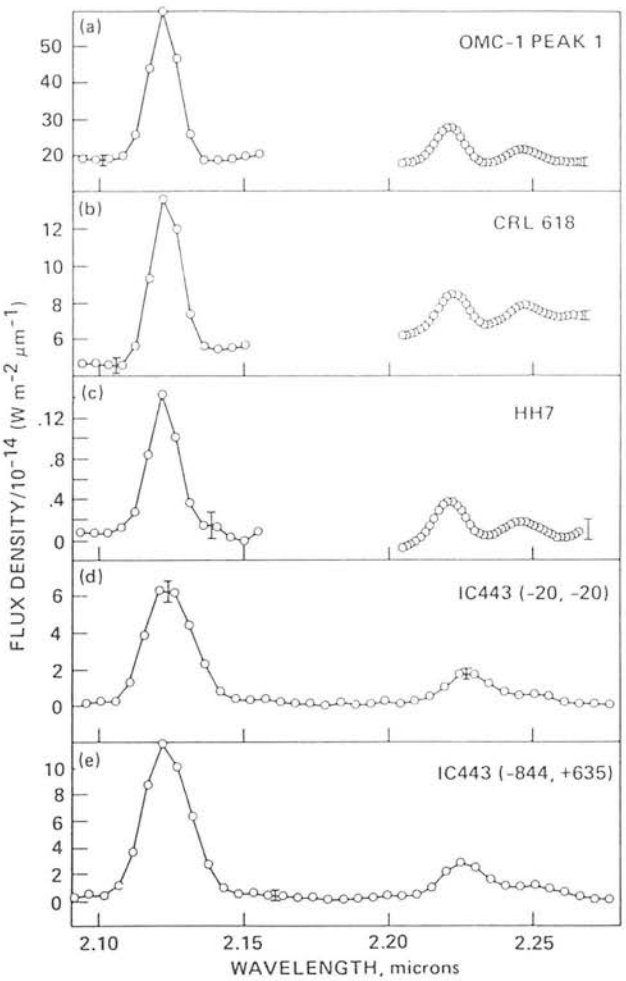


Figure 1. Spectra of the H_2 emission lines $1-0\,S(1)$ (2.122 μm), $1-0\,S(0)$ (2.223 μm) and $2-1\,S(1)$ (2.248 μm) in four different shock-excited sources; a region of massive star formation (OMC-1), a proto-planetary nebula (CRL 618), a Herbig–Haro object (HH 7) and a supernova remnant (IC 443; two positions). The spectra have been smoothed with a Gaussian filter of FWHM equal to the instrumental resolution. The data were taken using the UKIRT and either a grating spectrometer (CGS) or circular variable filter (CVF). Representative $\pm 1\sigma$ error bars are shown. Table 1 lists details of the observations and Table 3 the line fluxes.

10.1–13.4 in the $1-0\,S(1)/2-1\,S(1)$ ratio, 2.2–3.1 for $1-0\,S(0)/2-1\,S(1)$, and 1.0–2.3 for $S(13)/O(7)$. Two of the $1-0\,S(1)/2-1\,S(1)$ ratios agree with the OMC-1 result to better than 10 per cent; the other two (in IC 443) are 20 and 30 per cent higher, but were observed with an instrument which did not resolve the $2-1\,S(1)$ line from the stronger $1-0\,S(0)$. As a result we have some doubts regarding the reality of our reported $1-0\,S(1)/2-1\,S(1)$ ratio in IC 443.

3.2 EXCITATION TEMPERATURES

An excitation temperature, T_{ex}, can be defined for a line ratio by:

$$I_1/I_2 = \frac{A_1 \nu_1 g_1}{A_2 \nu_2 g_2} \exp\left(\frac{E_2 - E_1}{kT}\right) \exp(\Delta\tau). \qquad (1)$$

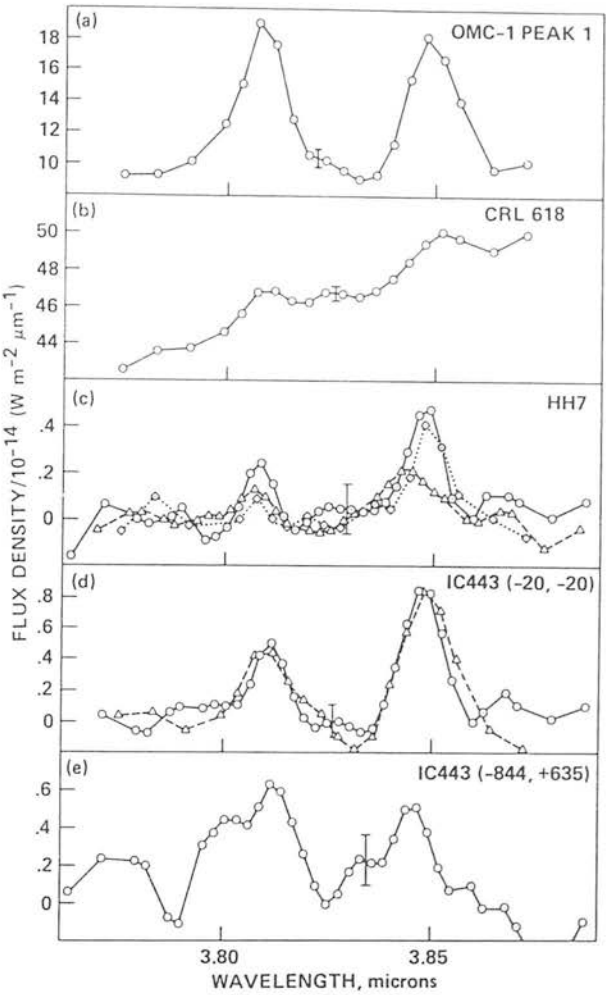


Figure 2. As for Fig. 1, spectra of the $1-0\,O(7)$ (3.807 μm) and $0-0\,S(13)$ (3.846 μm) H_2 emission lines. Three spectra for HH 7 are shown [one taken in 1984 December (◇), one in 1986 January (△) and one in 1987 January (○)]. For IC 443, two spectra are shown [one from 1984 December (△) and one from 1987 January (○)]. For HH 7 and IC 443 the representative $\pm 1\sigma$ error bars are for the 1987 January data in each case.

Here I is the observed line intensity, A the radiative decay rate for the transition (taken from Turner, Kirby-Docken & Dalgarno 1977), g the statistical weight of the level, ν the line frequency, E the energy level of the upper state of the transition (taken from Dabrowski & Herzberg 1984) and $\Delta\tau$ the differential optical depth of the interstellar extinction between the wavelengths of the two lines. The 1 and 2 refer to the two lines in question.

Table 5 shows the excitation temperatures derived for the 2- and 3.8-μm lines. For the calculations, it was first assumed that there was no differential extinction between 2.12 and 2.25 μm towards the emitting H_2 in all sources, and then the calculation was repeated with 0.1 mag of extinction. The excitation temperatures derived from the 2-μm lines are all found to be close to 2000 K whichever value is chosen for the extinction; the variation is less than 200 K between sources. It was assumed that there was no differential extinction between the 3.8-μm lines. The $S(13)/O(7)$ excitation temperature was found to show more variation, from 2400 to 3100 K. In every source, the $S(13)/O(7)$ excitation temperature is higher than that derived from the 2-μm lines.

Table 3. Line fluxes. The fluxes are in W m^{-2}; for each object, the first row is the line flux and the second row the 1σ error. The figure in parenthesis is the exponent [i.e. (−16) denotes 10^{-16}]. Measurements for the 2 μm lines in IC 443 were made through 12 arcsec [at (−20 arcsec, −20 arcsec)] and 19 arcsec [at (−844 arcsec, +635 arcsec)] apertures. All other measurements were through 5 arcsec apertures. For the 3.8 μm lines in HH 7 and IC 443 (−20 arcsec, −20 arcsec), all the individual spectra shown in Fig. 2 have been coadded to obtain the line flux.

Object	1-0 S(1)		1-0 S(0)		2-1 S(1)		1-0 O(7)		0-0 S(13)	
HH 7	1.80	(-16)	4.38	(-17)	1.84	(-17)	1.94	(-17)	4.26	(-17)
	1.0	(-17)	8.1	(-18)	7.4	(-18)	4.3	(-18)	4.3	(-18)
CRL 618	1.13	(-15)	2.45	(-16)	1.16	(-16)	2.22	(-16)	2.17	(-16)
	5	(-17)	2.4	(-17)	2.3	(-17)	4.4	(-17)	4.4	(-17)
OMC-1	5.25	(-15)	1.26	(-15)	5.08	(-16)	1.03	(-15)	1.08	(-15)
	1.3	(-16)	7	(-17)	5.7	(-17)	3	(-17)	3	(-17)
IC 443	1.26	(-15)	2.90	(-16)	9.6	(-17)	5.38	(-17)	9.80	(-17)
(-20",-20")	8	(-17)	1.5	(-17)	1.2	(-17)	7.6	(-18)	7.6	(-18)
IC 443	2.28	(-15)	5.14	(-16)	1.85	(-16)	5.98	(-17)	4.81	(-17)
(-844",+635")	2.3	(-17)	2.3	(-17)	2.3	(-17)	1.3	(-17)	1.3	(-17)

When the gas is thermalized at a single temperature the excitation temperature is equal to the kinetic temperature and is independent of the lines used to calculate it. Knacke & Young (1981) found for OMC-1 that the relative line intensities (for lines with upper-J levels up to $J=17$, with $v=0$) are consistent with LTE at 2000 ± 400 K. Our present results for OMC-1 lie within their uncertainties, but demonstrate that a single excitation temperature does not characterize the shocked H_2. A similar conclusion is drawn from our results for the other three objects, so we find it to be incorrect to state that post-shock H_2 is in LTE at any single temperature. This conclusion is supported by Beck & Beckwith's (1983) observation of the $3-2\,S(1)$ line in OMC-1, from which an excitation temperature of 2700 ± 300 K for the $v=2$ and $v=3$ levels was derived, compared to 2000 ± 200 K for the $v=1$ and $v=2$ levels. It is also supported by the derivation by Lane & Bally (1986) of different vibrational and rotational temperatures in the shocked gas in NGC 2071.

4 Discussion

4.1 USING THE LINE RATIOS TO MEASURE TEMPERATURE AND DENSITY IN THE SHOCKED GAS?

The ratio of the intensities of the $0-0\,S(13)$ and $1-0\,O(7)$ lines is apparently particularly sensitive to the excitation conditions in the shocked molecular hydrogen. The emitted photons for these lines come from very different upper state energy levels; the $S(13)$ line comes from a pure rotational level 17 450 K above the ground state and the $O(7)$ line from a vibration–rotation level 8300 K above ground. It is difficult to populate the $v=0$, $J=15$ level of the molecule for two reasons, both of which stem from the high value of J:

Table 4. Line ratios. For each source, three rows of data analysis are presented. In the first row is the expected value of line ratio. The best estimate of this parameter is given by the statistic $\mu_X/\mu_Y \times [1+(\sigma_Y/\mu_Y)^2]$, where μ and σ are the mean and the variance of the line pair, X and Y. The second row is the standard deviation, estimated from the statistic $(\mu_X/\mu_Y) \times \sqrt{[(\sigma_X/\mu_X)^2+(\sigma_Y/\mu_Y)^2]}$. The third row is the range of values for the ratio derived from the 1σ errors for each line in the pair.

Object	1-0 S(1)/2-1 S(1)	1-0 S(1)/1-0 S(0)	1-0 S(0)/2-1 S(1)	0-0 S(13)/1-0 O(7)
HH 7	11.3	4.2	2.8	2.3
	4.0	0.8	1.1	0.5
	[6.6 – 17.2]	[3.3 – 5.3]	[1.4 – 4.7]	[1.6 – 3.1]
CRL 618	10.1	4.7	2.2	1.0
	2.0	0.5	0.5	0.3
	[7.8 – 12.7]	[4.0 – 5.3]	[1.6 – 2.9]	[0.65 – 1.5]
OMC-1	10.5	4.2	2.5	1.05
	1.2	0.3	0.3	0.04
	[9.1 – 11.9]	[3.9 – 4.5]	[2.1 – 3.0]	[1.00 – 1.11]
IC 443 (–20",–20")	13.4	4.4	3.1	1.9
	1.6	0.2	0.4	0.3
	[11.5 – 15.2]	[4.1 – 4.6]	[2.6 – 3.6]	[1.5 – 2.3]
IC 443 (–844",+635")	12.5	4.4	2.8	0.84
	1.6	0.2	0.4	0.27
	[10.8 – 14.2]	[4.2 – 4.7]	[2.4 – 3.3]	[0.49 – 1.3]

Table 5. Excitation temperatures.

Object	1-0 S(1) / 2-1 S(1)		1-0 S(0) / 2-1 S(1)	0-0 S(13) / 1-0 O(7)
	$\Delta\tau = 0.0$	$\Delta\tau = 0.1$		
HH 7	2000	1900	1900	3100
	[1700–2400]	[1700–2300]	[1700–2500]	[2800–3400]
CRL 618	2100	2000	2100	2400
	[1900–2300]	[1800–2200]	[1900–2400]	[2200–2700]
OMC-1	2020	1950	2000	2440
	[1930–2140]	[1870–2060]	[1900–2130]	[2410–2480]
IC 443	1900	1900	1900	2900
(–20",–20")	[1800–2000]	[1800–2000]	[1800–2000]	[2700–3100]
IC 443	1900	1800	1900	2300
(–844",+635")	[1800–2000]	[1800–1900]	[1800–2000]	[2000–2600]

(i) Collisional excitation of the molecule occurs mainly via quadrupole transitions, which tend to increase J in units of 2 (Green, Ramaswamy & Rabitz 1978). Excitation to the $J=15$ level from the ground state therefore requires many more collisions than does excitation to the $v=1, J=5$ level.

(ii) The *Einstein-A* values for radiative decay from high-J levels are considerably higher than from low-J levels (Turner *et al.* 1977). As an H_2 molecule is collisionally excited towards the $J = 15$ level, it is more and more likely to decay to a lower J level.

On the other hand, to collisionally excite to the (1,5) level requires a single large change in energy, represented by a small value of the Boltzmann factor in the rate coefficient. At 2000 K, the rate coefficient to excite from $v = 0$ to $v = 1$ is $\sim 1.2 \times 10^{-14}$ cm^3 s^{-1}, whereas the excitation rate from $J = 13$ to $J = 15$ in the ground vibrational state is $\sim 6 \times 10^{-13}$ cm^2 s^{-1} (using rate coefficients for H_2–H_2 collisions from Draine, Roberge & Dalgarno 1983). The radiative decay rate out of the (1,5) level is 7.8×10^{-7} s^{-1}, and out of the (0,15) level is 1.6×10^{-6} s^{-1}; so the critical densities required to populate these levels are of the order of a few $\times 10^6$ cm^{-3}. In the rest of this section we discuss the relative population of these levels as a function of the gas density and temperature.

For each energy level (v,J) there exists a critical density, n_{crit}, below which the level is not thermally populated. For densities above n_{crit} the line intensity will depend on the local kinetic temperature of the gas, whereas below n_{crit} the intensity will also be sensitive to the density. We

Table 6. Critical densities for molecular hydrogen. H_2 number densities at which the populations of the $(v,J) = (0,15)$, (1,3) and (2,3) levels are a given percentage of the LTE population. These critical densities are calculated from the statistical equilibrium population levels, using the collision cross-sections for H_2 with H_2 from Draine *et al.* (1983 – DRD).

Temperature (K)	Percentage of LTE Population	(0,15) (cm^{-3})		(1,3) (cm^{-3})		(2,3) (cm^{-3})	
1000	50%	2	(6)	1.4	(7)	3	(7)
	75%	5	(6)	4	(7)	7	(7)
	90%	1.3	(7)	1	(8)	2	(8)
	95%	3	(7)	2	(8)	4	(8)
1500	50%	1.4	(6)	3	(6)	6	(6)
	75%	4	(6)	9	(6)	1.7	(7)
	90%	1	(7)	3	(7)	5	(7)
	95%	2	(7)	5	(7)	1	(8)
2000	50%	1	(6)	1	(6)	2	(6)
	75%	3	(6)	3	(6)	6	(6)
	90%	1	(7)	9	(6)	2	(7)
	95%	2	(7)	2	(7)	4	(7)
2500	50%	6	(5)	4	(5)	8	(5)
	75%	1.6	(6)	1.3	(6)	2	(6)
	90%	4	(6)	4	(6)	7	(6)
	95%	7	(6)	7	(6)	1.3	(7)
3000	50%	3	(5)	2	(5)	4	(5)
	75%	8	(5)	7	(5)	1.3	(6)
	90%	2	(6)	2	(6)	4	(6)
	95%	4	(6)	4	(6)	6	(6)
3500	50%	1.3	(5)	1	(5)	2	(5)
	75%	3	(5)	3	(5)	4	(5)
	90%	9	(5)	9	(5)	2	(6)
	95%	2	(6)	2	(6)	4	(6)
4000	50%	6	(4)	6	(4)	1.3	(5)
	75%	1.6	(5)	2	(5)	4	(5)
	90%	4	(5)	5	(5)	1	(6)
	95%	9	(5)	1	(6)	2	(6)

have investigated the H_2 level populations in statistical equilibrium in order to determine n_{crit} for the energy levels of interest. In Table 6 the critical densities required to achieve a specified percentage of the thermal population for the $(v,J)=(0,15)$, $(1,3)$ and $(2,3)$ levels are shown; the results for the $(1,5)$ and $(1,2)$ levels are much the same as those for the $(1,3)$ level. Collisional rates for H_2 molecules with H_2 molecules were taken from Draine *et al.* (1983). It can be seen that n_{crit} for the $(0,15)$ and $(1,3)$ levels is about the same for $T \geq 2000$ K, taking values of 10^6–10^7 cm^{-3}. The critical density for the $(2,3)$ level is a factor of 2 higher than that for the $(1,3)$ level.

There are large uncertainties in the values of the collisional rate coefficients for molecular hydrogen. The critical densities have therefore also been recalculated using different rate coefficients taken from Shull & Hollenbach (1977) and Shull & Beckwith (1982) for comparison. Table 7 lists values of n_{crit} for these rate coefficients. It can be seen that the calculated values of n_{crit} differ from the above values typically by factors of 2–3. For densities well below n_{crit} (10^4–10^5 cm^{-3}) the effect of the different rate coefficients is dramatic. Table 8 shows the line ratios calculated for $1-0/2-1\,S(1)$ and $0-0\,S(13)/1-0\,O(7)$ for gas at 2000 K and densities from 10^4 to 10^7 cm^{-3}, for each of the three sets of rate coefficient. It is clear that interpreting the observed H_2 line ratios in terms of physical parameters for the shocked gas is severely hindered by the current uncertainty in the rate coefficients. Further calculations of these quantities would be most welcome.

Despite the uncertainties in the rate coefficients, the calculations show that the population of the $(0,15)$ level is considerably more sensitive to variations in density in the range 10^4–10^6 cm^{-3} (below the critical density) than are the $(1,3)$ and $(2,3)$ levels. This behaviour is shown in Table 9 where the relative populations of the three levels are given as a function of density and temperature. (These results were obtained using the collision rates of Draine *et al.*; for the other

Table 7. Comparison of collision rate coefficients. H_2 number densities at which the populations of the $(v,J)=(0,15)$, $(1,3)$ and $(2,3)$ levels are a given percentage of the LTE population. The critical densities (in cm^{-3}) are calculated as in Table 6, except that the collision rates have been taken from Shull & Hollenbach (1977 – SH) and Shull & Beckwith (1982 – SB). Critical densities derived for the rate coefficients from DRD are listed in Table 6.

Temperature	% of	(0,15)		(1,3)		(2,3)		(0,15)		(1,3)		(2,3)	
(K)	LTE	Shull & Hollenbach						Shull & Beckwith					
1000	50	7	(5)	5	(6)	1.3	(7)	7	(5)	6	(7)	1.3	(8)
	75	2	(6)	1.4	(7)	4	(7)	2	(6)	1.4	(8)	4	(8)
	90	7	(6)	4	(7)	1	(8)	1	(7)	5	(8)	1	(9)
	95	1.4	(7)	8	(7)	2	(8)	3	(7)	1	(9)	>1	(9)
2000	50	4	(5)	4	(5)	1	(6)	4	(5)	1.6	(5)	4	(6)
	75	1	(6)	1	(6)	3	(6)	1.3	(6)	5	(6)	1.3	(7)
	90	3	(6)	3	(6)	8	(6)	4	(6)	1.4	(7)	4	(7)
	95	6	(6)	6	(6)	2	(7)	1	(7)	3	(7)	7	(7)
4000	50	1.5	(5)	3	(4)	1.3	(5)	2	(5)	1	(5)	4	(5)
	75	4	(5)	9	(4)	4	(5)	5	(5)	3	(5)	1	(6)
	90	1	(6)	2	(5)	1	(6)	1.4	(6)	8	(5)	3	(6)
	95	2	(6)	4	(5)	2	(6)	3	(6)	1.4	(6)	6	(6)

Table 8. Variation of line ratios with different rate coefficients. These line ratios are calculated for molecular hydrogen in statistical equilibrium at 2000 K. The different sets of rate coefficients used are from Draine *et al.* (1983 – DRD), Shull & Hollenbach (1977 – SH) and Shull & Beckwith (1982 – SB).

Density	1-0 S(1) / 2-1 S(1)			0-0 S(13) / 1-0 O(7)		
(cm^{-3})	DRD	SH	SB	DRD	SH	SB
1 (4)	29	670	3.6 (3)	7.9 (-4)	1.1 (-2)	6.0 (-2)
1 (5)	27	76	2.3 (2)	1.0 (-1)	0.29	1.2
1 (6)	18	19	35	0.45	0.45	0.84
1 (7)	13	13	15	0.46	0.46	0.50

Table 9. Equilibrium population as a function of temperature and density. The statistical equilibrium populations for the $(v,J)=(0,15)$, $(1,3)$ and $(2,3)$ levels (given as a percentage of the LTE population at each temperature) are calculated using DRD's collision rates.

Density	(0,15)			(1,3)			(2,3)		
(cm^{-3})	1000K	2000K	4000K	1000K	2000K	4000K	1000K	2000K	4000K
1 (4)	4 (-5)	2 (-3)	6	8 (-2)	1	13	3 (-2)	4 (-1)	6
5 (4)	4 (-2)	4 (-1)	44	4 (-1)	5	44	1 (-1)	2	27
1 (5)	4 (-1)	2	65	7 (-1)	9	62	3 (-1)	4	43
2 (5)	2	8	80	1	16	77	6 (-1)	8	62
1 (6)	32	49	96	7	49	95	3	33	90
2 (6)	53	68	98	12	66	97	6	51	95
1 (7)	87	92	100	41	91	100	26	85	99

two sets the same type of behaviour occurs, although quantitatively the results disagree.) Below the critical density, therefore, the $S(13)/O(7)$ ratio is considerably more sensitive to the gas density than is the $1-0/2-1\,S(1)$ ratio.

It should be possible, in principle, to use the combination of the $1-0/2-1\,S(1)$ and $0-0\,S(13)/1-0\,O(7)$ line ratios to probe the temperature and density conditions in the shocked gas. Table 10 lists these line ratios calculated for $T=1000$–4000 K and $n=10^4$–10^7 cm^{-3} for molecular hydrogen in statistical equilibrium, using the cross-sections of Draine *et al.* (1983). For each temperature, the $1-0/2-1\,S(1)$ line ratio varies by a factor of about 2 in the density range 10^4–10^7 cm^{-3}. On the other hand, the variation in this ratio over the temperature range 1000–4000 K, while at a constant density, is by a factor of 10. The line ratio can therefore be used to constrain the temperature. The $S(13)/O(7)$ ratio, on the other hand, is most sensitive to the gas density, at least for $n < n_{crit} \sim 10^6$ cm^{-3}.

Thus, if the collision cross-sections for molecular hydrogen were accurately known it would be possible to determine both the temperature and density of a slab of gas (if the density were below critical) by measuring the $1-0\,S(1)/2-1\,S(1)$ and $0-0\,S(13)/1-0\,O(7)$ line ratios.

The critical densities calculated above are for a purely molecular gas. If the gas is partly atomic then collisions with atomic hydrogen are expected to alter these values. The collision rates for H–H_2 are equally uncertain, but are probably 1–2 orders of magnitude greater than those for H_2–H_2 (Draine *et al.* 1983). If the gas is 10 per cent atomic then H–H_2 collisions may dominate, and even for 1 per cent atomic gas they may be significant. The net effect is to lower

Table 10. Equilibrium values of line ratios. The line ratios are calculated from the statistical equilibrium level populations at each temperature and density. Collision rate coefficients from DRD were used.

Temperature (Kelvin)	Density (cm^{-3})	1-0 S(1) / 2-1 S(1)	0-0 S(13) / 1-0 O(7)
1000	1 (4)	5.3 (2)	3.4 (-6)
	1 (5)	5.1 (2)	2.3 (-3)
	1 (6)	4.8 (2)	2.3 (-2)
	1 (7)	3.2 (2)	1.0 (-2)
	LTE	2.0 (2)	4.9 (-3)
2000	1 (4)	29	7.9 (-4)
	1 (5)	27	0.10
	1 (6)	18	0.45
	1 (7)	13	0.46
	LTE	12	0.46
3000	1 (4)	11	0.10
	1 (5)	8.5	1.3
	1 (6)	5.5	2.0
	1 (7)	4.9	2.1
	LTE	4.8	2.1
4000	1 (4)	6.4	1.9
	1 (5)	4.3	4.6
	1 (6)	3.2	4.5
	1 (7)	3.0	4.4
	LTE	3.0	4.4

the critical density for a given level, perhaps to values of $\sim 10^5$–10^6 cm^{-3}. The H–H_2 collision rates are relatively larger for vibrational transitions than for pure rotational transitions than the H_2–H_2 rates. Therefore, critical densities for low-J levels will be reduced more than for high-J levels in $v=0$, but otherwise the qualitative features of the H_2 line ratios discussed above will still apply. In view of the additional uncertainties which would be introduced, we have not calculated line ratios for mixtures of molecular and atomic gas.

4.2 APPLICATION TO SHOCK–EXCITATION MODELS

When a shock wave passes through molecular gas it rapidly excites the translational degrees of freedom of the molecules, heating the gas to temperatures of several thousand degrees. Inelastic collisions then populate the vibrational and rotational levels and can dissociate the molecules. Radiative decay from these excited levels, and from fine structure lines of atoms, and the collisional dissociation of the molecules (notably H_2), are responsible for the cooling of the gas, resulting in the dissipation of the kinetic energy of the shock wave. Lines-of-sight through the shocked gas will usually sample a wide range of conditions, from cool pre-shock gas, through a thin layer of extremely hot post-shock gas and through a thicker layer of cooling gas. The H_2 lines originate in portions of the column where the relevant energy levels are populated.

Calculation of line intensities in shocked gas is hindered by the uncertainty in the collision cross-sections. We have adopted a simple model, discussed in Burton (1986) and in Brand *et al.* (1988b), for a one-dimensional steady-state hydrodynamic shock, to aid our interpretation of the line ratios observed. The model calculates the instantaneous level populations of H_2 along a section through a shock front. Crucial to the model are the dominant cooling processes for the gas and the collision cross-sections for the hydrogen molecule. It appears that the cooling rate can be approximated by a power of the temperature which is close to 5 for

$300 < T < 4000$ K. However, by using cross-sections from three different papers (Shull & Hollenbach 1978; Shull & Beckwith 1982; Draine *et al.* 1983), it has been found that computed line ratios can vary by as much as a factor of 10. The values of the cross-sections are particularly critical in modelling the $S(13)$ line emission. Numerical calculations for the $S(13)/O(7)$ and $1-0/2-1\,S(1)$ ratios give the predicted density dependence, but for $n < n_{crit}$, the detailed behaviour is very sensitive to the cross-sections.

4.3 INTERPRETING THE PRESENT OBSERVATIONS

The $S(13)/O(7)$ and $1-0\,S(1)/2-1\,S(1)$ ratio in the OMC-1 shocked gas region have been found by Brand *et al.* (1988a) to be 1.05 and 10.5, *independent of position in the region*. The constancy of these line ratios in OMC-1 suggests that the values 1.05 and 10.5 are characteristic of 'thermalized' post-shock gas and hence that the post-shock gas density is above the critical density of the (0,15), (1,3) and (2,3) levels in OMC-1. This implies densities in the post-shock gas are greater than $\sim 10^7$ cm^{-3}, if the gas remains molecular.

Indeed the above model, for a hydrodynamic shock driven by the hot gas in an isobaric layer behind it, with the gas in thermal equilibrium (i.e. $n > n_{crit}$) at the local kinetic temperature, produces the observed ratios for $S(13)/O(7)$ and $1-0/2-1\,S(1)$ in OMC-1. In fact, the relative intensities of 19 lines of H_2 observed in OMC-1 fit the model well (Brand *et al.* 1988b).

The small variation of the $1-0/2-1\,S(1)$ ratio among the four sources reported here suggests that the post-shock gas is also thermalized in all of these shocked sources, so the densities are all greater than n_{crit}. The $S(13)/O(7)$ ratio, however, has been found to be significantly higher than 1.05 in two sources, IC 443 and HH 7. In the remainder of this section we consider several possible explanations for this large variation.

(i) The variation in $S(13)/O(7)$ could occur if the density were near, but below critical in the sources with the lowest $S(13)/O(7)$ ratios [OMC-1, CRL 618 and IC 443 (-844, $+635$)], and above critical in those with the highest ratios [HH 7 and IC 443 (-20, -20)]. The $S(13)/O(7)$ ratio would then show more variation than the $1-0/2-1\,S(1)$ ratio with the changes in density between sources (see Table 10). The uniformity of the $S(13)/O(7)$ ratio in OMC-1 (Brand *et al.* 1988a), however, argues strongly for the densities there being above critical.

(ii) A consequence of our simple model is that the $S(13)$ line is emitted in lower density (and hotter) gas than in the $1-0\,O(7)$ line (Brand *et al.* 1988b); if this density is below the critical value in some sources, it will lead to variations in the $S(13)/O(7)$ ratio from source to source. The critical density of the (2,3) level is, however, higher than that of any of the other levels (for all three sets of molecular constants), yet the $1-0/2-1\,S(1)$ line ratio is the same in all four objects. This implies that the density is in excess of n_{crit} for this level. Thus this explanation cannot be correct, unless the relative critical densities of the various upper levels are in serious error.

(iii) A third possible explanation for the variation in the ratio is that the shock model, which fits the OMC-1 observations so well, does not apply to the other sources. We therefore consider the possibilty that magneto-hydrodynamic *C*-shocks occur in the gas rather than hydrodynamic *J*-shocks (e.g. Draine *et al.* 1983). In such shocks the heating is magnetically mediated, holding the gas at a lower, but more constant, temperature over a larger column density than could be provided in a field-free shock. Draine & Roberge (1982) and Chernoff, Hollenbach & McKee (1982) have modelled the line emission from OMC-1 using such a shock, predicting the shock speed and pre-shock density to be (38 km s^{-1}, 7×10^5 cm^{-3}) and (36 km s^{-1}, 2×10^5 cm^{-3}) respectively. A feature of these models is their sensitivity to the

physical conditions in the gas; for instance in the model of Draine & Roberge, changing the density by a factor of 1.5, or the shock velocity by 10 per cent, greatly degrades the agreement between theory and observation. The uniformity of the $S(13)/O(7)$ and $1-0\,S(1)/2-1\,S(1)$ line ratios in OMC-1 would then argue for extremely uniform shock conditions over the entire source, which seems unlikely.

Applying the *C*-shock models of Draine *et al.* (1983) to the present data does not produce a satisfactory fit. Table 11 shows shock parameters derived from such an exercise. The pre-shock densities were chosen to be 10^4 and 10^6 cm^{-3}, and the excitation temperatures derived from the observed $1-0/2-1\,S(1)$ ratio were adopted. Pre-shock densities of 10^4 cm^{-3} are, however, too low to thermalize the levels behind the shock; in addition the predicted value of the $S(13)/O(7)$ ratio is much higher than that observed. For pre-shock densities of 10^6 cm^{-3} the predicted $S(13)/O(7)$ ratio is close to the range observed. The predicted $1-0\,S(1)$ line intensities, on the other hand, are at least an order of magnitude stronger than observed in IC 443 and HH 7, even when corrected for extinction. This would therefore require a low filling factor for the shocked gas in the beam. In addition, the shock velocity is restricted to the same narrow range (30–35 km s^{-1}) in all the sources observed. Although possible, this seems unlikely for four different types of astronomical object.

(iv) In *C*-shock models, different distributions of shock speeds can affect relative line intensities and might possibly explain the observed variations in line ratios. Many of the line intensities calculated in the *C*-shock models, including those of lines observed here, are sensitive (and monotonic) functions of velocity over most or all of the allowed velocity range. These functions differ from line to line, and are most rapidly varying for the highest excitation lines. Therefore, different distributions of shock velocities among the various sources could result in a larger source-to-source variation in the $S(13)/O(7)$ ratio than the $1-0/2-1\,S(1)$ ratio. In this scenario, the H_2 line emission from *C*-shocks is dominated by emission from shocks with the fastest velocities present, up to just less than that required for molecule dissociation. In addition, for any pairs of line ratios that are observed, the source-to-source variation would be largest for that pair arising from the highest upper-state energy levels. Further theoretical investigation of this conjecture is required. Given the large uncertainties in

Table 11. Derived parameters from *C*-shocks. These figures have been taken from the *C*-shock model of Draine *et al.* (1983). The excitation temperature has been derived from the $1-0/2-1\,S(1)$ line ratio. The limits on it cover the observed range of the ratio. The shock velocity, $1-0\,S(1)$ and $0-0\,S(13)$ line intensities, and $S(13)/O(7)$ line ratio are those predicted by the model for these conditions (i.e. density, magnetic field and excitation temperature). The $S(13)/O(7)$ line ratio has been derived assuming that the $1-0\,S(1)$ line intensity is ~ 8 times that of the $1-0\,O(7)$ line.

Pre-Shock Density (cm^{-3})	Pre-Shock Magnetic Field (μG)	Ionisation Fraction	T_{ex} (K)	Shock Speed (km s^{-1})	1-0 S(1) (Wm^{-2}sr^{-1})	0-0 S(13) (Wm^{-2}sr^{-1})	S(13)/O(7)
1 (4)	50	1 (-7)	1800	36	1 (-6)	9 (-7)	7
			2050	46†	2 (-6)	7 (-6)	25
1 (4)	100	1 (-7)	1800	44	2 (-6)	1 (-6)	4
			1900	48†	3 (-6)	4 (-6)	11
1 (6)	500	1 (-8)	1800	29	5 (-5)	2 (-6)	0.3
			2100	33	1 (-4)	2 (-5)	1.6
1 (6)	1000	1 (-8)	1800	32	7 (-5)	4 (-6)	0.5
			2100	35	1 (-4)	1.4(-5)	1.1

†Denotes the dissociation speed of the shock for the specified excitation temperature.

collision rates, it is possible that the other objections of the *C*-shock models (see above) could be reconciled.

(v) It may be possible that 'resonant collisions', which can occur when the energy difference between the initial and final states of a pair of colliding molecules is small (see Shull & Hollenbach 1978), could selectively populate the $(v,J)=(0,15)$ level in comparison with the (1,5) level. For instance, a resonant encounter may allow a molecule in (2,1) to be excited to (0,15) in just one collision. Resonant collisions are, however, only expected to be important if the gas density is just below critical. The constancy of the $1-0/2-1\,S(1)$ line ratio between sources suggests this cannot be true. Thus, it seems unlikely that this mechanism is operating in any of the sources observed.

We are unable to identify any line that might be blended with the $S(13)$ line at the resolution at which it has been observed, to produce an anomalous $S(13)$ line intensity. We also are unaware of any selective excitation mechanism which could be partially responsible for the over-population of the (0,15) level. If fluorescence were occurring in the gas, it would be expected to populate low lying rotational levels of excited vibrational states in preference to high-J levels in the ground vibrational state. Thus, the result that the $S(13)/O(7)$ ratio is higher in some objects than in OMC-1 is quite puzzling, and requires further study.

5 Conclusions

We have observed the $1-0\,S(1)$, $1-0\,S(0)$ and $2-1\,S(1)$ lines of molecular hydrogen at 2.1–2.2 μm and the $1-0\,O(7)$ and $0-0\,S(13)$ lines at 3.8 μm in four different types of galactic sources whose emission is believed to be shock-excited. We have demonstrated that:

(i) The relative intensities of the $1-0\,S(1):1-0\,S(0):2-1\,S(1)$ lines are approximately 10.5:2.5:1 in all the objects observed. There does not appear to be any major variation in these ratios between the sources.

(ii) The $0-0\,S(13):1-0\,O(7)$ line ratio varies from 1.05 in OMC-1 to ~2.3 in the Herbig–Haro object HH 7.

(iii) The excitation temperature derived from the 2 μm lines for the gas is in the range 1800–2100 K for all the sources, whereas for the 3 μm lines it is in the range 2300–3100 K. The gas cannot be characterized by a single excitation temperature in any of the sources observed.

The $1-0/2-1\,S(1)$ and $S(13)/O(7)$ line ratios can be used to probe the temperature and density conditions in a slab of shocked molecular gas, but the present uncertainty in the collision rate coefficients for H_2 molecules prevents an accurate determination of these quantities. The small variation of the $1-0/2-1\,S(1)$ ratio between sources, together with the uniformity of this ratio and of the $S(13)/O(7)$ ratio in OMC-1 (Brand *et al.* 1988a), argues strongly that the post-shock gas is thermalized in each of the objects studied. This implies that the post-shock gas density in molecular gas is greater than 10^7 cm^{-3}. Although the line ratios are well explained by a hydrodynamic jump-shock model in OMC-1, the variation in the $S(13)/O(7)$ ratio in other sources is not easily explained in this way. Several possible explanations for the variation are discussed, including the role of *C*-shocks, but none are entirely satisfactory. The variation in the $S(13)/O(7)$ line ratio requires further study.

Acknowledgments

We wish to thank the UKIRT staff for their assistance in obtaining these observations and PATT for the allocation of telescope time. MGB acknowledges the support of the SERC

through a SERC studentship. This work was done, in part, while MGB held a National Research Council NASA Research Associateship at Ames Research Center.

References

Beck, S. C. & Beckwith, S., 1983. *Astrophys. J.*, **271**, 175.
Beckwith, S., Beck, S. C. & Gatley, I., 1984. *Astrophys. J.*, **280**, 648.
Beckwith, S., Evans, N. J. II, Gatley, I., Gull, G. & Russell, R. W., 1983. *Astrophys. J.*, **264**, 152.
Beckwith, S., Persson, S. E., Neugebauer, G. & Becklin, E. E., 1978. *Astrophys. J.*, **223**, 464.
Brand, P. W. J. L., Toner, M., Webster, A. S., Geballe, T. R. & Williams, P. M., 1988a. *Mon. Not. R. astr. Soc.*, in press.
Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. .R., Bird, M. & Wade, R., 1988b. *Astrophys. J.*, in press.
Burton, M. G., 1986. *PhD thesis*, University of Edinburgh.
Burton, M. G., Geballe, T. R., Brand, P. W. J. L. & Webster, A. S., 1988. *Mon. Not. R. astr. Soc.*, **231**, 617.
Chernoff, D. F., Hollenbach, D. J. & McKee, C. F., 1982. *Astrophys. J.*, **259**, L97.
Dabrowski, I. & Herzbeg, G., 1984. *Can. J. Phys.*, **62**, 1639.
Draine, B. T. & Roberge, W. G., 1982. *Astrophys. J.*, **259**, L91.
Draine, B. T., Roberge, W. G. & Dalgarno, A., 1983. *Astrophys. J.*, **264**, 485.
Fischer, J., Simon, M., Smith, H. A., Geballe, T. R. & Storey, J. W. V., 1987. *Astrophys. J.*, **320**, 667.
Gatley, I., Hasegawa, T., Suzuki, H., Garden, R., Brand, P., Lightfoot, J., Glencross, W., Okuda, H. & Nagata, N., 1987. *Astrophys. J.*, **318**, L73.
Green, S., Ramaswamy, R. & Rabitz, H., 1978. *Astrophys. J. Suppl. Ser.*, **36**, 483.
Knacke, R. F. & Young, E. T., 1981. *Astrophys. J.*, **249**, L65.
Lane, A. P. & Bally, J., 1986. *Astrophys. J.*, **310**, 820.
Nadeau, D., Geballe, T. R. & Neugebauer, G., 1982. *Astrophys. J.*, **253**, 154.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G. & Ridgway, S. T., 1982. *Astrophys. J.*, **253**, 136.
Shull, J. M. & Hollenbach, D. J., 1978. *Astrophys. J.*, **220**, 525.
Shull, J. M. & Beckwith, S., 1982. *Ann. Rev. Astr. Astrophys.*, **20**, 163.
Toner, M. P., 1986. *MPhil thesis*, University of Edinburgh.
Treffers, R. R., 1979. *Astrophys. J.*, **233**, L17.
Treffers, R. R., Fink, U., Larson, H. P. & Gautier, T. N., 1976. *Astrophys. J.*, **209**, 793.
Turner, J., Kirby-Docken, K. & Dalgarno, A., 1977. *Astrophys. J. Suppl. Ser.*, **35**, 281.
Zealey, W. J., Williams, P. M. & Sandell, G., 1984. *Astr. Astrophys.*, **140**, L31.

Mon. Not. R. astr. Soc. (1989) **237**, 1009–1018

The velocity profile of the 1 − 0 S(1) line of molecular hydrogen at Peak 1 in Orion

P. W. J. L. Brand and M. P. Toner * *Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ*

T. R. Geballe and A. S. Webster *Joint Astronomy Centre, 665 Komohana Street, Hilo HI 96720, USA*

Accepted 1988 October 31. Received 1988 October 28; in original form 1988 September 20

Summary. The profile of the 1–0 S(1) line of molecular hydrogen at Peak 1 is presented, and is shown to pose problems for many models of the excitation region.

1 Introduction

Since the work of Nadeau & Geballe (1979) and Nadeau, Geballe & Neugebauer (1982) it has been known that the velocity profiles of excited molecular hydrogen lines in Orion, with widths of tens of km s^{-1}, imply widespread highly supersonic motion. Many star-forming regions are known to contain supersonic gas flows, which are thought to originate near, and be powered by, an infrared point source (e.g. IRc2 in the case of Orion). The high velocity range of the excited H_2 has presented a considerable puzzle to model builders, since Kwan (1977) established that molecular hydrogen would be dissociated in a shock travelling through the gas at a velocity greater than 25 km s^{-1} (or 40 km s^{-1} in the case of shocks with magnetic precursors in high density gas, Draine 1980). We present a velocity profile of the $v = 1 - 0$ S(1) line of molecular hydrogen which we measured at Peak 1 (Beckwith *et al.* 1978), at the relatively high spectral resolution of 12 km s^{-1}, and analyse possible explanations of the profile.

2 Observations

The observations of Peak 1 in Orion [$5^h\,32^m\,46\overset{s}{.}2$, $-5°\,24'\,02''$, (1950)] were made in November 1983, at the UK Infrared Telescope, using the UKIRT Fabry–Perot interferometer with Cooled Grating Spectrometer as order isolater. The sky aperture was 5 arcsec, and the full width at half maximum of the interferometer bandpass was 12 km s^{-1}. The interferometer was stepped and intensities were measured at 8 km s^{-1} intervals, ensuring nearly ideal Nyquist sampling. Wavelength and wavelength stability were checked with a krypton lamp, and no drift was detected throughout the observation. Nine spectra, each with 246 s integration

* Present address: ITeC, Dumfries, Scotland.

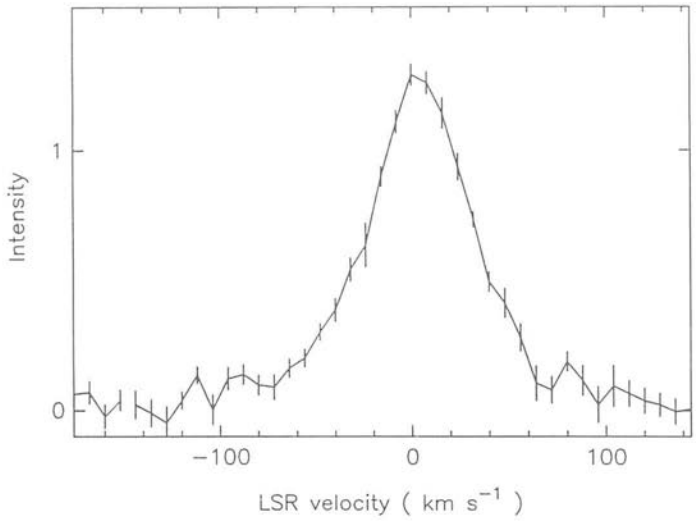


Figure 1. The observed velocity profile at Peak 1 in OMC-1. The beam diameter is 5 arcsec. The velocity is with respect to the local standard of rest.

time, were obtained and coadded, the result being shown in Fig. 1. The velocities in the figure have been corrected to local standard of rest. The full width at zero height is 140 km s^{-1}, while the peak of the profile is within one or two km s^{-1} of the velocity of the surrounding CO cloud (Toner 1986; Goldsmith, Plambeck & Chiao 1975). The spectrum appears rather smooth at this resolution, which is higher than that used by the above-mentioned previous observers, and shows no sign of breaking into narrow emission features. This is investigated in Section 3.2.

3 Discussion

Difficulties abound in any model of the flow which gives rise to this velocity profile. It is known (Nadeau & Geballe 1979; Nadeau *et al.* 1982) that high velocities occur throughout the Orion outflow, and that the gas is shocked to high temperatures throughout (Beckwith *et al.* 1978; Brand *et al.* 1988). Thus the process underlying this observation at Peak 1 is widespread. Furthermore, the shocked gas shows a range of velocities that is highly supersonic, within a region bounded by the viewing aperture whose projected size is small compared with the distance to the assumed source of the wind, IRc2. This poses major problems in the re-direction of momentum, as is now shown.

It will be assumed for the moment that we are not observing hydrogen molecules reformed after being dissociated in a shock. As a consequence, the high velocities we observe cannot be those with which the gas has passed perpendicularly through a shock, since these velocities are restricted to less than either 40 km s^{-1} or 25 km s^{-1}, as mentioned in the introduction. A possibility is that the observed velocities are different projections of the velocity of the gas parallel to a shock front, since the parallel velocity – at least in non-magnetic flows – is not constrained. This will be investigated by examining three types of source structure: (1) a single dynamical system with variable aspect which will provide a velocity range for shocked gas (a bow shock), (2) the possibility of a 'swarm' of shocks reproducing the profile, and finally (3) less well-defined dynamical schemes.

3.1 A SINGLE SHOCK STRUCTURE

First we investigate the structure of a bow shock in a molecular wind. The shock front will be wrapped around a dense intrusion, and there will therefore be a range of angles with respect to the shock front with which the wind impacts the shock (Hartmann & Raymond 1984). This in turn implies a range of shocked gas velocities, each with a particular projection in the line-of-sight. In order to determine the constraints on observable velocities, the relationship between line-of-sight velocity and impact angle in a plane oblique shock is examined in the appendix.

This yields a simple result. The range of line-of-sight velocities V'_{obs} from stationary shocks in a given wind must be less than V_{max}, whether V_{max} is 40 km s^{-1} or 25 km s^{-1}. In particular for $V_{max} = 40$ km s^{-1}, if $V_w = 40$ km s^{-1} then $\Delta V'_{obs} = 39.7$ km s^{-1}. If $V_w \gg 40$ km s^{-1} then $\Delta V'_{obs} = 35$ km s^{-1}, and the viewing angle with respect to the wind approaches a right angle. It is quite impossible to produce a range of observed velocities greater than V_{max}.

To demonstrate this limit in a realistic situation, the velocity profile from a bow shock in a 200 km s^{-1} molecular wind was calculated. Details of the calculation are in the appendix. The following assumptions simplify the calculation without surrendering the essentials:

(i) The shocked material is 'stunned' by the shock, and retains only its velocity parallel to the front. This is reasonable, since it is shown in the appendix that the perpendicular velocity of the shocked gas is small.

(ii) The observed intensity per unit area is proportional to the perpendicular incident component of velocity, $I = I_0 \sin \alpha$. This is correct for a constant density wind and emission by a transition that is in thermal equilibrium, as is probably the case for the 1 − 0 S(1) line.

(iii) The shape of the bow shock is chosen to be paraboloidal, to approximate the conditions of bluntness and of curvature in the regions of interest. Large deviations from the paraboloidal shape (e.g. Mach cone behaviour) occur only at low values of emissivity (e.g. Raga & Böhm 1986).

(iv) The range of angles of incidence, α, is limited by the incident velocity perpendicular to the shock. Above V_{max} the molecules are dissociated, and below V_{min} they are not excited.

The result derived in the appendix is that if $V_{max} = \infty$ and $V_{min} = 0$ (i.e. dissociation does not occur, and excitation always does), then the intensity per unit velocity interval is

$$J(V) = (1 + \cos^2 \theta - 2(V/V_w) \cos \theta)/|(V/V_w) - \cos \theta|^3, \quad (1)$$

where θ is the angle between viewing direction and the axis of symmetry.

Viewed end on, the profile is $(1 - V/V_w)^{-2}$, while sideways on it is $|V/V_w|^{-3}$. In any case it is a very concave profile, which will not produce observed profiles with FWHM greater than ~0.1 V_w, and only provides a profile with zero central velocity if viewed sideways on.

Furthermore, substitution of $V_{min} = 5$ km s^{-1} and $V_{max} = 40$ km s^{-1} actually makes a very minor difference (suppressing the very weak line wings). The results for $V_w = 200$ km s^{-1} at various viewing angles are shown in Fig. 2. The original restriction, to flows in which molecular hydrogen is not dissociated, can be relaxed. Even if dissociation and re-formation occurs, this type of flow cannot explain the observation.

3.2 MULTIPLE SHOCKS

It appears to be impossible to produce the observed broad profile in a single bow shock flow. Consider next an ensemble of emitters, each with its own velocity, covering a range of up to or greater than 140 km s^{-1}. That is, each contributor to the profile must consist of rapidly moving molecular material, through which a shock is currently propagating. There must be a sufficient

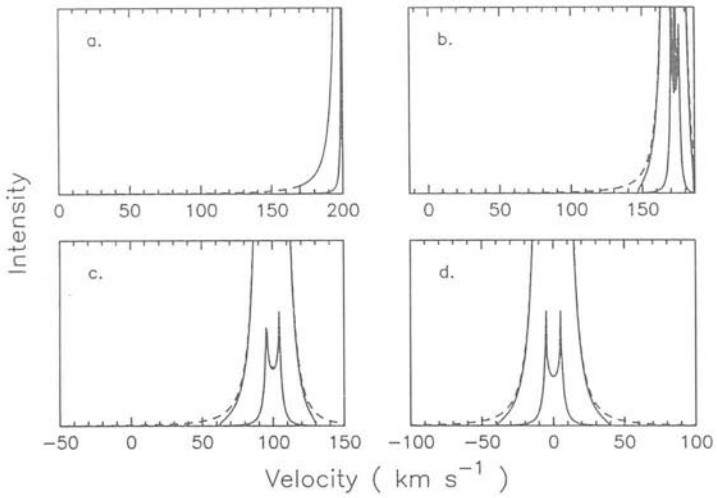


Figure 2. Profiles of intensity versus velocity from the bowshock model, observed at various angles with respect to the axis of symmetry: (a) along the axis, (b) at 30°, (c) at 60°, (d) at right angles. The cavity in the centre is due to the suppression of radiation from sub-5 km s^{-1} gas.

number of these contributors that they emit a distribution of intensity which explains the profile of Fig. 1.

A simple model consists of instrumental velocity profiles (a Lorentz profile with full width at half maximum of 12 km s^{-1}) centred at 8 km s^{-1} intervals, the measured positions, throughout the 140 km s^{-1} range of the observation. The amplitudes A_j of these profiles are allowed to be free parameters. Such a model was least-squares fitted to the observation of Peak 1. Unsurprisingly, a perfect fit was achieved.

Next, the amplitudes were quantized. That is, each amplitude A_j was replaced by

$$A_j' = \eta \times (\text{integer nearest } A_j/\eta) \tag{2}$$

where $\eta = \Sigma A_j / N_{\text{tot}}$ and N_{tot} is a chosen value of 'number of emitters'. This produces a profile which would be observed if the emitting region consisted of a number of equally bright regions, each with zero internal velocity, the regions being distributed in radial velocity and with total intensity equal to that observed at Peak 1. The model profile was recomputed from these amplitudes, and the χ-squared value of the fit calculated as a function of N_{tot}. Ideally, the individual emitters should have been fitted by moving them in velocity space, but since this was virtually fully sampled (8 km s^{-1} with 12 km s^{-1} FWHM in a Lorentz profile), it was considered that negligible error was made by computing the number of emitters at each sampled interval, instead.

The results are shown in Fig. 3. A minimum of 30–40 cloudlets is necessary to reproduce the observed profile satisfactorily.

In order for this scheme to form the basis of an explanation, a means of accelerating a large number of cloudlets in a variety of directions must be found. The cloudlets must achieve velocities of up to ±70 km s^{-1} in the line-of-sight without being completely dissipated, and – if this process has dissociated the hydrogen molecules – time must elapse for the molecules to reform, before a range of weaker shocks, with velocities in the range 10–30 km s^{-1}, propagate through the cloudlets. The range of velocities along the line-of-sight must be moderately symmetrical about $V = 0$ km s^{-1}, and the observed spread has to occur in a cylinder with cross-section 4×10^{14} m across the line-of-sight through Peak 1, since the observation was made with

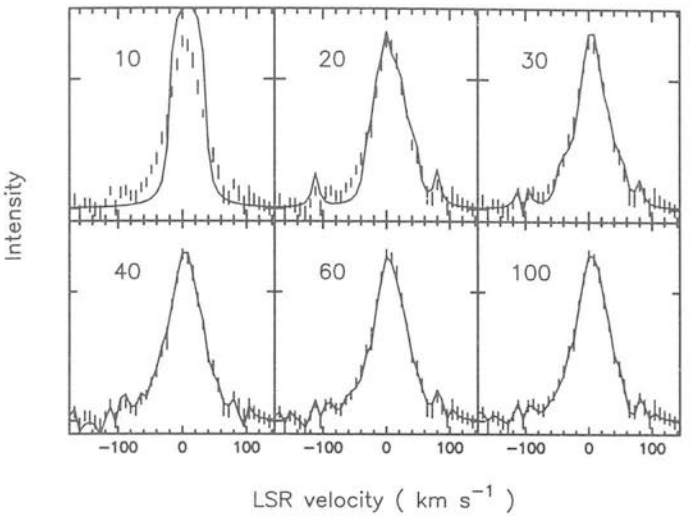


Figure 3. Fits to the data shown in Fig. 1, with several values of the number of equal intensity components. Each component has a velocity profile identical to the instrumental profile. The smoothness of the empirical profile implies that more than 30 independent components are required.

a 5 arcsec beam diameter. A very efficient momentum redistribution into the line-of-sight must occur, presumably by a pressure-driven – rather than momentum-driven – mechanism. Otherwise it is possible to envisage 'bullets' sprayed into a cone centred on IRc2, with a velocity range and angular distribution matched to the profile. The axis of the 'gun' must be aligned precisely perpendicular to the line-of-sight, to explain the high degree of symmetry about 0 km s^{-1}. This introduces so many *ad hoc* assumptions that it requires much more research before it can be regarded as a plausible explanation.

3.3 RE-FORMING MOLECULES

It should be noted that the above multiple-shock scheme probably requires re-formation of molecules behind a strong shock which provides acceleration. Then we rely on re-formation of molecules on the surfaces of dust grains downstream after the molecular material has cooled. However, it is known (Brand *et al.* 1988) that the emitting molecular gas is very hot and therefore cannot have cooled sufficiently to have re-formed on dust grains. Furthermore, observations of 3 μm H_2 lines by Geballe *et al.* (1986) also display broad profiles. It is very important to determine whether the profiles of all the emission lines are the same, in which case it is hard to escape the conclusion that the emission cannot be due to the shock that accelerated the gas, and the molecules must have re-formed, with all the complications mentioned in the last section, or that some still physically ill-defined process is in operation, which is considered below. If, on the other hand, the profiles are significantly different, then there is at least some hope that it is possible to disentangle a series of superposed shocks, although at present this seems the less likely alternative.

3.4 OTHER POSSIBILITIES

There may be other means of producing this breadth of line.

If the wind from IRc2 is collimated, and Peak 1 is a working surface, then the flow may look like the diagram in Fig. 4, which displays a situation considered by Dyson (1987) to be a

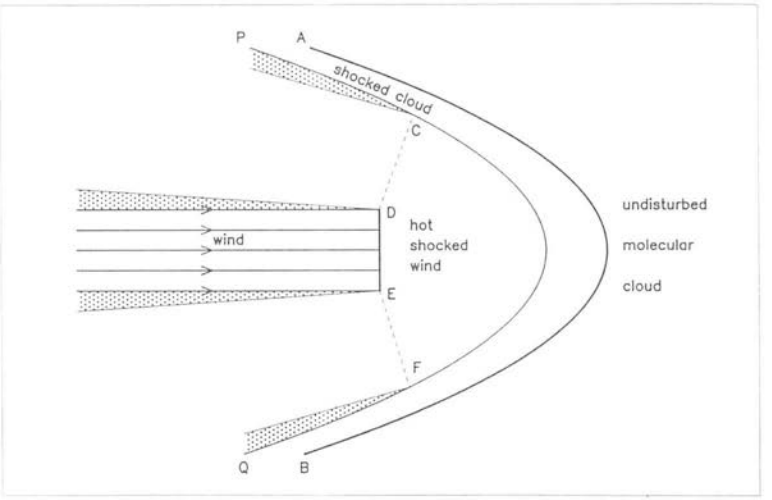


Figure 4. A representation of the region at the end of a high-velocity wind jet driving a shock AB into a molecular cloud. The jet is stopped by shock DE, and the resulting hot gas forms a contact discontinuity PQ with the shocked cloud material. The shocked wind material may escape down the sides of the wind, creating turbulent boundary layers shown by the stippled regions in the diagram.

possible mechanism for Herbig–Haro objects. The wind (presumed to be highly supersonic) is stopped at a shock DE. Behind this shock, the high-pressure shocked wind drives a second shock AB into the molecular cloud. Between the two shocks, the hot shocked wind gas is bounded by shock DE and the contact discontinuity with the shocked cloud gas, PQ. If the density between the cavity and the wind is low, the hot gas will accelerate, and the annular constriction at CD, EF will cause the sonic point to occur near there. The hatched areas are boundary layers, the one of interest being that next to PQ, since it lies next to molecular gas.

In this flow, where does the rapidly moving molecular gas appear? If we suppose we view this more or less sideways on, then the outward velocity of the shock AB can contribute. However, the bow shock analysis can be easily applied to show that we cannot produce high enough velocities or a broad enough velocity profile here.

The remaining possibility seems to be in the boundary layers at P, Q. In the vicinity of F, the net lateral velocity of the hot gas will be $\sim V_w/4$ (at the sonic point) and will be highly supersonic with respect to the molecular gas behind shock AB, presumed to have cooled by the time it reaches PQ. The Kelvin Helmholtz instability is suppressed by supersonic motion. Nevertheless, we seem to be forced to consider this region – both in the collimated jet picture, and in the bow shock – to find sufficient material in differential motion at high enough velocity to produce the width of profile required. Allowing for projection effects in both schemes, wind velocities of at least ~ 300 km s^{-1} seem to be required.

Another possibility may rescue the plane shocks we earlier considered. It was suggested (Scoville *et al.* 1982) that instabilities might broaden the emission lines. Currently we (Moorhouse, Brand, Dalgarno & Hartquist, in preparation) are investigating thermal instabilities (McCray, Stein & Kafatos 1975; Falle 1981) in molecular shocks to see whether the profile we have observed could be due to a non-stationary formation of cascades of shocks in a molecular flow, analogous to the multiple-shock structures produced behind supernova shocks (Innes, Giddings & Falle 1987).

Finally, and most speculatively, it is conceivable that magnetic stresses are sufficient to provide a range of velocities in magnetically accelerated material, where dissipation is reduced

to a level that excites, but does not destroy, the molecules. This provides tight constraints on ambipolar diffusive heating (which cannot be too great), while requiring Alfvén velocities of well over 100 km s^{-1}. If we use the densities suggested by observation (Brand *et al.* 1988), the required field strengths are of the order of 60 milligauss. However, there is no theory to describe turbulent magnetic field dynamics.

The above arguments concerning the structure of the Orion outflow are of a global nature, but are based on measurements at only one position. Although it is possible to make progress in this way, it is imperative to extend the high-resolution study to many adjacent positions. The original data of Nadeau & Geballe (1979) show that significant changes in the velocity structure occurs over the face of the nebula, and it is now a pressing matter to obtain high velocity resolution data over the entire outflow region, as originally delineated by Beckwith *et al.* (1978).

4 Conclusion

The profile of the H_2 1 − 0 S(1) line in the Peak 1 of the Orion Molecular Cloud, which has a FWZI of 140 km s^{-1} measured as a resolution of 12 km s^{-1}, sharpens the difficulties inherent in any gas-dynamic explanation of the observed range of velocities. It is shown to be an urgent matter (1) to determine the intrinsic width of a high-excitation line of H_2, to distinguish between cold and hot re-formed H_2 models, (2) to investigate the various non-standard mechanisms suggested in greater depth, to assess their potential as explanations of this very broad profile, and (3) to obtain high velocity resolution spectra at many spatial positions in Orion.

Acknowledgments

We thank the staff of the United Kingdom Infrared Telescope for their friendly and competent assistance. MPT was supported during this work by an SERC studentship.

References

Beckwith, S., Persson, S. E., Neugebauer, G. & Becklin, E. E., 1978. *Astrophys. J.*, **223,** 464.
Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. R., Bird, M. & Wade, R., 1988. *Astrophys. J.*, **334,** L103.
Draine, B. T., 1980. *Astrophys. J.*, **241,** 1021.
Dyson, J. E., 1987. In: *Circumstellar Matter, IAU Symp. 122,* eds Appenzeller, I. & Jordan, C., Reidel, Dordrecht.
Falle, S. A. E. G., 1981. *Mon. Not. R. astr. Soc.*, **195,** 1011.
Geballe, T. R., Persson, S. E., Simon, T., Lonsdale, C. J. & McGregor, P. J., 1986. *Astrophys. J.*, **302,** 693.
Goldsmith, P. F., Plambeck, R. L. & Chiao, R. Y., 1975. *Astrophys. J.*, **196,** L39.
Hartmann, L. & Raymond, J. C., 1984. *Astrophys. J.*, **276,** 560.
Innes, D. E., Giddings, T. R. & Falle, S. A. E. G., 1987. *Mon. Not. R. astr. Soc.*, **226,** 67.
Kwan, J., 1977. *Astrophys. J.*, **216,** 713.
McCray, R., Stein, R. F. & Kafatos, M., 1975. *Astrophys. J.*, **196,** 565.
Nadeau, D. & Geballe, T. R., 1979. *Astrophys. J.*, **230,** L169.
Nadeau, D., Geballe, T. R. & Neugebauer, G., 1982. *Astrophys. J.*, **253,** 154.
Raga, A. C. & Böhm, K. H., 1986. *Astrophys. J.*, **308,** 829.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G. & Ridgway, S. T., 1982. *Astrophys. J.*, **253,** 136.
Toner, M. P., 1986. *PhD thesis,* University of Edinburgh.

Appendix A

An oblique shock

Fig. A1 shows an oblique plane shock in a wind which is travelling left to right with velocity V_w. The shock is stationary in the frame of the figure, to whose plane it is perpendicular, and makes an angle α with the wind. The shocked – and therefore refracted – wind makes a smaller angle δ with the shock plane and has velocity V'_w. The region is being viewed at an angle β with respect to the wind and, for simplicity, in the plane of incidence.

It will be shown that, in such an oblique shock in a wind, the possible range of line-of-sight velocity is small. Refer to Fig. A1 for notation. The velocity component parallel to the shock front is unchanged:

$$V_w \cos \alpha = V'_w \cos \delta. \tag{A1}$$

The observed velocity V'_{obs} is

$$V'_{obs} = V'_w \cos(\alpha - \beta - \delta) \tag{A2}$$

$$= V_w \cos \alpha \cos(\alpha - \beta) + (V'_w \sin \delta) \sin(\alpha - \beta). \tag{A3}$$

Now, $V'_w \sin \delta$ is the perpendicular component of shocked gas velocity, $V_\perp$, and can be estimated as follows. The adiabatic exponent is taken to be 7/5, and the shock is strong, so the compression ratio is 6, and the maximum temperature behind the shock is $T_{max}/K \approx 30\ [(V_w \sin \alpha)/\mathrm{km\ s^{-1}}]^2$. The cooling occurring behind the shock is nearly isobaric, so the temperature through the flow, T', is approximately proportional to perpendicular velocity, or

$$V'_\perp \approx (V'_w \sin \delta)(T'/T_{max}). \tag{A4}$$

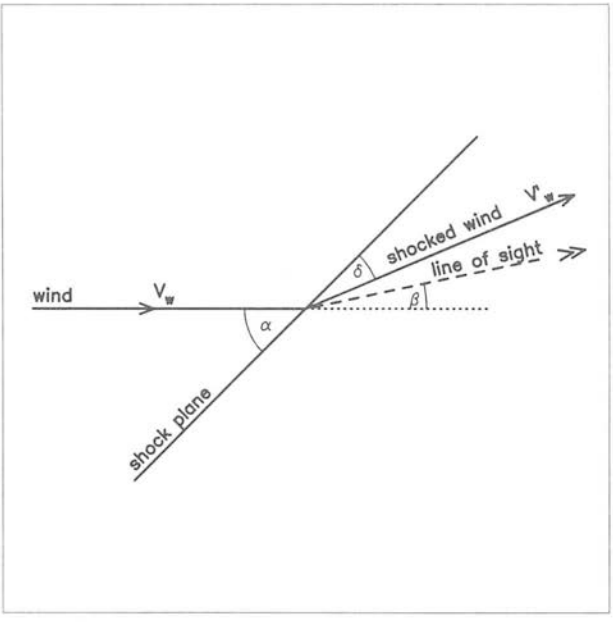


Figure A 1. Diagram of an oblique shock, showing the wind and shocked wind vectors, and the line-of-sight. The shock plane is perpendicular to the diagram, and the shock is stationary in this diagram.

When $T' = 2000$ K, the temperature at which the majority of 1 − 0 S(1) radiation is emitted is

$$V'_{\perp} = 11/(V_w \sin \alpha), \tag{A5}$$

where velocities are in km s^{-1}. Since $V_w \sin \alpha$ must be greater than 5–10 km s^{-1} to excite significantly the $v = 1$ state of H_2, it is clear that $V'_{\perp}$ is of order one or two km s^{-1}. Consequently the second term in equation (A3) will be ignored. The gas is virtually slipping along the back of the shock wave.

With this approximation, the maximum range of observed shocked wind velocity due to varying shock aspect angle α can be deduced from equation (A3), and is given by

$$(\Delta V'_{obs})^2 = V_{max}^2 + V_{min}^2 - 2 V_{max} V_{min} \cos(\alpha_2 - \alpha_1). \tag{A6}$$

Here the velocities V_{max} and V_{min} are the greatest and least permitted incident perpendicular shock velocities, and $\sin \alpha_2 = V_{max}/V_w$, $\sin \alpha_1 = V_{min}/V_w$. V_{max} is either 25 km s^{-1} (J-shock, Kwan 1977) or 40 km s^{-1} (C-shock, Draine 1980). V_{min} is approximately 5 km s^{-1}, required to excite the upper level of 1 − 0 S(1) H_2 transition. The viewing angle with respect to the wind direction that maximizes $\Delta V'_{obs}$ is given by

$$\tan \beta = \frac{V_{min}\sqrt{V_w^2 - V_{min}^2} - V_{max}\sqrt{V_w^2 - V_{max}^2}}{V_{max}^2 - V_{min}^2} \tag{A7}$$

The velocity profile from a bow shock

The paraboloid representing a bow shock (Fig. A2) is taken to be

$$z = aR^2 \tag{A8}$$

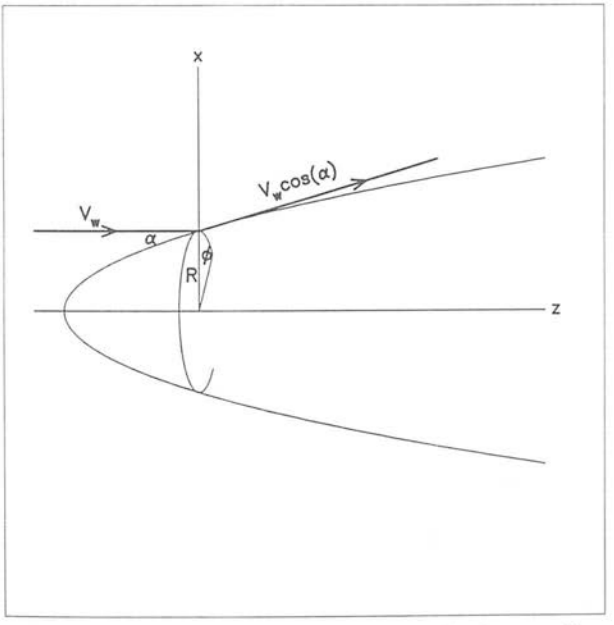


Figure A2. Diagram of the bow shock discussed in the appendix.

with $R^2 = x^2 + y^2$, so that

$$\cot \alpha = \frac{dz}{dR} = 2aR. \tag{A9}$$

The observed velocity component of shocked gas, viewed in the zx plane at angle θ to the z axis – the axis of symmetry – is

$$V/V_w = \cos\theta \cos^2\alpha + \sin\theta \cos\alpha \cos\phi, \tag{A10}$$

where ϕ is the azimuth angle in the xy plane. This result is obtained by assuming [(i) in section 3.1 and first part of appendix] that the perpendicular component of gas velocity with respect to the shock is zero. The element of observed intensity [with $I = I_0 \sin\alpha$, assumption (ii)] is

$$dI = I\,dA = I_0 \sin\alpha \frac{dz}{\cos\alpha} R\,d\phi = I_0 R\,dR\,d\phi \tag{A11}$$

and therefore the profile $J(V)$ is given by

$$J(V) = \frac{dI/dV}{I_0/a^2 V_w}$$

$$= (1/4)\int (\partial\phi/\partial(V/V_w)) \cos\alpha\, d(\cos\alpha). \tag{A12}$$

Substituting from equation (A10) gives

$$J(V) = \int \frac{dc}{(1-c)^2 \sqrt{(pc^2 + qc + r)}} \tag{A13}$$

where $p = 1$, $q = 2\xi\cos\theta$, $\xi = 2V/V_w - \cos\theta$, $r = \sin^2\theta - \xi^2$. The integral can be done analytically to give

$$J(V) = -\frac{\sqrt{R}}{(\cos\theta - \xi)^2} + \frac{1 - \xi\cos\theta}{|\cos\theta - \xi|^3} \arcsin \frac{(\cos\theta - \xi)^2 t - (1 - \xi\cos\theta)}{\sin\theta \sqrt{(1-\xi^2)}} \tag{A14}$$

where $t = 1/(1-c)$, $c \equiv \cos 2\alpha$, and $R = -1 + 2(1 - \xi\cos\theta)t - (\cos\theta - \xi)^2 t^2$. If $V_{max} = \infty$ and $V_{min} = 0$, then the limits of integration are $\arccos 2\alpha = \pm 1$, the quadratic R is zero at the limits, and the arcsin term is $\pm \pi/2$. This proves equation (1) of the paper.

If V_{max}, V_{min} are real limits of dissociation and excitation, then the limits of the integral for $J(V)$ are $c = \cos(2 \arcsin(V_L/V_w))$ where 'L' ≡ 'min' or 'max', *c.f.* the end of the first section of the appendix. These limits reduce the strength of the line wings as shown in Fig. 2 of the paper.

Mon. Not. R. astr. Soc. (1989) **238**, 1513–1527

Shocked molecular hydrogen in the bipolar outflow NGC 2071

Michael G. Burton,[1,2] T. R. Geballe[3,4] and P. W. J. L. Brand[1]

[1] *Department of Astronomy, University of Edinburgh, The Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*

[2] *NASA Ames Research Center, Space Science Division, MS: 245-6, Moffett Field, CA 94035, USA*

[3] *Joint Astronomy Centre, 665, Komohana St, Hilo, HI 96720, USA*

[4] *Foundation for Astronomical Research in the Netherlands – ASTRON*

Accepted 1989 January 16. Received 1988 November 28

Summary. Emission from the $v=1-0$ S(1) line of molecular hydrogen has been mapped in the bipolar outflow NGC 2071. The line emission peaks at six positions distributed irregularly along the two lobes, which extend over a distance of ~ 1 pc. These lobes are parallel to, but offset ~ 20 arcsec from, the lobes of high-velocity CO line emission. Spectra from 2.1 to 2.45 μm of the H_2 emission lines are typical of shock-excited emission. The total H_2 line luminosity is estimated to be $\sim 4.5\, L_\odot$. Profiles of the 1—0 S(1) line are relatively narrow (< 30 km s^{-1} FWHM) for shocked gas. The peak velocity varies systematically across the source, in a manner consistent with the observed bipolarity of the millimetre-wave CO line emission. The H_2 lines probably originate from dense clumps of molecular gas, lying along one edge of the outflow cavity, which are being shocked by a high-velocity wind. There is evidence for partial dissociation of the molecular gas into atomic gas by the shock, but the lack of detectable Br γ line emission puts a low upper limit on the degree of ionization of the gas. The energetics and composition of the high-velocity gas seem consistent with a model in which the driving agent is a bipolar atomic wind, arising from the vicinity of the central IR sources, sweeping up and shocking the surrounding molecular cloud and evacuating a cavity within it.

1 Introduction

Shocked molecular hydrogen lines, emitted at various infrared wavelengths, provide a direct probe of energetic interactions occurring within molecular clouds. The line emission traces the regions where high-velocity gas, typically participating in an outflow from a young stellar source, drives into the molecular cloud, converting some of its bulk energy of motion into

internal energy of the gas, which then radiatively cools. Studying this line emission allows one not only to examine the structure of, and physical processes occurring within, a molecular shock, but also to determine the energetics of the outflowing gas, which facilitates the study of the cause and mechanism of the enigmatic outflow phenomenon.

One particular source which seems well suited to this kind of investigation is NGC 2071, a region of active star formation, located inside the northern part of the Orion B molecular cloud complex at a distance of 500 pc. A molecular outflow near the reflection nebula NGC 2071 exhibits two lobes of well collimated, high-velocity blue- and red-shifted CO gas (e.g. Bally 1982; Snell *et al.* 1984, hereafter SSSE). Coincident with this flow is a region of far-infrared emission with total luminosity $\sim 1000\,L_{\odot}$ (Harvey *et al.* 1979; Sargent *et al.* 1981), which contains a cluster of four 10-μm sources (Persson *et al.* 1981). Three of the 10-μm peaks have associated compact H II regions (Bally & Predmore 1983; Snell & Bally 1986), and H_2O and OH masers are located near the cluster (Schwartz & Buhl 1975; Pankonin, Winnberg & Booth 1977; Campbell 1978; Genzel & Downes 1979). In addition to the high-velocity CO line emission (with FWZI $> 70\,\mathrm{km\,s^{-1}}$), high-velocity lines of CS (Takano *et al.* 1984; Takano 1986) and NH_3 (Takano *et al.* 1985) have been detected. The region of high-velocity line emission is about 5 arcmin in extent, corresponding to 0.75 pc, and the width of the flow observed in the $J=2-1$ CO line is 1–2 arcmin. The high-velocity line emission has been modelled as originating from numerous dense clumps of gas swept up into a thin shell by a collimated stellar wind and ploughing into the ambient molecular cloud (SSSE). Low-velocity CS and NH_3 line emission originates from an elongated structure orthogonal to the high-velocity CO bipolar flow, which has been interpreted as a rotating molecular disc (Takano *et al.* 1986).

Emission from the $1-0$ S(1) line of H_2 in NGC 2071 has been mapped by Bally & Lane (1982), Simon & Joyce (1983) and Lane & Bally (1986, hereafter LB), and recently at high

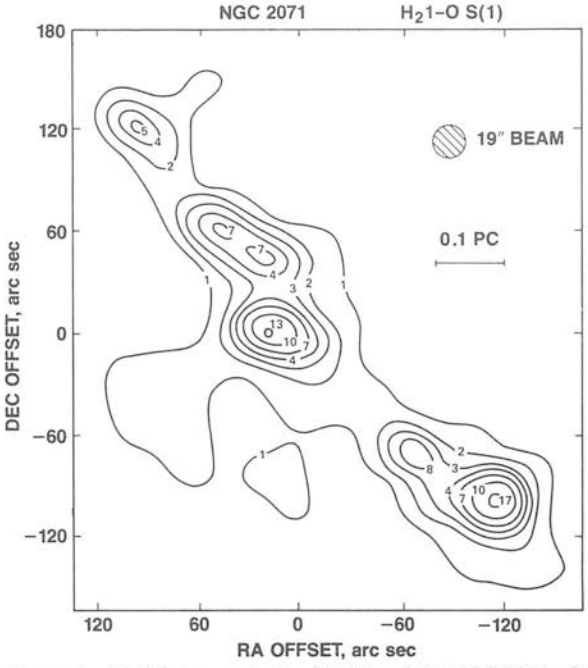


Figure 1. Map of the H_2 $v=1-0$ S(1) line emission (2.122 μm) in NGC 2071. Contours are in units of 7×10^{-21} Wcm^{-2} through the 19 arcsec beam. The (0″,0″) position for the map is the IR source IRS1, at $5^{h}44^{m}30.6^{s}$, $+0°20'42''$ (1950.0). The scale bar assumes that the source distance is 500 pc.

(1 arcsec) resolution by Garden, Russell & Burton (1989). LB found that much of the emission originates from several bright spots distributed along the two lobes. The lobes are parallel to those of the high-velocity CO line emission (SSSE), but are offset by ~15 arcsec. There is a prominent 1 arcmin gap in the emission between the two H_2 lobes. The previous observers have concluded that the H_2 line emission arises in cooling gas, lying behind shock waves generated by the interaction of a high-velocity wind, from a central source, with dense molecular gas.

Of particular interest to the study of shock interactions in molecular clouds is the detection of high-velocity 21-cm atomic hydrogen (H I) line emission (Bally & Stark 1983; Bally 1986). This emission comes from two lobes containing high-velocity blue- and red-shifted gas, with overall morphology similar to that of the H_2 line emission. There is, however, a significant anti-correlation of the H_2 and H I emission peaks, and Bally concludes that the H I is probably produced in dissociative shocks and is unlikely to play an important role in the dynamics of the outflow.

In this paper we present new data on the near-infrared H_2 line emission, particularly the 1–0 S(1) line, in this source. The line emission has been mapped, and found to be more extended than by previous workers. Low-resolution spectra have been obtained at four locations, allowing a determination of the excitation mechanism and extinction to the source. Velocity-resolved profiles of the S(1) line were obtained at three positions in order to study source dynamics. Combining these data with radio data, we discuss possible origins for the high-velocity atomic and molecular gas present, and also a scenario for the outflow involving a neutral, atomic wind sweeping up and shocking ambient cloud gas.

2 Observations

The observations were obtained with the 3.8-m United Kingdom Infrared Telescope (UKIRT), on Mauna Kea, Hawaii, during 1985 January and 1986 January. They involved the use of a circular variable filter (CVF) of velocity resolution $\sim 2500\,\mathrm{km\,s^{-1}}$ and Fabry–Perot interferometers (FPs) with 130 and $35\,\mathrm{km\,s^{-1}}$ resolution, and involved various observing techniques.

The map of the 1–0 S(1) line was obtained by placing the FP with $\sim 130\,\mathrm{km\,s^{-1}}$ velocity resolution in front of the CVF tuned to 2.122 μm. The transmitted light was viewed by an InSb detector through an aperture diameter of 19 arcsec FWHM. The FP was frequency-switched between the line wavelength and a reference frequency at a nearby uncontaminated part of the atmospheric spectrum. The map was made by raster scanning the telescope along a grid of positions covering the source. The integration time for each position was 16 s, of which 8 s were spent on line. As velocity shifts are less than $40\,\mathrm{km\,s^{-1}}$ between positions on the map (see Section 3.3), with linewidths $\ll 130\,\mathrm{km\,s^{-1}}$, we believe that the line intensity distribution derived is accurate.

The resulting S(1) line emission map, smoothed by a Gaussian filter of width 20 arcsec, is presented in Fig. 1. It contains about 400 pixels. Pixel separation within the lobes is 10 arcsec (i.e. fully sampled) and is 20 arcsec outside. The (0,0) position is located on IRS1, at $5^h44^m30.6^s$, $+0°20'42''$ (1950.0), with absolute pointing errors of any pixel estimated to be accurate to better than 10 arcsec. The relative pointing between nearby pixels is accurate to better than 5 arcsec. Wavelength and line flux calibrations were made on the peak of the 1–0 S(1) line in OMC-1, which we have measured to be $(3.5 \pm 0.5) \times 10^{-18}\,\mathrm{W\,cm^{-2}}$ through the same aperture. Absolute intensities are estimated to be reliable to about 20 per cent and the relative intensities of different pixels are accurate to about one unit, $7.0 \times 10^{-21}\,\mathrm{W\,cm^{-2}}$, on the flux scale used on the map.

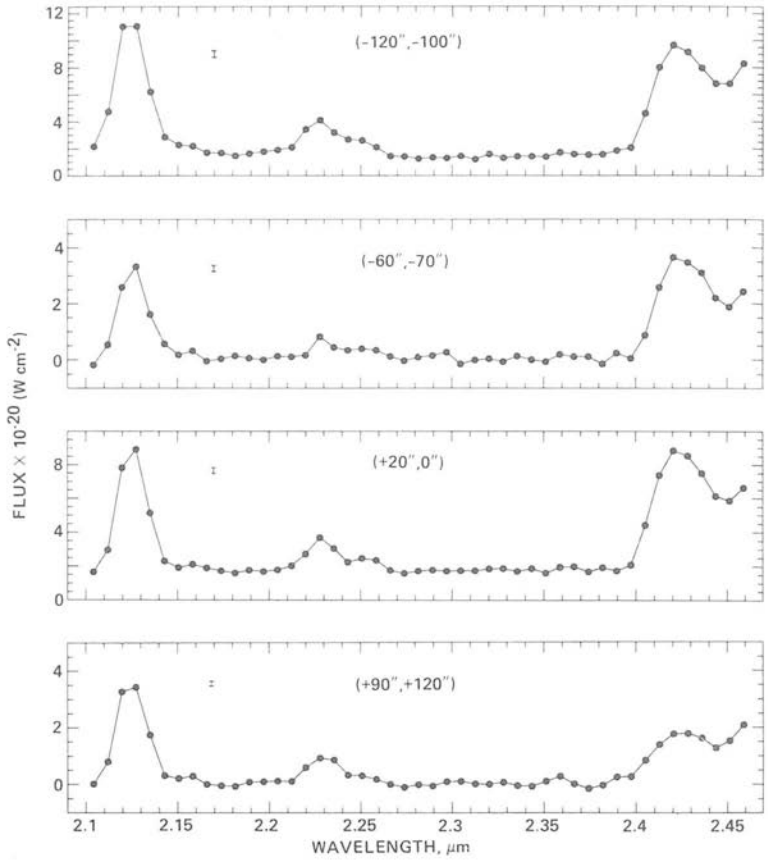


Figure 2. Spectra of the H_2 emission lines from 2.1 to 2.45 μm at four locations in NGC 2071. The beam size is 19 arcsec and the spectral resolution, $\lambda/\Delta\lambda$, is 115. The flux, in W cm^{-2}, is per resolution element. Important lines are labelled. Coordinates are given as offsets from the (0", 0") position of IRS1. A sample $\pm 1\sigma$ error bar is shown for each spectrum.

Low-resolution, fully sampled spectra from 2.10 to 2.45 μm were obtained at four of the brighter positions, and are shown in Fig. 2. This wavelength range includes the H_2 1 − 0 S(1), S(0) and several Q-branch lines, the 2 − 1 S(1) line, and the H I Brγ line. These observations utilized the CVF, a 19 arcsec aperture and a chopper throw of 120 arcsec E–W. Total integration times were about 30 s per position, per wavelength setting (including time spent chopping onto the sky). The intensities were calibrated by observing HD 84800 and HR 2241, assumed to have K magnitudes of 7.53 and 4.00, respectively.

High-resolution profiles of the 1 − 0 S(1) line were obtained at three locations. These observations utilized an FP of velocity resolution 35 km s^{-1}, in series with the CVF (tuned to 2.122 μm) and InSb detector, viewing a 12 arcsec aperture. The FP scanned over a 330 km s^{-1} range in steps of 15 km s^{-1}. Integration times were about 20 min for the two brightest positions, and 50 min for the faintest position. The resultant profiles have been deconvolved using the maximum-entropy method (MEM) (Willingale 1981; Skilling & Bryan 1984). A discussion of the technique appears in Burton & Geballe (1986). The S(1) lines profiles are shown in Fig. 3; for each position the observed profile, deconvolved profile and simulated profile (convolution of the data with the instrumental profile; calculated as a consistency check), are shown. All three observed profiles, and the profile of an argon line, are presented in

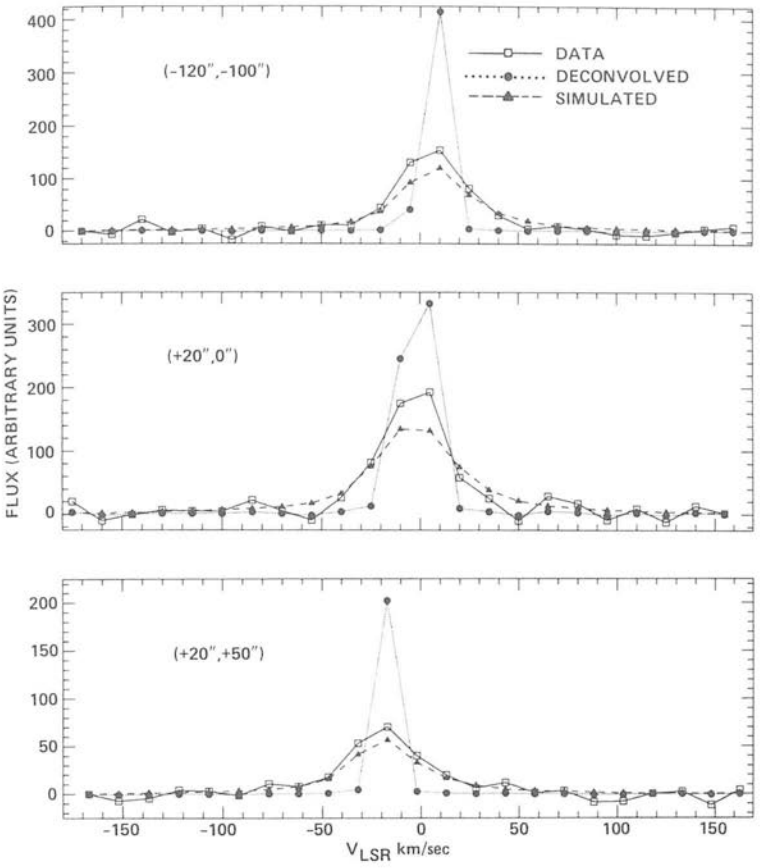


Figure 3. Profiles of the H_2 $v = 1 - 0$ S(1) line at three locations in NGC 2071. For each location three profiles are shown: the observed profile (□'s and continuous line); the deconvolved profile (●'s and dotted line) obtained from the maximum entropy algorithm; and the simulated profile (▲'s and dashed line) (the convolution of the deconvolved profile with the instrumental response, obtained from the profile of an argon line). Fluxes are in arbitrary units. Velocities, in km s^{-1}, are with respect to the local standard of rest. Positions given are offsets from IRS1.

Fig. 4. Since the lines are narrow, the deconvolution algorithm has difficulty in converging to the MEM solution, which would be a δ-function for an unresolved line. This results in the simulated profiles having too weak a flux at line centre.

3 Results

3.1 MORPHOLOGY

From Fig. 1, it is evident that much of the $1 - 0$ S(1) line emission originates from several clumps distributed along two well-defined bipolar lobes, centred on, and extending roughly equal distances from, the infrared source, IRS1. The emission region is comparable in length, ~6 arcmin (~0.9 pc, at an assumed distance to the source of 500 pc), and orientation to that of the high-velocity CO line emission from the source, but is considerably narrower, ~30 arc sec (~0.07 pc), than the 1–2 arcmin width of the CO. The extent of the S(1) line emission is greater than shown in previous maps (e.g. LB). Six distinct emission peaks have been resolved,

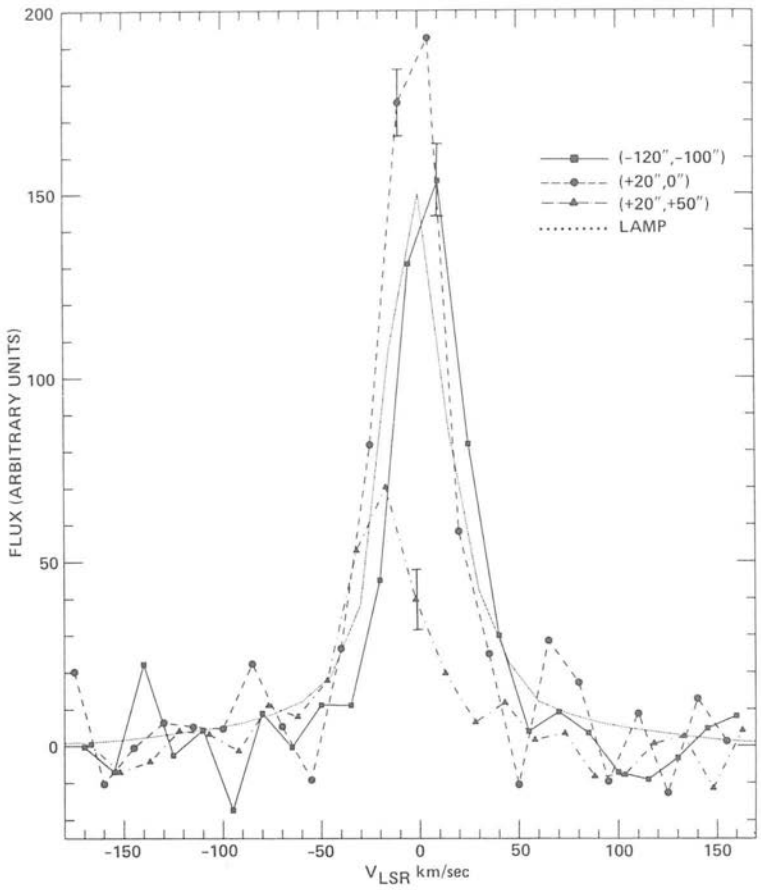


Figure 4. Observed profiles of the $H_2 v=1-0$ S(1) line at three locations in NGC 2071; $(-120'', -100'')$ (■'s wih continuous line), $(+20'', 0'')$ (●'s with dashed line) and $(+20'', +50'')$ (▲'s with dot-dash line). The instrumental response is shown by the lamp profile (dotted line).

four along the north-east (blue-shifted) lobe, and two along the south-west (red-shifted) lobe. These peaks are not co-linear on the sky, but, as described by LB, curve slightly in an extended S-like distribution. The fluxes from the emission peaks are given in Table 1.

Although the morphology of the S(1) line emission determined by us is generally similar to that found by LB, there are some differences. The positions of some of the peaks differ by 5–15 arcsec in the two maps. We find an offset of ~20 arcsec between the H_2 and CO line emission (as mapped by SSSE), slightly larger than that found by LB. The IR cluster of sources (Persson *et al.* 1981) appears to be centred about 15 arcsec south of the brightest peak of the blue lobe as mapped by LB; in our map this clump of emission, although enclosing the IR cluster, is centred about 15 arcsec to the east of it. It is clear from both maps, however, that the IR sources are not located in the gap between the two emission lobes. Finally, we detect faint extended line emission in and near the gap between the two bright lobes, as well as an arc of faint emission extending to the SE of the brightest of the peaks of the north-east lobe; neither of these was detected by LB. The arc may be a region of reflected line emission, as in the source OMC-1 (Hough *et al.* 1986). Polarimetry of the line emission would test this conjecture.

Table 1. H_2 $v = 1-0$ S(1) line fluxes for emission peaks in NGC 2071.

RA Offset	DEC Offset	Line Flux	Column Density
arcsec	arcsec	$\times 10^{-20}$ W cm^{-2}	$\times 10^{+17}$ cm^{-2}
−115	−100	12	13
−65	−70	5.5	5.9
+20	0	10	11
+20	+50	5.1	5.5
+50	+60	6.1	6.5
+95	+120	3.8	4.1

The 1 − 0 S(1) line was measured through the CVF+FP and a 19 arcsec aperture. The offsets refer to a (0,0) position at $5^h44^m30.6^s$, $+0°20'42''$ (1950.0). Absolute positions for the peaks are accurate to better than 10 arcsec. Absolute line fluxes are accurate to about ±20 per cent and relative fluxes to ±10 per cent. The shocked H_2 column densities assume 1 mag of extinction to the source.

3.2 EXCITATION CONDITIONS

The excitation of the emitting gas has been investigated using the spectra in Fig. 2. The 1 − 0 S(1), 1 − 0 S(0) and 2 − 1 S(1) line intensities are listed in Table 2, together with the column densities of the upper-level of each line. The hydrogen Br γ line, at 2.166 μm, was not detected at any position and, adopting a conservative 3σ upper limit, has line flux $< 3 \times 10^{-21}$ Wcm^{-2}

Table 2. H_2 line fluxes and column densities.

RA offset	DEC Offset		1–0 S(1)		1–0 S(0)		2–1 S(1)		1–0 Q	
arcsec	arcsec		2.122μm		2.223μm		2.248μm		2.41–2.42μm	
-120	-100	Flux	1.1	(-19)	2.8	(-20)	1.3	(-20)	1.9	(-19)
		Error	4	(-21)	3	(-21)	3	(-21)	9	(-21)
		N_{upper}	6.4	(+15)	2.3	(+15)	5.4	(+15)		
-60	-70		3.7	(-20)	6.5	(-21)	3.1	(-21)	7.8	(-20)
			9	(-22)	9	(-22)	9	(-22)	6	(-21)
			2.2	(+15)	5.5	(+14)	1.3	(+14)		
+20	0		8.4	(-20)	2.1	(-20)	8.1	(-21)	1.5	(-19)
			2	(-21)	2	(-21)	2	(-21)	1	(-20)
			4.9	(+15)	1.8	(+15)	3.5	(+14)		
+90	+120		4.1	(-20)	1.1	(-20)	2.6	(-21)	4.6	(-20)
			2	(-21)	1	(-21)	1	(-21)	6	(-21)
			2.4	(+15)	9.2	(+14)	1.1	(+14)		

Fluxes are in Wcm^{-2}; for each object, the first row is the line flux and the second row the 1σ error. The third row is the upper-level column density in cm^{-2}. The figure in parenthesis is the exponent [i.e. (−19) denotes 10^{-19}].

Table 3. Line ratios.

RA	DEC	1–0 S(1)/2–1 S(1)	1–0 S(0)/2–1 S(1)	1–0 S(1)/1–0 S(0)	1–0 Q-branch/1–0 S(1)
-120	-100	9.1	2.3	4.0	1.7
		[7.1 – 11.3]	[1.7 – 3.0]	[3.5 – 4.5]	[1.5 – 1.8]
-60	-70	13.3	2.3	5.8	2.1
		[9.3 – 17.7]	[1.4 – 3.4]	[4.9 – 6.8]	[1.9 – 2.3]
+20	0	10.7	2.7	4.0	1.8
		[8.5 – 13.0]	[2.1 – 3.5]	[3.6 – 4.3]	[1.7 – 2.0]
+90	+120	18.4	5.0	3.7	1.1
		[10.9 – 27.5]	[2.8 – 7.8]	[3.3 – 4.2]	[0.9 – 1.3]

Two rows of data analysis are presented. In the first is the mean value of line ratio. In the second is the range for the ratio, derived from the 1 σ errors of the line fluxes.

through a 19 arcsec beam. Line ratios for the observed H_2 lines are presented in Table 3. The relative line intensities are similar at the four positions observed. The average value of the 1 – 0 S(1):1 – 0 S(0):2 – 1 S(1) line ratio is 10.7:2.5:1.0. This is very similar to the value observed in OMC-1 (10.3:2.5:1.0 – see Burton *et al.* 1988), a source in which the H_2 line emission is shock-excited.

For any line pair, the excitation temperature corresponding to the emitting levels is given by

$$I_1/I_2 = \frac{A_1 \nu_1 g_1}{A_2 \nu_2 g_2} \exp[(E_2 - E_1)/kT], \quad (1)$$

where I is the observed line intensity, A the radiative decay rate for the transition (taken from Turner, Kirby-Docken & Dalgarno 1977), g the statistical weight of the level, ν the line frequency and E the energy level of the upper state of the transition (taken from Dabrowski & Herzberg 1984). The (small) differential extinction between the wavelengths of the above three H_2 lines has been neglected. These temperatures are listed in Table 4. The average vibrational

Table 4. Excitation temperatures.

RA	DEC	1–0 S(1) / 2–1 S(1)	1–0 S(0) / 2–1 S(1)	1–0 S(1) / 1–0 S(0)
offset	offset	Kelvin	Kelvin	Kelvin
-120	-100	2100	2100	1200
		[2000 – 2400]	[1900 – 2300]	[900 – 1600]
-60	-70	1900	2100	*
		[1700 – 2100]	[1800 – 2500]	
+20	0	2000	1900	1100
		[1900 – 2200]	[1800 – 2100]	[900 – 1400]
+90	+120	1700	1600	1000
		[1500 – 2000]	[1500 – 1900]	[800 – 1400]

For each position, the first row contains the excitation temperature derived from the mean value of the line ratio. The second line is the range in this value obtained from the 1 σ range in the line ratios. * Excitation temperature is ill-determined.

excitation temperature between the $v=1$ and $v=2$ levels is 1950 ± 75 K, and the average rotational temperature between $J=2$ and $J=3$ in $v=1$ is 1100 ± 175 K. These are lower than the corresponding values measured by LB (2690 ± 240 K and 1420 ± 240 K), but typical of values measured in shocked molecular sources (e.g. Burton *et al.* 1989). We note, however, that H_2 line emission from a dense photo-dissociation region can simulate that from a shocked region, if the molecular gas has been heated to temperatures of the order of 1000 K by the process of UV-fluorescence followed by collisional de-excitation, in densities $\geq 10^{5-6}$ cm^{-3} (e.g. Hollenbach 1988; Burton, Hollenbach & Tielens, in preparation). For conclusive evidence of shock-excitation it is therefore necessary to study the motions of the emitting gas.

3.3 DYNAMICS

Profiles of the $v=1-0$ S(1) line in three locations are shown in Figs 3 and 4. Line parameters derived are listed in Table 5. The line was essentially unresolved at this resolution, at all three positions, implying that the intrinsic FWHMs at these locations are <30 km s^{-1}. In addition, we do not find evidence for a 100 km s^{-1} wide profile, as reported by Persson *et al.* (1981), who used a 6-arcsec beam on the peak located near the IR cluster. There is a systematic gradient in the S(1) line emission velocity, running from $+10$ km s^{-1} V_{LSR} at the peak of the 'red' lobe, through 0 km s^{-1} at the peak of the 'blue' lobe, to -17 km s^{-1} in the peak at $(+20'', +50'')$. The velocity range in the 'blue' lobe is comparable to the range in velocities of the CO line emission from that lobe (SSSE), but the S(1) line velocity at the single position observed in the 'red' lobe is similar to that of the ambient molecular cloud.

Although our data do not provide unequivocal evidence for high-velocity gas motions in the H_2 line emitting gas at any one location, they show that relative velocities of at least 25 km s^{-1} exist between emitting regions, and that the velocities are in some way related to the high-velocity outflow seen in the CO lines. The velocity shifts, the overall similarity between the CO and H_2 line maps, the relative line ratios at the peaks, and the evidence for high velocity motions in other species such as H I and CS, lead us to conclude that the H_2 line-emitting gas is shock-excited, driven by the high-velocity outflow present in the source. This conclusion is in agreement with those of previous workers (e.g. LB).

Table 5. Parameters of the H_2 $1-0$ S(1) line profiles.

RA	DEC	Total Flux	Central Flux	V_{LSR}	FWZI		FWHM	
offset	offset	arbitrary units		km/s	km/s		km/s	
-120	-100	4.7	4.2	+10	45	(-20,+25)	16	(+2,+18)
+20	0	6.2	3.3	0	45	(-25,+20)	28	(-15,+13)
+20	+50	2.2	1.4	-17	45	(-42,+3)	16	(-25,-9)

These parameters were derived from the deconvolved line profiles. Observations were made through an 11 arcsec aperture. Fluxes are in arbitrary units; the 'total flux' is the area under the line and the 'central flux' is the flux in the central velocity channel (of width 15 km s^{-1}). An approximate calibration is that 1 unit $\equiv 8.4 \times 10^{-21}$ W cm^{-2}. V_{LSR} is the estimated velocity of the peak emission with respect to the local standard of rest (with error ± 5 km s^{-1}). FWZI and FWHM are the full-width zero-intensity and half-maximum velocities of the line. Figures in parenthesis give the red- and blue-extent of the emission. The ambient molecular cloud has velocity $V_{LSR} = +10$ km s^{-1} (Snell *et al.* 1984).

3.4 EXTINCTION

The extinction to the H_2 line emitting gas can be crudely estimated from the ratio of the $v=1-0$ Q-branch lines (2.4 μm) to the $1-0$ S(1) line (2.12 μm) (see Scoville *et al.* 1982, Burton *et al.* 1988). We compare the Q/S(1) ratio with that measured in OMC-1 (2.5 ± 0.1) on the same night. The observed ratios in NGC 2071 suggest that the extinction is smaller than to OMC-1; indeed some of the ratios imply zero extinction. Because of the narrow widths of the H_2 lines, however, the observed Q-branch intensities can be severely affected by telluric line absorption. The extinction at 2.2 μm to OMC-1 probably lies in the range 0.5–2 mag (e.g. Beckwith *et al.* 1983; Brand *et al.* 1988), with the most recent measurements suggesting it lies towards the low end of this range. We adopt 1 mag as the extinction at 2.2 μm to NGC 2071, but recognize that there are severe uncertainties in this estimate. LB adopted 2.7 mag based on estimates of the molecular column density to the core, but we believe this estimate to be too high from our comparison with OMC-1.

3.5 PHYSICAL PARAMETERS FOR THE SHOCKED GAS

The total $1-0$ S(1) line flux detected from NGC 2071 is 1.5×10^{-18} W cm^{-2}, which corresponds to a source luminosity of 0.12 $L_\odot$ and a mass of hot gas in the $(v,J)=(1,3)$ level of 2.3×10^{-6} $M_\odot$ (assuming the source distance is 500 pc and not correcting for extinction). Adopting one magnitude of extinction at 2 μm, and correcting for the fraction of emitting gas in the (1,3) level (see Burton *et al.* 1988a), the total luminosity and mass of hot H_2 are $L_{H_2} \sim 4.5\ L_\odot$ and $M_{Hot\,H_2} \sim 4 \times 10^{-4}\ M_\odot$. The former value is about 3 per cent of the luminosity of the shocked H_2 in OMC-1 (Beckwith *et al.* 1983). Column densities of hot, shocked H_2 at each emission peak are listed in Table 1. The maximum value, $\sim 1 \times 10^{18}$ cm^{-2} in each lobe, is about 1 per cent of the greatest column density observed in OMC-1. Assuming that a 10 km s^{-1} shock wave is propagating into the molecular gas, these column densities imply that the average pre-shock density in the beam is of the order of 10^5 cm^{-3}.

4 Discussion

4.1 LOCATION OF THE HOT MOLECULAR HYDROGEN

The present data, in conjunction with the observations of other molecular and atomic lines, provide compelling evidence that the observed H_2 line emission arises largely in formerly ambient cloud material that has been shocked by a high-velocity bipolar wind. It is possible that some of the emission occurs within the wind itself; however, the lack of evidence for high-velocity emission in the observed H_2 line profiles strongly constrains the degree of this activity. The relative uniformity of the CO high-velocity line emission along each lobe, in comparison to the clumpiness of the H_2 line morphology, is further evidence for the interaction of a rather freely outflowing gas with molecular cloud material of varying concentration. Finally, the apparent offset of the CO emission from that of the H_2 implies that the bulk of the H_2 line emission occurs at the edges of the outflow, rather than within it. This is consistent with the model derived by SSSE from their CO data. They suggest that the interaction of a wind with ambient gas occurs mostly at the end of a jet, rather than continuously along its length. In the model, a high-velocity wind fills an evacuated cavity and the swept-up gas, lying along the walls of the cavity, produces the bulk of the high-velocity CO line emission.

Although there are some differences between our S(1) line map and that of LB, we both find that the major axis of the $2-1$ CO line emission region, as measured by SSSE, is offset by some 15–20 arcsec from the axis of the S(1) line map. The IR cluster is located centrally along

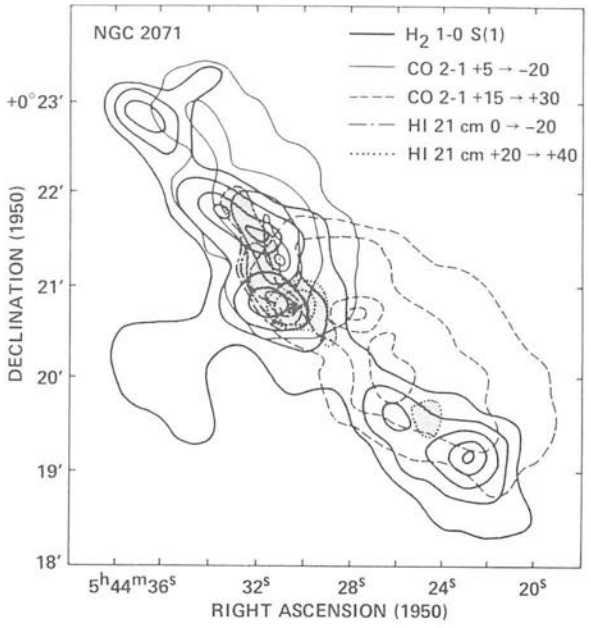


Figure 5. Overlay of maps of the H_2 $v = 1 - 0$ S(1) line (this work, thick solid line), the high-velocity blue- and red-shifted atomic gas (21-cm H I, Bally 1986, shaded regions surrounded by thin dashed and dotted lines) and the blue- and red-shifted molecular outflow ($J = 2 - 1$ CO, SSSE, thin solid and dashed lines) in NGC 2071.

the S(1) axis. It would appear that this offset is real. SSSE estimate their absolute pointing to be accurate to better than 10 arcsec, and determine the centroid of the outflowing CO gas to lie 15 arcsec north-west of the IR cluster. An H_2O maser (Schwartz & Buhl 1975; Campbell 1978; Genzel & Downes 1979) and an OH maser (Pankonin, Winnberg & Booth 1977) are located symmetrically about the IR cluster and offset from the CO axis. Fig. 5 presents an overlay of the S(1) line map and the high-velocity H I and $2 - 1$ CO maps. If this offset is real, it would imply that the shocked molecular gas lies along just one side of the molecular outflow.

The irregular distribution of the H_2 line peaks in the lobes, and the velocity gradient along them, suggests that the jets are not linear features. The distribution is reminiscent of that of the S(1) line emission from the bipolar outflow DR 21 (Garden *et al.* 1986). In DR 21, however, the velocity profiles are broader (up to 150 km s^{-1} wide), with the peak velocity oscillating along the lobes. The narrow widths in NGC 2071 would suggest that at each peak the lines are dominated by emission from a single shock, rather than from an ensemble of clumps moving with varying velocities. The systematic velocity shift could result from a tilt out of the plane of the sky of the outflow, with the blue lobe curving towards us and the red lobe being in the plane of the sky (to account for the emission velocity there being at the same velocity as that of the ambient cloud). However, profiles have only been observed at three of the peaks. Velocity-channel maps may reveal more complex velocity structure.

4.2 ORIGIN OF OBSERVED H I AND H_2

The presence of high-velocity atomic hydrogen and shocked molecular hydrogen, together with the absence of ionized gas in NGC 2071, constrains models of the shock. The mass of high-velocity atomic hydrogen present is only $\sim 0.1\ M_\odot$, and consists of $0.05\ M_\odot$ in the core around the IR cluster and $0.05\ M_\odot$ in the jets (Bally & Stark 1983; Bally 1986), while the mass

of outflowing molecular gas is $\sim 10\ M_\odot$ (SSSE). The 3σ upper limit on the Br γ line intensity, from this work, limits the amount of ionized hydrogen to less than 2 per cent of the $0.0004\ M_\odot$ of hot, post-shock molecular gas.

As discussed in Burton *et al.* (1988) there are several possible origins for the presence of high-velocity atomic and molecular gas in a source. The observed H_2 may be shocked or it may have reformed after being dissociated by the shock. The atomic hydrogen may be dissociated H_2, ambient H I (i.e. cloud material), outflowing H I in the wind, or recombined H II. The spectrum of the H_2 does not resemble that predicted from models of recombining gas (e.g. Duley & Williams 1986; Hollenbach & McKee 1989). This implies that the observed H_2 is indeed shocked. The low limit of ionized gas from Br γ suggests that the observed atomic hydrogen is not recently recombined ionized gas. Likewise, the high-velocity atomic hydrogen clearly cannot be a component of the ambient molecular cloud. Therefore the H I might only be either wind material, or shocked and dissociated H_2. A 21-cm map of the H I (Bally 1986), however, reveals that the H I is offset from the CO in much the same manner as the H_2 line emission (see Fig. 5). This implies that it is associated with the shocked gas, rather than with the freely expanding outflow. A detailed comparison of the H_2 and H I peaks shows that they are anti-correlated. As the degree of dissociation behind a shock depends on its velocity, which is inversely proportional to the density of the cloud, this anti-correlation suggests that the H I peaks correspond to low density regions of the cloud. Where the H_2 line emission peaks, then, the clumps are denser, the shock velocity lower, and the degree of dissociation smaller. This effect could be tested further by obtaining velocity-resolved maps of the H_2 and H I line emission. Similar maps of the O I 63-μm line, which has already been detected in NGC 2071 (Haas, private communication) could also be revealing, as this line is predicted to be particularly intense in fast shocks (Werner *et al.* 1984). In conclusion, we infer that the high-velocity atomic gas in the jets is mostly dissociated molecular material behind shock fronts. We note that if the H I does not then recombine, then the shocks in NGC 2071 cannot be strongly dissociating. The amount of atomic hydrogen generated per year, $\sim 6 \times 10^{-6}\ M_\odot$ (using a dynamical time-scale for the outflow of $\sim 1.6 \times 10^4$ yr; SSSE), is $\ll 4 \times 10^{-4}\ M_\odot$, the amount of hot H_2, which has a lifetime of about 1 yr.

4.3 NATURE AND ENERGETICS OF THE OUTFLOW

In this section we estimate the energy budget of the high-velocity and shocked material in NGC 2071 in order to determine the mechanisms controlling the dynamics of the outflow. In particular we investigate the possibility, suggested by SSSE and Snell & Bally (1986), that the outflow is driven by a neutral wind from a central source. We also attempt to determine the dominant cooling mechanisms at work.

Estimates of the mass loss and wind speed, assuming a steady momentum-driven wind, required to account for the CO data (SSSE) and also for a jet model to account for the shocked H_2 luminosity (LB), give $dM_w/dt \sim 4 \times 10^{-5}\ M_\odot\ \mathrm{yr}^{-1}$ and $V_w \sim 100\ \mathrm{km\ s}^{-1}$ (the 'jet model' estimate is not changed substantially by the data presented here). Over the life of the outflow the total mass loss is $\sim 0.6\ M_\odot$, much less than the mass of outflowing molecular gas ($\sim 10\ M_\odot$). Most of the high-velocity molecular gas must therefore be swept-up material and not original wind material. The mass loss through ionized winds from the compact H II regions around the IR sources is $\sim 2 \times 10^{-6}\ M_\odot\ \mathrm{yr}^{-1}$ (Snell & Bally 1986), much less than the required mass loss rate to power the outflow. Thus the ionized gas can at most be a small fraction of the wind material.

Conserving momentum between wind material, filling an evacuated cavity, and outflowing

molecular gas, lying in a surrounding shell, the wind mass, M_w, is given by

$$M_w = M_f \left(\frac{V_f}{V_w} \right)^2 , \qquad (2)$$

where M_f is the mass of outflowing molecular gas and V_f the mean outflow velocity ($\sim 8\,\mathrm{km\,s^{-1}}$) (SSSE; see their equation 9). Thus $M_w \sim 0.06\,M_\odot$, comparable to the mass of high-velocity 21-cm H I in the vicinity of the IR sources (Bally & Stark 1983). As the wind expands it is diluted and the emission from the highest velocity gas becomes too weak to detect. The H I gas observed in the lobes is probably dissociated gas, as discussed in the previous section.

The mechanical luminosity injected into the neutral wind, $1/2\ dM_w/dt\ V_w^2$, is $\sim 30\,L_\odot$. This is very much less than the stellar luminosity ($L_{FIR} \sim 1000\,L_\odot$, Sargent *et al.* 1981), but an order of magnitude larger than the mechanical luminosity of the outflowing molecular gas ($\sim 4\,L_\odot$, SSSE). The majority of the energy input to the flow must be radiated away after shock heating the gas, or go into dissociation of the molecules. This energy transfer occurs behind the shock front. Our estimate of the H_2 line luminosity ($\sim 4.5\,L_\odot$) demonstrates that it is a significant cooling mechanism for the gas. These estimates are, of necessity, crude, and we cannot yet determine whether molecular hydrogen line radiation is the dominant coolant for the gas, or whether it may be, say, [O I] 63-μm line emission or molecular dissociation.

To summarize, the momentum estimates for the molecular, atomic and ionized material present are consistent with the molecular outflow being driven by a fast, atomic wind from the neighbourhood of the IR cluster. Molecular hydrogen line radiation is an important cooling agent, but may not be the dominant cooling mechanism at work.

4.4 LOCATION OF THE SOURCE OF THE OUTFLOW

In all likelihood, the source of the wind is located near the gap between the lobes of H_2 line emission. Clearly, one or more of the several infrared sources found near the gap, approximately centred relative to the axis of the bipolar flow, are plausible candidates for the source of the outflow. Compact H II regions are associated with these objects (Snell & Bally 1986), but the near-infrared recombination lines have not been detected (Smith *et al.* 1987), indicating an extinction to these objects that is much higher than to the coincident H_2 line emitting region.

If one or more of the above infrared objects provides the driving wind, then the brightest peak in the north-east (blue-shifted) lobe, on which the infrared sources are superposed, may not actually be a part of the bipolar flow, as it would not be situated in one of the lobes of the flow. Instead it would constitute shocked gas either in front of, or behind the outflow source, or surrounding it. In addition, the gap itself would then probably have little significance; it would not to be due to a disc-like structure, as such a structure would heavily obscure the 10-μm sources and the superimposed H_2 emission. However, the H_2 line profile at this position is narrow, giving no indication that this is shocked gas directly along the line-of-sight to the wind source. Therefore, one perhaps cannot rule out the possibility that an object which is hidden in the gap between the lobes is the actual source of the outflow in NGC 2071. One should note, however, that such an object would not be symmetrically placed between the two CO lobes.

5 Conclusions

A map of the H_2 1 – 0 S(1) line has been obtained in NGC 2071. The emission displays a bipolar morphology, similar to that of the high-velocity CO and H I line maps. The two lobes of

emission are centred on a cluster of IR sources and offset from the axis of the CO lobes. The bulk of the S(1) line emission arises from 6 peaks distributed irregularly along the lobes. Spectra of the 1 − 0 S(1), 1 − 0 S(0) and 2 − 1 S(1) lines demonstrate that the emission is typical of shock-excited sources. Velocity profiles of the S(1) line are narrow ($< 30\,\mathrm{km\,s^{-1}}$ FWHM), but show a systematic variation of the emission velocity, from 'red' to 'blue', moving from the peak of the red lobe out to the blue lobe. The total 1 − 0 S(1) line flux is $1.5 \times 10^{-18}\,\mathrm{W\,cm^{-2}}$, which corresponds to a total shocked H_2 line luminosity of $\sim 4.5\,L_\odot$ assuming 1 mag of extinction. A model for the outflow is discussed, involving a high-velocity, bipolar, neutral wind from the vicinity of the IR sources sweeping up ambient gas and evacuating two cavities. The shock, which is concentrated alongside one edge of the cavity, partially dissociates the molecular gas into atomic gas, but the total mass of the dissociated gas, and the mass lost through the wind, is small compared to the total mass of molecular gas swept up and accelerated. Molecular hydrogen line radiation behind the shocks is an important mechanism for ridding the wind of its energy.

Acknowlegments

We wish to thank the UKIRT staff for their assistance in obtaining these observations and PATT for the allocation of telescope time. The United Kingdom Infrared Telescope is operated by the Royal Observatory Edinburgh on behalf of the UK Science & Engineering Research Council. MGB acknowledges the support of the SERC through a SERC studentship and also the National Research Council for a NASA Research Associateship held at Ames Research Centre.

References

Bally, J., 1982. *Astrophys. J.*, **261,** 558.
Bally, J., 1986. In: *Masers, Molecules and Mass Outflows in Star Forming Regions,* Proc. of Meeting held at Haystack Observatory, Westford, MA, USA on 1985 May 15–16, ed. Haschick, A. P., p. 179. Haystack Observatory Publication.
Bally, J. & Lane, A. P., 1982. *Astrophys. J.*, **257,** 612.
Bally, J. & Predmore, C. R., 1983. *Astrophys. J.*, **265,** 778.
Bally, J. & Stark, D. A., 1983. *Astrophys. J.*, **266,** L61.
Beckwith, S., Evans, N. J. II, Gatley, I., Gull, G. & Russell, R. W., 1983. *Astrophys. J.*, **264,** 152.
Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. R., Bird, M. & Wade, R., 1988. *Astrophys. J.*, **334,** L103.
Burton, M. G. & Geballe, T. R., 1986. *Mon. Not. R. astr. Soc.*, **223,** 13P.
Burton, M. G., Geballe, T. R., Brand, P. W. J. L. & Webster, A. S., 1988. *Mon. Not. R. astr. Soc.*, **231,** 617.
Burton, M. G., Brand, P. W. J. L., Geballe. T. R. & Webster, A. S., 1989. *Mon. Not. R. astr. Soc.*, **236,** 409.
Campbell, P. D., 1978. *Publs astr. Soc. Pacif.*, **90,** 262.
Dabrowski, I. & Herzberg, G., 1984. *Can. J. Phys.*, **62,** 1639.
Duley, W. W. & Williams, D. A., 1986. *Mon. Not. R. astr. Soc.*, **223,** 177.
Garden, R., Russell, A. P. G. & Burton, M. G., 1989. *Astrophys. J.*, submitted.
Garden, R., Geballe, T. R., Gatley, I. & Nadeau, D., 1986. *Mon. Not. R. astr. Soc.*, **220,** 203.
Genzel, R. & Downes, D., 1979. *Astr. Astrophys.*, **72,** 234.
Harvey, P. M., Campbell, M. F., Hoffman, W. F., Thronson, H. A. Jr. & Gatley, I., 1979. *Astrophys. J.*, **229,** 990.
Hollenbach, D., 1988. *Astr. Lett. &Comm.*, **26,** 191.
Hollenbach, D. & McKee, C. F., 1989. *Astrophys. J.*, in press.
Hough, J. H. *et al.*, 1986. *Mon. Not. R. astr. Soc.*, **222,** 629.
Lane, A. P. & Bally, J., 1986. *Astrophys. J.*, **310,** 820 (LB).
Pankonin, V., Winnberg, A. & Booth, R. S., 1977. *Astr. Astrophys.*, **58,** L25.
Persson, S. E., Geballe, T. R., Simon, T., Lonsdale, C. J. & Bass, F., 1981. *Astrophys. J.*, **251,** L85.
Sargent, A. I., van Duinen, R. J., Fridlund, C. V. M., Nordh, H. L. & Aalders, J. W. G., 1981. *Astrophys. J.*, **249,** 607.

Schwarz, P. R. & Buhl, D., 1975. *Astrophys. J.*, **201,** L265.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G. & Ridgway, S. T., 1982. *Astrophys. J.*, **253,** 136.
Simon, T. & Joyce, R. R., 1983. *Astrophys. J.*, **265,** 864.
Skilling, J. & Bryan, R. K., 1984. *Mon. Not. R. astr. Soc.*, **211,** 111.
Smith, H. A., Fischer, J., Geballe, T. R. & Schwartz, P. R., 1987. *Astrophys. J.*, **316,** 265.
Snell, R. L. & Bally, J., 1986. *Astrophys. J.*, **303,** 683.
Snell, R. L., Scoville, N. Z., Sanders, D. B. & Erickson, N. R., 1984. *Astrophys. J.*, **284,** 176 (SSSE).
Takano, T., 1986. *Astrophys. J.*, **303,** L85.
Takano, T., Stutzki, J., Fukui, Y. & Winnewisser, G., 1986. *Astr. Astrophys.*, **158,** 14.
Takano, T., Stutzki, J., Winnewisser, G. & Fukui, Y., 1985. *Astr. Astrophys.*, **144,** L20.
Takano, T., Fukui, Y., Ogawa, H., Takaba, H., Kawabe, R., Fujimoto, Y., Sugitani, K. & Fujimoto, M., 1984. *Astrophys. J.*, **282,** L69.
Turner, J., Kirby-Docken, K. & Dalgarno, A., 1977. *Astrophys. J. Suppl.*, **35,** 281.
Werner, M. W., Crawford, M. K., Genzel, R., Hollenbach, D. J., Townes, C. H. & Watson, D. M., 1984. *Astrophys. J.*, **282,** L81.
Willingale, R., 1981. *Mon. Not. R. astr. Soc.*, **194,** 359.

Paper presented at the ASP Centennial meeting, Berkeley, CA, June 21-23, 1989 "Evolution of the Interstellar Medium", ed. L. Blitz, publisher Astronomical Society of the Pacific.

HIGH SPECTRAL AND SPATIAL RESOLUTION IMAGING OF THE SHOCK WAVES IN ORION

MICHAEL BURTON[1,2], JOSS BLAND[3], D. AXON[4], P. BRAND[5], R. GARDEN[2], T. GEBALLE[6], D. HOLLENBACH[1], J. HOUGH[7], I. McLEAN[6], A. MOORHOUSE[5]
[1]NASA Ames, MS:245-6, Moffett Field, CA 94035, [2]U.C. Irvine, [3]Rice University, [4]Jodrell Bank Radio Observatory, [5]University of Edinburgh, [6]Joint Astronomy Centre Hilo, [7]Hatfield Polytechnic

ABSTRACT Orion has been imaged in the v=1–0 S(1) line of molecular hydrogen, with 0.6" pixel scale and 12 km s^{-1} spectral resolution, and the most detailed picture yet of the shocked emission obtained.

Despite considerable effort over the past decade, the molecular hydrogen line emission from the OMC–1 in Orion has remained an enigma. It is clear that shocks are responsible for the phenomenom (*e.g.,* Beckwith *et al.* 1978, *Ap. J.,* 223, 463), and the velocity profiles of the excited lines, with widths of up to 150 km s^{-1} (*e.g.,* Nadeau *et al.* 1982, *Ap. J.,* 253, 154), imply widespread supersonic motions. Accounting for the shape and width of the profile has remained a problem since Kwan (1977, *Ap. J.,* 216, 713) established that H_2 would be dissociated in jump-shocks moving faster 25 km s^{-1} (or $\sim$ 40 km s^{-1} in magnetically-mediated shocks, Draine 1980, *Ap. J.,* 241, 1021). At the highest spectral and spatial resolution the line emission has been measured (12 km s^{-1} and 5", Brand *et al.* 1989, *M.N.,* 237, 1009) the profile has remained smooth, relatively symmetrical, has a strong blue wing, and shows no hint of breaking into narrow emission features. The most plausible model seems to be that it arises from a conglomeration of discrete clumps, each individually accelerated to different velocities by a wind from the source IRc2, and being shocked. However, as Brand *et al.* (1989) discuss, this requires a minimum of 30–40 cloudlets along *every* line of sight to the source (a cylinder with cross-section $\sim$ 0.01pc, small compared to the source size of $\sim$ 0.2pc, yet with a velocity dispersion $\pm$75 km s^{-1}). There are considerable difficulties in explaining the redirection of momentum from the source of the wind which ultimately drives the shocks.

The observations were obtained on the UKIRT telescope in January 1989 using the IR array camera IRCAM with 0.6" pixel scale. A Fabry–Perot interferometer (FP) with resolution 12 km s^{-1} was placed in the beam, and scanned in steps of 5 km s^{-1}, over the profile of the H_2 1–0 S(1) line, at 2.1218 μm, covering a 36" $\times$ 38" region at the peak of the H_2 line emission from OMC–1. The surface of constant phase (a paraboloid) was determined, and the phase shift across each frame deconvolved from the data (see *e.g.,* Bland *et al.* 1987, *M.N.,* 228, 595). In the figure are shown profiles of the line emission for one

quadrant of the image. It must be cautioned that the reduction is still preliminary and the wavelength scale is non-uniform. Further work remains to linearise the scale and remove some artifacts from the data. Nevertheless, the results clearly demonstrate that H_2 line profiles with excellent S/N can now be obtained, with both high spectral and spatial resolution.

The first impression from the data is, despite the increased spatial resolution, that the profiles remain wide, have extended blue wings, and are generally smooth, all over the source. The problem of redirection of the wind momentum is exacerbated. There is, however, some structure in the profiles. Line splitting is seen for the first time in several locations, with the secondary component being a high-velocity blue-shifted feature. Examination of the images reveals that this arises from several filamentary, or 'jet-like' features, which emanate from the vicinity of IRc2. It is tempting to speculate that they delineate the conduits along which a high-velocity wind from IRC2 is flowing, with the H_2 emission arising from accelerated molecular gas swept-up and entrained along them. This then poses the question of what kind of collimation mechanism might lead to more than one jet escaping from the source?

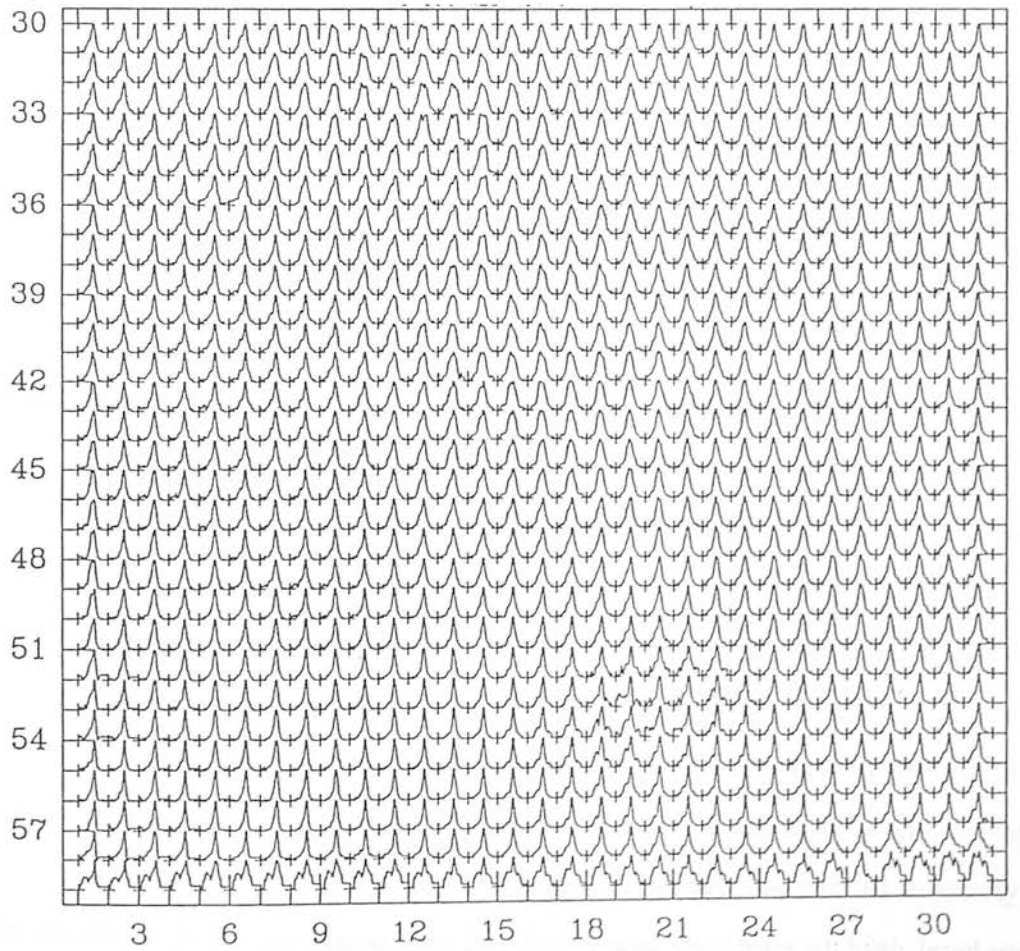


Figure 1: Line profiles of the H_2 1–0 S(1) line for one quarter of the data, covering a portion of the peak emission region of OMC–1. Fluxes are scaled to the peak intensity in each pixel. N is to the bottom and E to the left.

THE NATURE OF SHOCKS IN MOLECULAR CLOUDS

Peter W. J. L. Brand
Department of Astronomy,University of Edinburgh
Royal Observatory Edinburgh EH9 3HJ, U.K.

Abstract Evidence is presented to suggest that the shocked molecular hydrogen emission in the brightest part of the Orion outflow is produced in a J-shock and not a C-shock; that this is true throughout the entire flow; that it may be true in many outflow sources; and that this exacerbates problems with current explanations of the very wide velocity profiles observed in molecular hydrogen emission.

1. THE DATA

This paper summarizes work done at the U.K. Infrared Telescope (supported by the Science and Engineering Research Council) in Hawaii, by a group including M.Burton, M.Bird,T.Geballe, A.Moorhouse, M.Toner, R.Wade, A.Webster, and P.Williams. Following the discovery of shocked molecular hydrogen in Orion (Gautier *et al.* 1976), difficulties faced by early hydrodynamic models (*eg* Hollenbach 1981) led to the general acceptance of magnetically-moderated C-shock models (Draine 1980, Draine & Roberge 1982, Chernoff *et al.* 1982).

The results of our measurements are shown in figure 1. These inferred column densities cover a much wider range of level energy than previously used in model fits, and clearly are at variance with published C-shock models (the dashed lines). On the other hand, the cooling zone behind a hydrodynamic shock is modelled by the continuous line, and is evidently a good fit. Two parameters are required, the overall intensity (the curve is drawn through the strongest line 1-0 S(1)) and the slope at high level energy which is determined by the pressure driving the shock.

That is the evidence that the position at which these data were obtained, Peak 1 in the Orion molecular outflow (Beckwith *et al.* 1978), is excited not by a C-shock but by a hydrodynamic (or J-) shock. The evidence that this situation is outflow-wide is

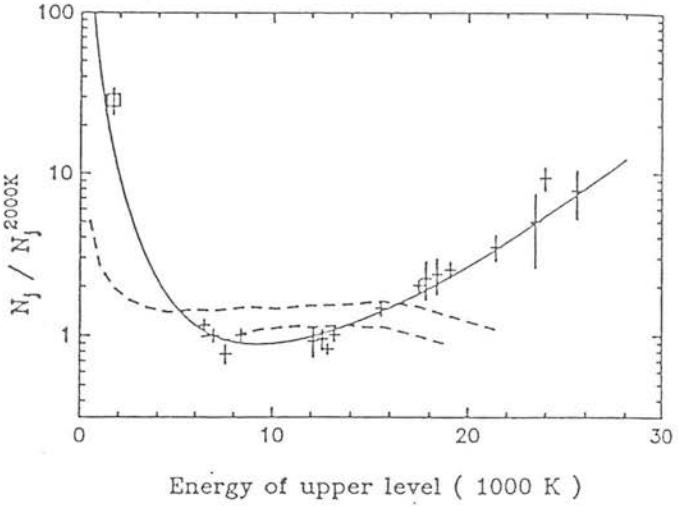
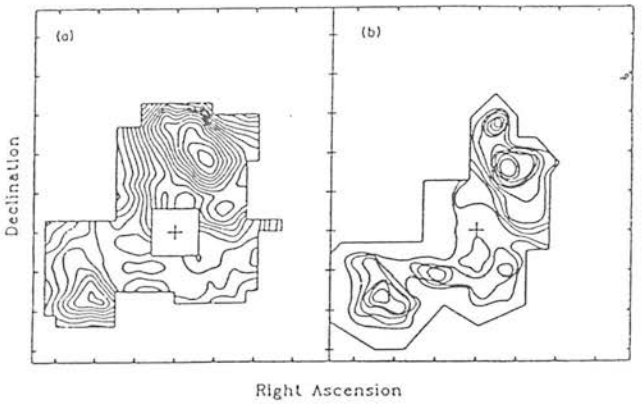


Figure 1(top). Column density derived from the observations at Peak 1 in Orion divided by those expected from 2000K gas. The dashed lines are C-shock models, the solid line is a J-shock model.



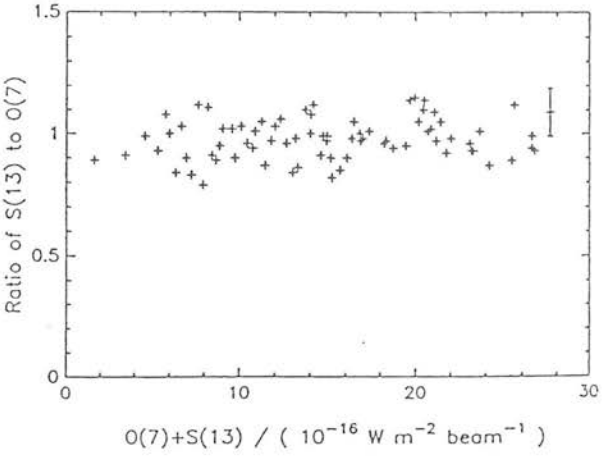
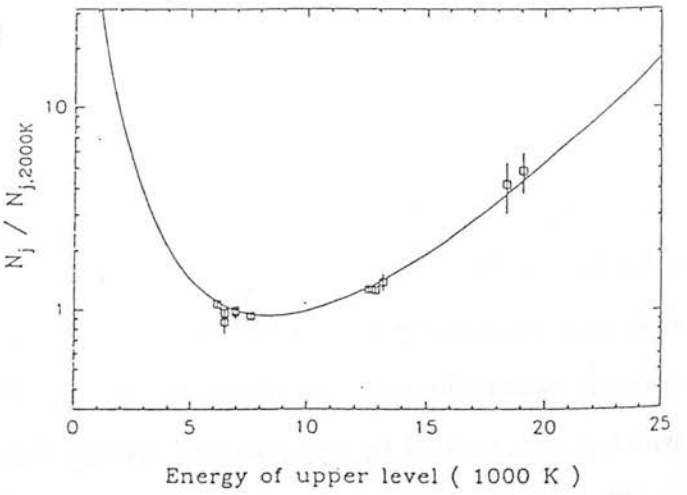


Figure 2(centre left). 3.8μm map(left) and 2.1μm map from Beckwith *et al*(right) of the Orion molecular outflow.

Figure 3(centre right). The 3.8μm line ratio from all parts of the outflow, showing no trend with intensity.

Figure 4(bottom). The same as figure 1, for data from the brightest part of the IC 443 H_2 emission



shown in figure 2, which consists of a map of intensity of the sum of two 3.8μm lines of H_2 throughout the outflow. The right-hand map is drawn from the original 2.1μm map of Beckwith *et al.* for comparison. It is clear the morphology is the same. The lines in question are the 1-0 O(7) line with an upper level energy of 8365 K and the 0-0 S(13) line with an upper level energy of 17445 K. Because the wavelengths are nearly identical, there are no problems de-reddening the data. From figure 1 it is clear that the ratio is a sensitive indicator of the shape of the right-hand part of the model curve.

Figure 3 displays this ratio from all positions in this map, ordered according to intensity. It is clear that there is no significant trend in the data, and that the curve of figure 1 applies to all parts of the outflow. That is the evidence for J-shocks throughout the outflow.

Similar data obtained from the brightest part of the H_2 emission in the supernova remnant IC 443 (Burton *et al.* 1988) is shown in figure 4. Again a J-shock-like pattern of column densities of H_2 lines is produced. Burton *et al.* (1989) show that the line ratios in several other sources are consistent with this interpretation. These are the data on which is based the assertion that J-shock excitation appears to be common in shocked molecular clouds.

2. THE THEORY

Figure 5(a) is a diagram of the cross section through a hydrodynamic J-shock, showing the cooling zone where the temperature drops at a rate determined by the local cooling function. The cooling is supposed due to collisionally excited H_2 either dissociating or de-exciting by line radiation, and to line emission from CO cooling. This can by crudely approximated by a steep power law, which is assumed in this argument. Figure 5(b) shows the same flow in temperature space, with superposed curves of the inverse cooling rate and a Boltzmann factor corresponding to a particular H_2 line. The local density of excited molecules contains these as factors, and their product is also shown in figure 5(b) as a dashed line. This peaks at a value depending only on the upper level energy of the line, and on the cooling function. It is clear from the diagram that for modest velocities (and therefore modest post-shock temperature maxima) that the integral under the dashed curve does not depend on velocity. In this approximation, and taking s to be the T-exponent in the cooling function, the column densities of the various lines with upper level energies of T_j are proportional to $(T_j)^{-s}$. This

approximates closely to the solid curve in figure 1.

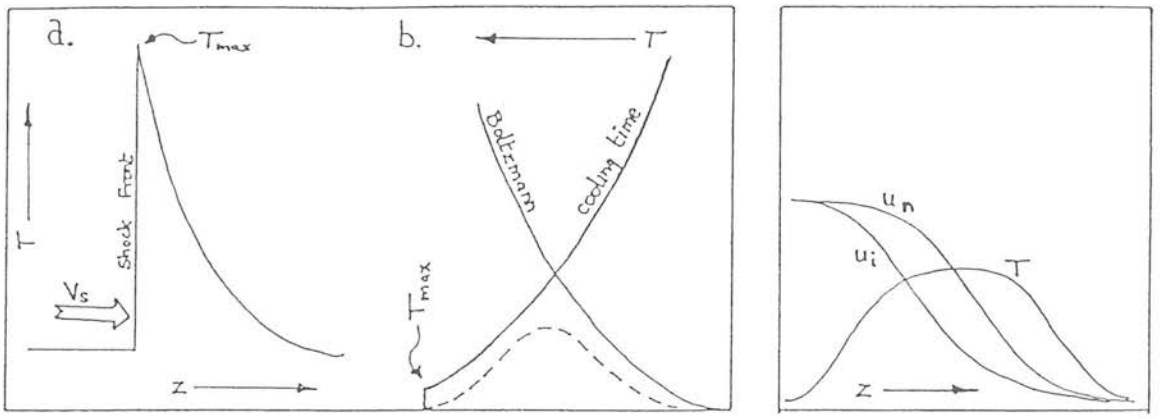


Figure 5(left). A schematic of a hydrodynamic (J-)shock. *'a'* shows the temperature profile as a function of distance through the cooling zone, while *'b'* is the same region in T-space showing the behaviour of $e^{-T_j/T}$ and of the cooling time of the shocked gas. Their product is shown as a dashed line.
Figure 6(right). shows the velocities of the ions u_i and of the neutrals u_n, and of the temperature in the frame of a C-shock.

On the other hand, a C-shock structure is quite different. The heating is by friction with the streaming ions, and is relatively slow compared with the cooling time, in contrast to the J-shock case. This means that the region where the radiation is generated reaches its highest temperature when heating balances cooling, and it behaves like a slab of gas at that maximum temperature (figure 6). Since the heating is velocity dependent, so is the maximum temperature. Thus the expected behaviour of column densities in a C-shock is that it would represent a single temperature, dependent on the shock velocity. This seems to be at variance with our observations. Because C-shock models imply a lower density in the shocked gas (there is a longer distance over which a given line is radiated), these models also show signs that the highest H_2 levels are not fully populated, and therefore the corresponding column densities fall below the Boltzmann line. On the other hand, the cooling time increases as temperature drops, and so the lowest levels find themselves in what is to all intents a cooling zone, with frictional heating no longer significant, and the pattern of column densities shows the characteristic rise at low excitation energies established also behind J-shocks. Both

these phenomena are apparent in the C-shock model in figure 1.

3. INTERPRETATION

If the evidence is accepted at face value, then there are several consequences. First of all, the very good fit of the model to the data in figure 1 implies that the cooling function is correct. This means that the cooling in these shocks is dominated by H_2 dissociation above approximately 3000 K, by H_2 line cooling, and then by CO and other molecular dipole cooling below 1000 K.

Secondly, the extinction to the Orion outflow is lower than generally accepted. The best fit value to these data is $A_K = 0.8$, which implies a visual extinction of only 9 magnitudes between it and the Trapezium stars.

Thirdly, dissociational cooling depends on the square of the density while line cooling, being saturated, depends only on the first power of density. The slope of the cooling curve at high excitation energies is therefore steeper if the pressure (and therefore the density at a given temperature) behind the shock is greater. Correspondingly, the slope of the curve in figure 1 flattens. The best fit pressure at Peak 1 in Orion (or indeed throughout the Orion flow, since the 0-0 S(13) to 1-0 O(7) ratio is constant) is high, with a value of 8×10^{10} K cm^{-3}. This provides fierce constraints on the properties of the wind driving the flow.

This seems a very neat package, but there are problems. The first is that measurements of rotationally excited CO in the Orion outflow made by several groups, and summarised by McKee (1982), indicate a post-shock density an order of magnitude lower than deduced from these models. However, the region from which this radiation was collected is much greater than the region delineated by the H_2 peak, and within the vicinity there are several sources of CO heating, with very different properties. Further work is required.

The second problem appears to be a problem for any shock model, but is especially acute for J-shock models, which completely dissociate H_2 at shock speeds above 25 km s^{-1}. The width of the profile observed (Brand *et al.* 1989) implies that in the line of sight there is a significant fraction of the excited gas travelling at up to $\pm$100 km s^{-1}, *i.e.* at over three times the dissociation speed.

However, problems or no, it is clear that the capability of high performance spectrometers on a large IR telescope to obtain accurate intensities for H_2 lines over a wide

range of intensity and wavelength is providing an excellent basis for a new and detailed look at the mechanics of energy deposition in molecular clouds by young stellar objects.

4. REFERENCES

Beckwith, S., Persson, S.E., Neugebauer, G., & Becklin, E.E. 1978, *Ap. J.* **234**, L213.

Brand, P.W.J.L., Toner, M.P., Geballe,T.R., & Webster, A.S., 1989, *MNRAS* **236**, 929.

Burton, M.G., Geballe, T.R., Brand, P.W.J.L. & Webster, A.S., 1988, *MNRAS.* **231**, 617.

Burton, M.G., 1989, *MNRAS. in press.*

Chernoff, D.F., Hollenbach, D.J., & McKee, C.F., 1982 *Ap. J.* **259**, L97.

Draine, B.T., 1980, *Ap. J.*, **241** 1021.

Draine, B.T., & Roberge, W.G., 1982, *Ap. J.* **259**, L91.

Gautier, T.N. III, Fink, U., Treffers, R.R., & Larson H.P., 1976, *Ap. J.* **207**, L29.

Hollenbach, D.J., 1981, *Symposium on the Orion Nebula to honour Henry Draper,* N.Y. Acad. Sci, New York, 242.

McKee, C.F., Storey, J.V.W., Watson, D.M., & Green S., 1982, *Ap. J.* **259**, 647.

Mon. Not. R. astr. Soc. (1990) **242**, 88–91

Velocity profiles of high-excitation molecular hydrogen lines

A. Moorhouse,[1] P. W. J. L. Brand,[1] T. R. Geballe[2] and M. G. Burton[3,4]

[1] *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*
[2] *Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720, USA*
[3] *N.A.S.A. Ames Research Center, MS 245-6, Moffett Field, CA 94035, USA*
[4] *University of California at Irvine, Irvine, CA 92717, USA*

Accepted 1989 June 6. Received 1989 May 11.

SUMMARY

Profiles of three lines of molecular hydrogen near 2.2 μm, originating from widely spaced energy levels, have been measured at a resolution of 32 km s^{-1} at Peak 1 in the Orion molecular outflow. The three lines [$1-0$ $S(1)$, $2-1$ $S(1)$ and $3-2$ $S(3)$] are found to have identical profiles. This result rules out any significant contribution to the population of the higher energy levels of molecular hydrogen at Peak 1 by fluorescence, is inconsistent with multiple C-shock models which produce higher excitation temperatures at larger shock velocities, and is generally consistent with emission from multiple J-type shocks.

1 INTRODUCTION

Observations in Orion of the $v=1-0$ $S(1)$ molecular hydrogen line (e.g. Nadeau & Geballe 1979) demonstrated that some of the shocked gas is undergoing motions in excess of 70 km s^{-1} relative to the quiescent molecular cloud. Similar highly supersonic motions of shocked H_2 are observed in other star forming regions (e.g. Doyon & Nadeau 1988; Garden *et al.* 1986), and are presumed to be associated with the interaction of a high-velocity wind from a young stellar object and a molecular cloud being impacted by the wind. However, the line emission at these high velocities defies easy explanation (e.g. Brand *et al.* 1989a).

The atomic and molecular line emission in Orion has been modelled by several authors as arising in a magnetically moderated C-type shock (Draine & Roberge 1982; Chernoff, Hollenbach & McKee 1982). However, the intensities of weak, high excitation, molecular hydrogen lines are considerably stronger than the predictions from the C-shock models, and appear consistent with a hydrodynamic or J-type shock (Brand *et al.* 1988). This conclusion presupposes that there is insignificant emission in these weak high-excitation lines from fluorescence, which would be observed as a narrow unresolved component to the line core. Conversely in C-shocks the excitation temperature increases as the shock velocity increases (Draine, Roberge & Dalgarno 1983). Comparison of the $1-0$ $S(1)$ with lines of similar excitation energy, but from longer wavelengths (Geballe *et al.* 1986; Scoville *et al.* 1982), show that slightly higher excitation lines display stronger high-velocity wings. This was interpreted by the above authors as an effect of extinction through the outflow and different locations of the high- and low-velocity gas. However, it is not clear if these differences are due, in part, to real differences in the velocity distributions of the different energy levels involved. By observing lines which are from widely different upper energy levels, but are very close in wavelength, effects of differential extinction are minimal, and any intrinsic differences in the velocity distributions can be ascertained.

2 OBSERVATIONS

The observations displayed in Fig. 1 were all made at the United Kingdom Infrared Telescope on the night of 1988 January 21. The indium antimonide photometer/low-resolution spectrometer (UKT9) was used in conjunction with an ambient temperature Fabry Perot interferometer (FP), whose pass-band has a full width at half maximum of 25 km s^{-1} in parallel light. The circular variable filter (CVF) in UKT9 acted as an order isolator, although it transmits several adjacent orders. The FP was stepped at intervals of 10 km s^{-1}, providing a fully sampled spectrum. The aperture size was 12 arcsec, degrading the resolution of the FP to 32 km s^{-1}. Standard nodding and chopping (60 arcsec east–west) were performed. All observations were of Peak 1 (RA = 5^h 32^m $46^s.2$, Dec. = $-5°$ $24'$ $02''$; Beckwith *et al.* 1978). Molecular hydrogen $1-0$ $S(1)$, $2-1$ $S(1)$ and $3-2$ $S(3)$ lines were observed. The blue wing of the $3-2$ $S(3)$ line is blended with the $4-3$ $S(5)$ line, which is observed for the first time. Further details are given in Table 1.

Absolute velocity calibration of the H_2 lines was achieved by comparing their wavelengths calculated from the energy levels of Dabrowski (1984) with nearby lines from an argon lamp. The FP was regularly monitored for velocity drift, by observing the argon lamp before and after each observation. The uncertainty in the velocity calibration is 5 km s^{-1}, the main source being drift of the FP during the observations,

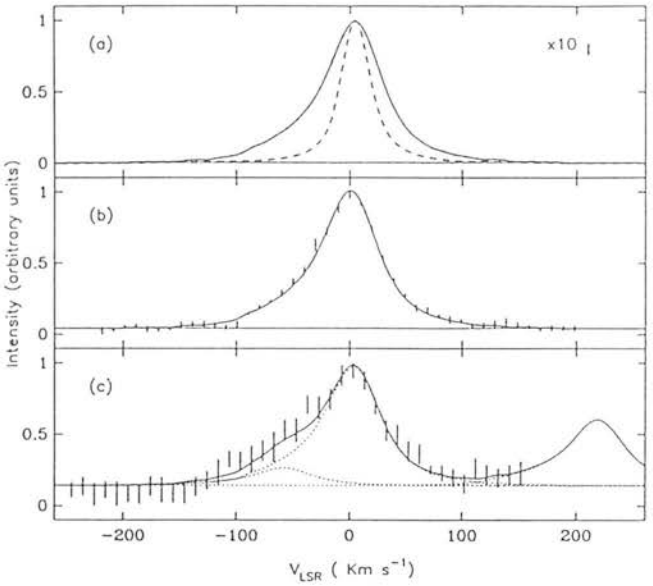


Figure 1. Profiles of molecular hydrogen lines at Peak 1 in OMC-1. (a) The $1-0$ $S(1)$ line (dots, with error bars); a cubic spline (solid curve) fitted through the data points and an Argon lamp line (dashed curve) profile. (b) The $2-1$ $S(1)$ line where the spline shown in (a) is fitted to the data (solid curve). (c) The $3-2$ $S(3)$ line, fitted by a spline model with components due to (dotted curves) the $3-2$ $S(3)$ ($v=3$ km s^{-1}), $4-3$ $S(5)$ ($v=58$ km s^{-1}), and the $1-0$ $S(0)$ ($v=215$ km s^{-1}) lines. In all of the above the dotted line represents the underlying continuum.

Table 1. Observing log.

Line	Wavelength[1] (microns)	Integration time[2] per point (sec)	Number of points
1-0 S(1)	2.1218	8	43
2-1 S(1)	2.2477	48	42
3-2 S(3)	2.2014	96	39
4-3 S(5)	2.2009	96	39

[1]Calculated from energy levels of Dabrowski (1984); [2]Total integration time includes time spent on and off source.

which thus very slightly degrades the instrumental profile for the long integrations. The measured peak velocity for each of the profiles is quoted in Table 2 as well as the relative intensities and upper energy levels of the lines. The peak of the $1-0$ $S(1)$ profile is at $+4$ km s^{-1} close to the velocity of the ambient molecular cloud ($+9$ km s^{-1}; Goldsmith, Plambeck & Chiao 1975). The velocity of the peaks of the other profiles are, within the measurement errors, identical to that of the $1-0$ $S(1)$ profile. The absolute flux scale was not determined because of uncertainties in the numbers and relative strengths of orders transmitted by the CVF.

3 RESULTS

The spectrum of the $1-0$ $S(1)$ line is shown in Fig. 1(a). The error bars are plus/minus one standard deviation. The signal-to-noise ratio is very high, so for clarity a bar ten times the length of the average error bar is plotted in the top right. The solid curve is a cubic spline fitted through the data points. The dashed line is the measured lamp profile which, since the lamp line is unresolved, may be taken to be the instrumental profile. The spline fit will be used to compare other lines to the $1-0$ $S(1)$ line. It is clear from Fig. 1 that the errors in the $1-0$ $S(1)$ data are very small. Errors in comparing the profiles of fainter lines with the spline fit will thus be due largely to the measurement errors in these lines.

Table 2. Line parameters.

Line	Energy level[1] (K)	Relative intensity	Velocity of peak[2] intensity (km s^{-1})
1-0 S(1)	6956	100	4
2-1 S(1)	12440	9	0
3-2 S(3)	19086	2	3
4-3 S(5)	25623	0.2	6[3]

[1]Calculated from energy levels of Dabrowski (1984); [2]With respect to local standard of rest; [3]Not well constrained by data, error is 10 km s^{-1}.

The $2-1$ $S(1)$ line profile is presented in Fig. 1(b). The solid line is the least squares fit to these data using the cubic spline from the $1-0$ $S(1)$ line plus a flat continuum. In this fit $\chi^2=39$ with 42 data points. It is thus apparent that there are no measurable differences between the observed $1-0$ $S(1)$ and $2-1$ $S(1)$ profiles.

The $3-2$ $S(3)$ line profile is displayed in Fig. 1(c). It is blended with the $4-3$ $S(5)$ line (2.009 μm) at -61 km s^{-1} relative to the $3-2$ $S(3)$ line. The rise in intensity beyond $+100$ km s^{-1} is the blue wing of the $1-0$ $S(0)$ line, approximately two FP orders away from the $3-2$ $S(3)$ line and partially transmitted by the CVF. On the LSR velocity scale in the diagram it would peak at 215 km s^{-1}. The solid line is again a least squares fit to the data, except that in this case a flat continuum plus three cubic splines with the correct velocity offsets were fitted to the data. The amplitude of the $4-3$ $S(5)$ line was set to 0.10 times the $3-2$ $S(3)$ line. This is the ratio calculated from a J-shock model fitted to many molecular hydrogen lines in the 2–3 μm region (Brand *et al.* 1988). The value of χ^2 is 22. There are 39 data points and three free parameters. The dashed lines show the individual contributions from the lines. Again it is evident that there are no differences between the $3-2$ $S(3)$ profile and the $1-0$ $S(1)$ profile. If the ratio of blended lines is left as a free parameter, the best fit value is 0.3, with a χ^2 of 20. It is noted that the values of χ^2 for these data are smaller than would be expected. This is probably due to the high continuum-to-line ratio (about 10), so the formal errors are dominated by continuum variations.

The $4-3$ $S(5)$ line is the highest excitation line yet detected in shocked gas. It is weak, and so it is not obvious that the data constrain its profile. This question was examined by fitting an instrumental profile instead of the $1-0$ $S(1)$ spline to this line. The result is shown in Fig. 2. It is clear that this is a worse fit (in this case the χ^2 increases to

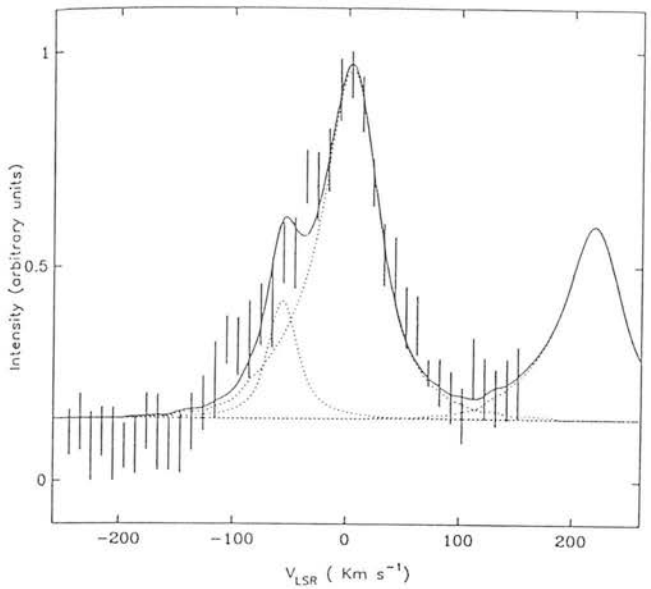


Figure 2. Same data as in Fig. 1(c), except that the 4 − 3 *S*(5) has been fitted with the instrumental profile.

26) and so we infer that the 4 − 3 *S*(5) line possesses a broad profile, which is consistent with the 1 − 0 *S*(1) profile.

In order to quantify the fraction of flux that a narrow (e.g. fluorescent) line could contribute to the broad lines, a combination of the instrumental profile and the 1 − 0 *S*(1) spline was fitted to the data, the relative amplitudes being left as free parameters. The 3-sigma upper limit to the fraction of each observed profile which could be produced by a narrow line was 0.03 for the 2 − 1 *S*(1) line, 0.20 for the 3 − 2 *S*(3) line and 0.6 for the 4 − 3 *S*(5) line.

4 DISCUSSION

Limits on fluorescence and reformation

The Orion molecular outflow lies behind the optical H II region M42. The interface layer between the two gives rise to fluorescent H_2 emission (Hayashi *et al.* 1985). Such fluorescent emission is a likely contaminant of the line emission observed toward the molecular outflow region. However, the fluorescent emission is considerably narrower than the present resolution (Burton *et al.* 1989b), so that line emission at moderate and high velocities must originate in shocked gas. The 60-arcsec chop was chosen in order to sample line emission away from the outflow, but still within the H II region, largely subtracting any foreground fluorescence. The upper limits to narrow features in the profiles represent limits to fluorescent contamination of the line emission seen toward the outflow region. From the similarity of all the lines we conclude that any fluorescent contribution from the photodissociation region is insignificant.

Theoretically the fluorescent contribution to the $v = 3$ and 4 lines is expected to be much less than the above limits. The average intensity of the 1 − 0 *S*(1) line emission observed from the photodissociation region is approximately 500 times smaller than that from Peak 1 (Tielens & Hollenbach 1985). At Peak 1 the 3 − 2 *S*(3) line is one fiftieth of the 1 − 0 *S*(1) line (Brand *et al.* 1988). Even if fluorescence was able to produce as much flux in the 3 − 2 *S*(3) line as the 1 − 0 *S*(1), the contribution from the photodissociation region to the 3 − 2 *S*(3) would be only 10 per cent of that from the shocked outflow. The fluorescent component is likely to be much less, since the fluorescent models of Black & van Dishoeck (1987) predict that at most the 3 − 2 *S*(3) is 20 per cent of the 1 − 0 *S*(1). A similar argument holds for the 4 − 3 *S*(5) line.

The above discussion refers only to the foreground fluorescence from the photodissociation region. Fluorescence in the fast-moving gas can be ruled out by the observed line ratios (Brand *et al.* 1988) and the present result of similar profiles, implying that the H_2 line ratios are independent of velocity. The ratios are totally unlike any fluorescence models (i.e. Black & van Dishoeck 1987), and show a smooth increase in excitation temperature as the energy of the upper level of the transition increases. Recently it has been pointed out that in regions of high density, collisions will thermalize the lower vibrational levels of UV-excited molecules (i.e. see review by Sternberg 1989, and references therein), however, this cannot produce the range of excitation temperatures observed and, more importantly, there would still be strong emission from vibrational levels far higher than are actually observed.

Molecular hydrogen is completely collisionally dissociated in J-type shocks with velocities greater than 25 km s^{-1} (Kwan 1977), and in magnetohydrodynamic C-type shocks at velocities greater than 45 km s^{-1} (Draine & Roberge 1982). Thus, the large width of these H_2 lines is not easy to explain. Molecular hydrogen is known to reform on the surfaces of cool dust grains (Hollenbach & Salpeter 1971; Duley & Williams 1984), at gas temperatures so low that H_2 line emission is no longer collisionally excited. It is likely that the H_2 will reform in an excited state (e.g. Duley & Williams 1986, Hollenbach & McKee 1989). However, such excited molecules will be collisionally de-excited in the dense ($\geq 10^{13}$ m^{-3}) gas behind the shock front. Therefore we do not believe that molecule reformation can account for the high-velocity, high-excitation line emission.

Shock models

It is clear from the foregoing discussion that an individual simple shock wave (either J- or C-type) cannot lead to the large range of emission velocities observed in excited H_2 lines. A high-velocity bow shock, either from a wind passing a high-density clump or a bullet propagating into a lower density medium, suffers from similar problems, and cannot in itself lead to broad profiles (Brand *et al.* 1989a). An oblique C-type shock can broaden the profile beyond what would be produced by a J-type shock, but the range is still too small to account for the observations. Finally, the possibility that the high-velocity line emission is due to scattering off fast-moving grains has been ruled out by measurements of profiles of the 1 − 0 *O*(7) line at 3.8 μm (Geballe *et al.* 1986) and by spectropolarimetry of the 1 − 0 *S*(1) line (Burton *et al.* 1988).

It therefore seems most plausible that the observed H_2 line profile along a given line-of-sight is produced by an ensemble of low-velocity shocks, some of which are occurring in molecular gas that is comoving with the molecular outflow,

as suggested by Chevalier (1980) and by Nadeau, Geballe & Neugebauer (1982). Observational evidence for this interpretation has been found by Scoville *et al.* (1982), Geballe *et al.* (1986) and, perhaps most strongly, by Geballe & Garden (in preparation), although geometric models of the emission region are not obvious (e.g. Brand *et al.* 1989a).

If indeed the line emission is due to a superposition of shocks, the similarity of the profile of lines with widely differing upper energy levels constrains the shock type. This is because, in a J-shock, the line-emitting H_2 has already been accelerated and is cooling rapidly, so the temperature distribution is independent of the shock velocity. In a C-shock, much of the line-emitting gas is in LTE at a temperature which is a strong (monotonically increasing) function of the shock speed. For example, J-shocks give significant emission only for shock velocities between 10 km s^{-1} (below this the temperature behind the shock front is not high enough to excite much H_2 before the gas cools) and 25 km s^{-1} (the dissociation limit). For shock velocities between these limits the excitation temperature of any particular line pairs is constant. The emission from J-shocks is more dependent on the pressure [through the amounts of dissociational to radiative cooling ($\propto n^2$ and n, respectively)], even over two orders of magnitude the excitation temperature derived from the $2-1\ S(1)/1-0\ S(1)$ ratio varies only from 1800 to 2200 K. The appearance of C-shocks is dependent on many parameters (i.e. shock velocity, magnetic field strength, ionization fraction and density); in a 36 km s^{-1} shock (applied to Orion by Draine, Roberge & Dalgarno 1983) the excitation temperature of the $2-1/1-0$ lines is 2700 K and drops to 800 K for 20 km s^{-1}. The line profiles of a superposition of J-shocks will be similar for all upper state energies, unless there are large variations in the shock pressure. However, for a superposition of C-shocks, higher excitation lines are very sensitive to the parameters of the shock. The present results then, when interpreted in terms of multiple shock events along the line-of-sight, clearly favour J-type over C-type shocks. Several recent and independent observational tests (Brand *et al.* 1988, 1989b; Burton *et al.* 1989a), in Orion and elsewhere, also favour J-shocks.

5 CONCLUSIONS

We have presented velocity profiles at Peak 1 of three molecular hydrogen lines arising from widely differing energy levels. All three are virtually identical in shape and peak at the same velocity. This implies that:

(i) fluorescent contributions from the photodissociation region cannot be large;

(ii) the bulk of the high-velocity emission is not due to H_2 that has reformed following a fast (dissociative) shock;

(iii) if the line profiles are due to a superposition of shocks along the line-of-sight, the shocks are more likely to be J-type than C-type.

ACKNOWLEDGMENTS

We thank the staff of UKIRT for their friendly and helpful assistance in obtaining the observations. AM was supported by an SERC studentship, and MGB held a National Research Council NASA Research Associateship at Ames Research Center during this work.

REFERENCES

Beckwith, S., Persson, S. E., Neugebauer, G. & Becklin, E. E., 1978. *Astrophys. J.*, **223**, 464.
Black, J. H. & van Dishoeck, E. F., 1987. *Astrophys. J.*, **332**, 412.
Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. R., Bird, M. & Wade, R., 1988. *Astrophys. J.*, **334**, L103.
Brand, P. W. J. L., Toner, M. P., Geballe, T. R. & Webster, A. S., 1989a. *Mon. Not. R. astr. Soc.*, **237**, 1009.
Brand, P. W. J. L., Toner, M. P., Geballe, T. R., Webster, A. S., Williams, P. M. & Burton, M. G., 1989b. *Mon. Not. R. astr. Soc.*, **236**, 929.
Burton, M. G., Brand, P. W. J. L., Geballe, T. R. & Webster, A. S., 1989a. *Mon. Not. R. astr. Soc.*, **236**, 409.
Burton, M., Geballe, T. R., Brand, P. W. J. L. & Moorhouse, A., 1989b. *Astrophys. J.*, submitted.
Burton, M. G., Hough, J. H., Axon, D. J., Hasegawa, T., Tamura, M., McCaughrean, M. J. & McLean, I. S., 1988. *Mon. Not. R. astr. Soc.*, **235**, 161.
Chernoff, D. F., Hollenbach, D. J. & McKee, C. F., 1982. *Astrophys. J.*, **259**, L97.
Chevalier, R. A., 1980. *Astrophys. J.*, **21**, L67.
Dabrowski, I., 1984. *Can. J. Phys.*, **62**, 1639.
Doyon, R. & Nadeau, D., 1988. *Astrophys. J.*, **334**, 883.
Draine, B. T., 1980. *Astrophys. J.*, **241**, 1021.
Draine, B. T. & Roberge, W. G., 1982. *Astrophys. J.*, **259**, L91.
Draine, B. T., Roberge, W. G. & Dalgarno, A., 1983. *Astrophys. J.*, **264**, 485.
Duley, W. W. & Williams, D. A., 1984. *Interstellar Chemistry*, Academic Press, London.
Duley, W. W. & Williams, D. A., 1986. *Mon. Not. R. astr. Soc.*, **223**, 177.
Garden, R., Geballe, T. R., Gatley, I. & Nadeau, D., 1986. *Mon. Not. R. astr. Soc.*, **220**, 203.
Geballe, T. R., Persson, S. E., Simon, T., Lonsdale, C. J. & McGregor, P. J., 1986. *Astrophys. J.*, **302**, 693.
Goldsmith, P. F., Plambeck, R. L. & Chiao, R. Y., 1975. *Astrophys. J.*, **196**, L39.
Hayashi, M., Hasegawa, T., Gatley, I., Garden, R. P. & Kaifu, N., 1985. *Mon. Not. R. astr. Soc.*, **215**, 31P.
Hollenbach, D. J. & McKee, C. F., 1989. *Astrophys. J.*, in press
Hollenbach, D. J. & Salpeter, E. E., 1971. *Astrophys. J.*, **163**, 155.
Kwan, J., 1977. *Astrophys. J.*, **216**, 713.
Nadeau, D. & Geballe, T. R., 1979. *Astrophys. J.*, **230**, L169.
Nadeau, D., Geballe, T. R. & Neugebauer, G., 1982. *Astrophys. J.*, **253**, 154.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G. & Ridgway, S. T., 1982. *Astrophys. J.*, **253**, 136.
Sternberg, A., 1989. *Infrared Spectroscopy in Astronomy, Proc. 22nd Eslab Symp.*, Salamanca, 1988 December 7–9th.
Tielens, A. G. G. M. & Hollenbach, D., 1985. *Astrophys. J.*, **291**, 747.

THE ASTROPHYSICAL JOURNAL, 354:242–246, 1990 May 1

INFRARED IMAGES OF IONIZED AND MOLECULAR HYDROGEN EMISSION IN S106

SAEKO S. HAYASHI,[1] TETSUO HASEGAWA,[2,3] MASUO TANAKA,[2,3] MASAHIKO HAYASHI,[4] COLIN ASPIN,[5] IAN S. MCLEAN,[5] PETER W. J. L. BRAND,[6] AND IAN GATLEY[7]

Received 1989 March 9; accepted 1989 November 6

ABSTRACT

We present high-resolution emission-line images of the central arcminute of S106 in Brγ and H_2 $v = 1$–0 $S(1)$. The images were taken at UKIRT through a Fabry-Perot filter of 130 km s^{-1} resolution using the infrared array camera IRCAM with a pixel size of 0".62. The Brγ emission shows a clear bipolar distribution which resembles that of the radio continuum emission. The central dark lane caused by the disk is wider in the Brγ image than in comparable radio continuum images. The molecular hydrogen emission is predicted to originate in the dynamically shocked regions and/or in the photodissociation regions formed at the interface between the H II region and the clumpy molecular gas. For the case of S106, the peaks of the H_2 emission are located $\sim 5''$ outside the corresponding Brγ peaks; that is, further from the central star. The assumption that this displacement corresponds to the typical depth of the H_2 emitting zone in the photodissociation region, at which $A_v \simeq 2$ mag, leads to a density estimate of $\sim 10^5$ cm^{-3}.

This high density, together with the low excitation temperature reported earlier for S106, suggests that the origin of the *thermal* component in the H_2 emission measured by Tanaka *et al.* is either *high-density fluorescence*, that is, collisional de-excitation of the UV pumped H_2 in photodissociation regions formed on the dense molecular clumps beyond the ionization fronts, or it is radiation from *slow shocks* driven by the wind from the central source, propagating into the dense clumps. In either case, a novel physical environment is required to understand the observations.

Subject headings: interstellar: molecules — nebulae: H II regions — nebulae: individual (S106) — stars: formation

I. INTRODUCTION

Vibrationally excited molecular hydrogen is often found in the vicinity of young stellar objects. Such emission was invariably interpreted in terms of thermal excitation in dynamical shocks (e.g., Beckwith *et al.* 1978) until recent excitation studies revealed the presence of fluorescent emission from radiatively excited molecular hydrogen (e.g., Hayashi *et al.* 1985; Gatley *et al.* 1987, hereafter Paper I; Hasegawa *et al.* 1987, hereafter Paper II). In practice, molecular hydrogen emission is generally found to be a mixture of thermal and fluorescent emission whose ratio varies from one source to another (Tanaka *et al.* 1989, hereafter Paper III). Theoretical calculations suggest that the radiatively excited molecular hydrogen exists in thin ($A_v \leq 2$ mag) layers in photodissociation regions (PDR) (Tielens and Hollenbach 1985), but there are, as yet, no observations of sufficiently high spatial resolution to display structural details of PDRs.

S106 is a bipolar H II region with a young massive star at the center. The distance is estimated to be 600 pc (Staude *et al.* 1982). The central star, which we will refer to as S106IR, has a bolometric luminosity of 2×10^4 $L_\odot$ (Harvey *et al.* 1982), corresponding to a zero-age main sequence star of spectral type late O or early B. The bipolar H II region is separated by a narrow gap, which is seen even in radio continuum images, suggesting that an almost edge-on disk still exists around the central star (e.g., optical/infrared: Gehrz *et al.* 1982; Hodapp and Becklin 1988; Aspin *et al.* 1989*a*; radio: Bally, Snell, and Predmore 1983; Felli *et al.* 1984). The molecular/dust disk is also reported at various scale sizes (e.g., Bally and Scoville 1982; Bieging 1984; Harvey, Lester, and Joy 1987; Mezger *et al.* 1987; Kaifu and Hayashi 1987). Two micron polarimetric observations have shown an extended reflection nebula of $\sim 3'$ radius, illuminated by S106IR (McLean *et al.* 1987), demonstrating directly that a large neutral envelope surrounds the ionized cavity delineated by the optical and radio continuum images. A ridge of molecular hydrogen emission has been found in this neutral envelope (Longmore, Robson, and Jameson 1986); its relationship with the H II region is not clearly understood.

In this paper, we present high-resolution images of H_2 $v = 1$–0 $S(1)$ and Brγ emission from S106 obtained with the infrared camera at UKIRT, which clearly show that the H_2 emission is located just outside the H II region.

II. OBSERVATIONS

The observations of Brγ ($\lambda = 2.166$ μm) and H_2 $v = 1$–0 $S(1)$ ($\lambda = 2.122$ μm) were made in 1987 August with the infrared array system IRCAM on the 3.8 m United Kingdom Infrared Telescope, Mauna Kea, Hawaii. The array used was a Santa Barbara Research Center 62 × 58 InSb device, cooled to 35 K. For detailed specifications of the instrument and the data acquisition, see McLean *et al.* (1986), McLean (1989), and Aspin *et al.* (1988). The pixel-defined spatial resolution was 0".62. A Fabry-Perot etalon manufactured by Queensgate Instruments was placed in front of the camera window and gave a spectral resolution (FWHM) of 130 km s^{-1}. A cold wide-band K filter in series with a cold narrow-band Brγ or H_2 $v = 1$–0 $S(1)$ line filter was used to specify the passband and to

[1] James Clerk Maxwell Telescope, Joint Astronomy Centre.
[2] Nobeyama Radio Observatory.
[3] Institute of Astronomy, University of Tokyo.
[4] Department of Astronomy, University of Tokyo.
[5] United Kingdom Infrared Telescope, Joint Astronomy Centre.
[6] Department of Astronomy, University of Edinburgh.
[7] National Optical Astronomy Observatories.

minimize the background radiation falling on the detector. The planetary nebula NGC 7027, which shows strong emission in both Brγ and H_2 lines, was used to tune the setting of the spectrometer. The fluxes of the current data are estimated using our NGC 7027 frames adopting the published line strengths for Brγ (Merrill, Soifer, and Russell 1975; Smith, Larson, and Fink 1981) and H_2 $v = 1$–0 $S(1)$ (Smith, Larson, and Fink 1981; Longmore, Robson, and Jameson 1986). The tracking of the telescope was adjusted by monitoring the optical image of S106 itself on a TV monitor. In addition, during the off-line data analysis, the infrared images were reregistered with respect to the strong point source S106IR.

Each image in the H_2 line is constructed from a set of four observations, taken at frequencies on and off the line and at positions on and off the source. That is, the data are both *frequency-switched* and *position-switched*. Five on-source spatial positions were measured in each line: one frame centered on S106IR and four others each offset by 15″ in both right ascension and declination. These five frames have been registered with reference to S106IR apparent in all frames before the sky subtraction and then been combined into a single image of 1 arcmin^2. The positional error caused in this registration process is mainly due to the identification of the centroid of S106IR and is not more than 1 pixel. Bright clumps and filaments are also used in supplement to assist the positioning. For the *off-line* wavelength, the Fabry-Perot filter was set 0.0027 μm shorter than the *on-line*. This difference, 0.13% in the wavelength, is sufficiently close to the maximum transmission within a narrow-band filter of 1% width. The observational sequence for Brγ was the same as that for H_2 except that only the frequency switch mode was used, and no sky frame was subtracted; the signal-to-noise ratio and the flatness are ample for the comparison with H_2 data. The total exposure times for Brγ and H_2 data were 1 minute and 7 minutes, respectively.

III. RESULTS

The Brγ image in the top left panel of Figure 1 (Plate 10) shows a clear bipolar distribution as seen in the radio continuum (e.g., Bally, Snell, and Predmore 1983; Felli *et al.* 1984). The central dark lane caused by a dust disk appears wider in the Brγ image than in the radio continuum because of the extinction in Brγ. The slight tilt of this almost edge-on disk causes more obscuration in the northern lobe. Note that the velocity field of the ionized gas is bipolar, with a receding northern lobe and an approaching southern lobe, as observed in optical line spectroscopy (Solf and Carsenty 1982). Limb brightening in the southern lobe indicates that the Brγ emission arises from a shell, not from a filled cone.

Most of the bright clumps and ridges coincide with IR sources detected at 10 and 20 μm by Gehrz *et al.* (1982). The brightest feature in the southwest corresponds to IRS 3. The spatial coincidence of the ionized clumps and the 10 and 20 μm emission is a subsidiary supporting case for the interpretation of Gehrz *et al.* (1982) that the infrared emission comes from warm dust associated with the ionization fronts.

The total flux of Brγ in S106 over the observed area is estimated to be $5(\pm 1) \times 10^{-18}$ W cm^{-2}. This uncertainty is introduced by the sky level in the current measurement and by the scatter of the measured flux of NGC 7027 in the literature. Our adopted calibration gives a flux for S106IR of about 5×10^{-19} W cm^{-2}, which agrees well with previous aperture photometry (Felli *et al.* 1984; Persson *et al.* 1984). The fluxes of the northern and southern lobes are 40% and 60% of the total Brγ emission, respectively. This is caused by the difference in the foreground extinction; typically $A_v = 12$ mag toward the northern lobe and 8 mag toward the southern lobe (Felli *et al.* 1984). The total flux of Brγ corrected for foreground extinction is $1.1(\pm 0.2) \times 10^{-17}$ W cm^{-2}.

While the Brγ image shows the familiar structure of the ionized bipolar lobes, the H_2 image resolved fine structures within the ridge known from the measurement by Longmore, Robson, and Jameson (1986). The molecular hydrogen emission, shown in the upper right panel of Figure 1, remains bipolar as a whole—separated to the north and south by the central dark lane—but the region occupied by the H_2 emission is wider in the east-west direction than the Brγ emission. The eastern and western ridges are equally bright in H_2, but the western region is brighter in Brγ. The bright H_2 emission comes from several clumps and filaments. The surface brightness of these H_2 clumps is typically $1–2 \times 10^{-4}$ ergs s^{-1} cm^{-2} sr^{-1}, which is 5–10 times brighter than the interclump region.

The most conspicuous difference between the distribution of Brγ and vibrationally excited H_2 emission is that the H_2 originates slightly farther from S106IR. Each clump is seen in both Brγ and H_2 with an almost perfect one-to-one correspondence, but the peaks of H_2 are located just outside the Brγ by 3″–7″, as shown in Figure 2 (Plate 11). These offsets are definitely larger than the positional uncertainties of the H_2 and Brγ images (see § II). We interpret this displacement as a direct measure of the density of the clumps (see § IV).

The total flux of H_2 $v = 1$–0 $S(1)$ for the observed area of S106 is estimated to be about $4(\pm 1) \times 10^{-19}$ W cm^{-2}. This uncertainty is mainly due to that in the sky level. The measured flux corresponds to half of that measured by Longmore, Robson, and Jameson (1986) which includes the diffuse emission extending beyond the boundaries of the current 1′ frame. In contrast to the Brγ flux, the H_2 flux in the northern lobe is 1.3 times brighter than that in the southern lobe in the raw data; after the extinction correction for foreground material, using the same values of A_v as for the Brγ flux, the flux of the northern lobe is 1.8 times that of the southern lobe. The total flux after this correction is $9(\pm 3) \times 10^{-19}$ W cm^{-2}. This H_2 flux imbalance between the north and south suggests some variation in the properties in the neutral envelope—probably in its density and clumpiness.

The bottom left panel of Figure 1 shows the distribution of the diffuse 2 μm continuum emission, which is more uniform than the Brγ and the H_2 emission, especially in the northern lobe. This is probably a consequence of scattering by dust. Polarization measurements with a 20″ aperture by McLean *et al.* (1987) show that an infrared reflection nebula extends out to 3′ from S106IR. In the inner region of the nebulosity, the radiation may be a mixture of the scattered light (Aspin *et al.* 1989*a*, *b*) and emission from transiently heated very small grains (e.g., Sellgren, Werner, and Dinerstein 1982).

IV. DISCUSSION

The peaks of the H_2 emission observed throughout S106 are located 3″–7″ (that is, $3–6 \times 10^{16}$ cm) farther from the central star than the corresponding Brγ peaks, a situation reminiscent of that found in the Orion bright bar, where the H_2 emission peaks $\sim$30″ (2×10^{17} cm) farther from the Trapezium stars than the ionized gas emission (Hayashi *et al.* 1985). The overall

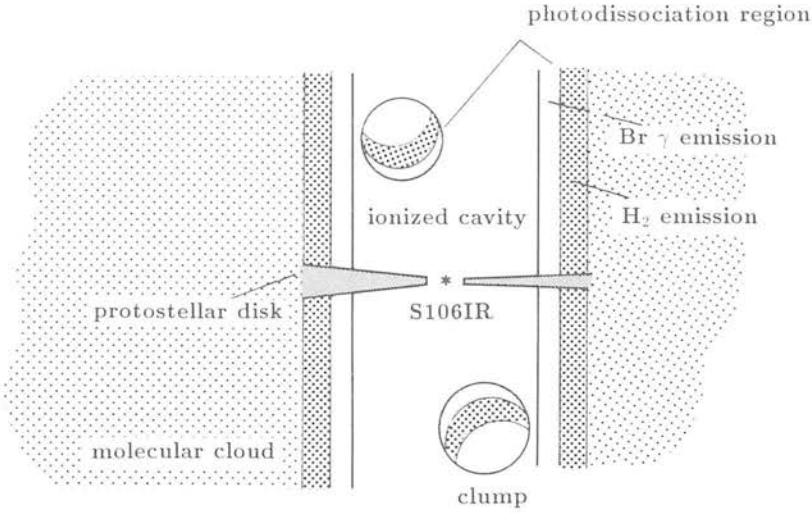


FIG. 3.—Schematic picture of the S106 region showing a cross section of the H II region, the H_2 emission region, and the ambient molecular cloud

geometry of the vibrationally excited molecular hydrogen and the ionized gas in S106 is sketched in Figure 3.

It has already been shown (Longmore, Robson, and Jameson 1986) that all the observed H_2 emission could be produced either by shock excitation or by fluorescence by UV radiation from S106IR. In the latter case, the observations can be understood in terms of a model of photodissociation regions (e.g., Tielens and Hollenbach 1985), in which the surface of a dense molecular cloud is illuminated by an intense far-ultraviolet radiation field. The nonionizing UV photons penetrate the ionization front into the neutral gas and excite H_2 molecules via the absorption in the Lyman and Werner bands (912–1108 Å). A tenth of the excited H_2 molecules dissociate, while the rest return to the electronic ground state with vibrational excitation and then cascade down by infrared fluorescence.

Theories of the fluorescent molecular hydrogen have shown that when the predissociation density is low [$n(H_2) \leq 10^4$ cm^{-3}], the level population distribution of fluorescent H_2, and thus the line intensity ratios, are insensitive to the physical parameters of the photodissociation region. That is, the line ratios in this case do not depend on the incident UV flux, the predissociation density, or the temperature (e.g., Black and van Dishoeck 1987; Takayanagi, Sakimoto, and Onda 1987; Sternberg 1988). The case of higher predissociation density has recently been considered (Hollenbach 1988; Sternberg and Dalgarno 1989). When $n(H_2) \geq 10^5$ cm^{-3}, collisional de-excitation in the photodissociation region thermalizes the population in the lower vibrational levels of molecular hydrogen at a gas temperature of ~ 1000 K, and so the infrared line ratios no longer adopt the values characteristic of fluorescence in the low-density case. Observationally, it becomes difficult to distinguish high-density radiative excitation from thermal excitation in shocks.

In Paper III, Tanaka *et al.* made sensitive measurements of six transitions of H_2 which originate from the energy levels 6471–13,890 K above the ground state. By adopting the theoretical level population distribution for *low-density* fluorescence as a template, they decomposed the observed H_2 emission into *fluorescent* and *thermal* components. For the case of S106, the results show that about 65% of the $v = 1$–0 $S(1)$ emission comes from thermally excited H_2 at an excitation temperature less than 1400 K and that the rest arises from *low-density* fluorescence (see Table 2 in Paper III).

Millimeter-wave observations of molecular lines have shown that the average density of the molecular cloud surrounding S106 H II region is $n(H_2) \simeq 10^4$ cm^{-3} (e.g., Lucas *et al.* 1978; Churchwell and Bieging 1982; Bally and Scoville 1982; Kaifu and Hayashi 1987). In this molecular cloud, however, there are clumps of higher density, as evidenced by the detection of the $J = 5$–4 and the $J = 7$–6 lines of CS (Hayashi *et al.* 1990) and these clumps are entangled with the expanding H II region.

The identification of the *fluorescent* component of S106 with the low-density [$n(H_2) \leq 10^4$ cm^{-3}] photodissociation regions surrounding the H II region seems secure. For the *thermal* component, however, there are two potential origins: collisional de-excitation of fluorescent H_2 (*high-density fluorescence*) or excitation in shocks driven by the expansion of the H II region and the outflow from the protostar. The excitation temperature for the *thermal* component derived in Paper III ($T_{ex} \leq 1400$ K) is lower than the temperature derived for a typical shock-excited source (e.g., $T_{ex} = 2000$ K in Orion-KL; Paper II, and references therein), but is in the range predicted for high-density radiative excitation. In what follows, we examine the relative importance of these two mechanisms with the new, high-spatial resolution images in hand.

First we examine the shock explanation. The reason that shock temperatures are usually measured to be ~ 2000 K is probably due to the fact that moderate-speed shocks produce a cooling zone in which the ratios of lines emitted are independent of shock speed. This mimics a 2000 K gas if the temperature is inferred from a pair of 2 μm H_2 lines. Since faster shocks are brighter, there is a strong selection effect at work. On the other hand, shocks slower than 10 km s^{-1} do not create hot enough gas for the mechanism to work. As the shock speed is reduced, the temperature derived from any pair of lines tends toward the maximum temperature behind the shock. The surface brightness also drops with shock speed, quite dramatically for the higher excitation lines.

If there are shocks in S106, they must be driven either by the pressure of the H II region, or by the wind from S106IR. It is difficult for H II gas with a sound speed of 10 km s^{-1} to produce shock speeds of more than a few km s^{-1}. Based on the optical spectroscopy, Solf and Carsenty (1982) suggested that

the expansion velocity of the toroid surrounding the protostar is 6 km s^{-1}. The molecular line at highly excited states imply that the dynamical shock, if there is any, will not exceed 10 km s^{-1} (Harris *et al.* 1987). The speed of the cloud concerned here is the expansion, or transverse velocity with respect to the bipolar cavity/outflow, and it is presumably much slower than the *outflow* velocity along the axis of the bipolar lobes (otherwise the shape of the cavity would have been different). This motion is also imposed by the wind from S106IR (e.g., Garden and Geballe 1986). In any event, these measurements imply that the shock speed is of comparable magnitude as the expansion velocity, say 6–10 km s^{-1}. The shock will have propagated at a speed ahead of the ionization front equal to the shock speed divided by the compression ratio in the shock, which will be of order 10. In these circumstances, a 10 km s^{-1} shock will have propagated 3×10^{16} cm in $\sim 10^4$ yr. This represents the separation between the Brγ and H_2 peaks. A calculation of a 7 km s^{-1} shock propagating into an initial density of 10^5 cm^{-3}, and with a cooling function including H_2 dissociation, H_2 line emission, and CO line emission gives the following result. The surface brightness is 10^{-4} ergs s^{-1} sr^{-1}, which is the observed value. The temperature inferred from the 1–0 $S(1)$ and 2–1 $S(1)$ lines of H_2 is 1200 K (Brand 1990). This calculation seems to be consistent with the observation of the H_2 clumps. However, the model is very sensitive to small changes. A decrease of 2 km s^{-1} in the shock speed reduces the surface brightness by a factor of 17. An increase by the same amount raises the apparent temperature to over 1700 K. One favorable note is that the required velocity falls in the velocity range observed in the ionized gas, as referred to above. In summary of this exercise, a slow shock can be one of the feasible explanations for the thermal H_2 emission in S106.

Now we turn to the high-density fluorescence explanation. We first estimate the density in the photodissociation regions toward the peaks of the H_2 emission assuming that the excitation is radiative. As the Brγ peaks delineate the ionization fronts, the displacement of 3″–7″ ($2.7–6.1 \times 10^{16}$ cm) between the Brγ and the adjacent H_2 peaks is a measure of the depth of the H_2 emitting zone from the ionization front. If we take $A_v \simeq 2$ mag for the extinction between the ionization front and the H_2 emitting region (Tielens and Hollenbach 1985), the total hydrogen density of the photodissociation region is estimated as $n(\text{total}) \simeq 10^5$ cm^{-3} [the relation between A_v and $N(H_2)$ is adopted from Dickman 1978]. Therefore, it is self-consistent to assume that the H_2 emission is fluorescent in these peaks, for the deduced density is high enough to thermalize the level populations. The resultant thermal gas temperature is predicted to be ~ 1000 K (Sternberg and Dalgarno 1989). A correspondingly low temperature for the thermal component in S106 has been deduced in Paper III.

The present high-resolution image resolves the H_2 peaks for the first time and gives a good measure of the surface brightness. The peak brightness of the H_2 $v = 1–0$ $S(1)$ emission clumps is $(1–2) \times 10^{-4}$ ergs s^{-1} cm^{-2} sr^{-1} with a clump-interclump contrast of 5 : 1 to 10 : 1. Sternberg and Dalgarno (1989) calculated the surface brightness of the fluorescent H_2 emission for a grid of the intensity of the incident UV radiation and the density in the photodissociation region. The UV radiation field incident upon the ionization front in S106 is estimated as $\sim 10^4$ times stronger than the general background. For this radiation field, the observed surface brightness can be reproduced, within an uncertainty caused by various possible geometries of the emitting zone, if the density is $n(\text{total}) \geq 10^5$ cm^{-3} in the clumps and $n(\text{total}) \simeq 10^4$ cm^{-3} between clumps. These densities agree with the estimates from the millimeter-wave observations.

In conclusion, the high density in the photodissociation region in S106 [$n(\text{total}) \simeq 10^5$ cm^{-3}] estimated from the present H_2 image suggests either that the *high-density fluorescence* is the major origin for the *thermal* component revealed by Paper III, or that this component is produced by *low-velocity shocks*. It is this component which comprises the emission in the bright clumps visible on the H_2 image. In the present model, the *fluorescent* component identified in Paper III arises naturally from the less dense [$n(\text{total}) \simeq 10^4$ cm^{-3}] portions of the neutral envelope around S106.

T. H., M. T., P. W. J. L. B., and I. G. are Visiting Astronomers at the United Kingdom Infrared Telescope, which is operated by the Royal Observatory Edinburgh on behalf of the Science and Engineering Research Council of the United Kingdom. The authors thank Mark Casali and Tom Geballe for their help during this experiment, and Professor Norio Kaifu for the promotion of the international collaboration which produced this paper. Suggestions from an anonymous referee were instructive and helped to improve this paper. T. H. and M. T. were supported by the Japanese Society for the Promotion of Science.

REFERENCES

Aspin, C., McLean, I. S., Rayner, J. T., and McCaughrean, M. J. 1988, *Astr. Ap.*, **197**, 242.
Aspin, C., McLean, I. S., Schwarz, H. E., and McCaughrean, M. J. 1989*a*, *Astr. Ap.*, in press.
Aspin, C., Rayner, J. T., McLean, I. S., and Hayashi, S. S. 1989*b*, *M.N.R.A.S.*, submitted.
Bally, J., and Scoville, N. Z. 1982, *Ap. J.*, **255**, 407.
Bally, J., Snell, R. L., and Predmore, R. 1983, *Ap. J.*, **272**, 154.
Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978, *Ap. J.*, **223**, 464.
Bieging, J. H. 1984, *Ap. J.*, **286**, 591.
Black, J., and van Dishoeck, E. 1987, *Ap. J.*, **322**, 412.
Brand, P. W. J. L. 1990, in preparation.
Churchwell, E., and Bieging, J. H. 1982, *Ap. J.*, **258**, 515.
Dickman, R. L. 1978, *Ap. J. Suppl.*, **37**, 407.
Felli, M., Staude, H. J., Reddmann, T., Massi, M., Eiroa, C., Hefele, H., Neckel, T., and Panagia, N. 1984, *Astr. Ap.*, **135**, 261.
Gatley, I., *et al.* 1987, *Ap. J. (Letters)*, **318**, L73 (Paper I).
Garden, R. P., and Geballe, T. R. 1986, *M.N.R.A.S.*, **220**, 611.
Gehrz, R. D., Grasdalen, G. L., Castelaz, M., Gullixson, C., Mozurkewich, D., and Hackwell, J. A. 1982, *Ap. J.*, **254**, 550.
Harris, A. I., Stutzki, J. R., Lugten, J. B., Stacey, G. J., and Jaffe, D. T. 1987, *Ap. J. (Letters)*, **322**, L49.
Harvey, P. M., Gatley, I., Thronson, H. A., Jr., and Werner, M. W. 1982, *Ap. J.*, **258**, 568.
Harvey, P. M., Lester, D. F., and Joy, M. 1987, *Ap. J. (Letters)*, **316**, L75.
Hasegawa, T., Gatley, I., Garden, R. P., Brand, P. W. J. L., Ohishi, M., Hayashi, M., and Kaifu, N. 1987, *Ap. J. (Letters)*, **318**, L77 (Paper II).
Hayashi, M., Hasegawa, T., Gatley, I., Garden, R., and Kaifu, N. 1985, *M.N.R.A.S.*, **215**, 31P.
Hayashi, S. S., *et al.* 1990, in preparation.
Hodapp, K.-W., and Becklin, E. E. 1988, cited in Gatley, I. 1988, *Science*, **242**, 1264.
Hollenbach, D. J. 1988, *Ap. Letters Comm.*, **26**, 191.
Kaifu, N., and Hayashi, S. S. 1987, in *IAU Symposium 115, Star Forming Regions*, ed. M. Peimbert and J. Jugaku (Dordrecht: Reidel), p. 369.
Longmore, A. J., Robson, E. I., and Jameson, R. F. 1986, *M.N.R.A.S.*, **221**, 589.
Lucas, R., Le Sequeren, A. M., Kazes, I., and Encrenaz, P. J. 1978, *Astr. Ap.*, **66**, 155.
McLean, I. S. 1989, in *Proc. 3d NASA Ames Workshop on IR Detectors*, ed. C. McCreight, in press.
McLean, I. S., *et al.* 1987, *M.N.R.A.S.*, **225**, 393.
McLean, I. S., Chuter, T. C., McCaughrean, M. J., and Rayner, J. T. 1986, *Proc. SPIE*, **627**, 430.
Merrill, K. M., Soifer, B. T., and Russell, R. W. 1975, *Ap. J. (Letters)*, **200**, L37.
Mezger, P. G., Chini, R., Kreysa, E., and Wink, J. 1987, *Astr. Ap.*, **182**, 127.

Persson, S. E., Geballe, T. R., McGregor, P. J., Edwards, S., and Londsdale, C. J. 1984, *Ap. J.*, **286**, 289.
Sellgren, K., Werner, M. W., and Dinerstein, H. L. 1983, *Ap. J. (Letters)*, **271**, L13.
Smith, H. A., Larson, H. P., and Fink, U. 1981, *Ap. J.*, **244**, 835.
Solf, J., and Carsenty, U. 1982, *Astr. Ap.*, **113**, 142.
Staude, H. J., Lenzen, R., Dyck, H. M., and Schmidt, G. D. 1982, *Ap. J.*, **255**, 95.
Sternberg, A. 1988, *Ap. J.*, **332**, 400.
Sternberg, A., and Dalgarno, A. 1989, *Ap. J.*, **338**, 197.
Takayanagi, K., Sakimoto, K., and Onda, K. 1987, *Ap. J. (Letters)*, **318**, L81.
Tanaka, M., Hasegawa, T., Hayashi, S. S., Brand, P. W. J. L., and Gatley, I. 1989, *Ap. J.*, **336**, 207 (Paper III).
Tielens, A. G. G. M., and Hollenbach, D. 1985, *Ap. J.*, **291**, 722.

COLIN ASPIN and SAEKO S. HAYASHI: Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720

PETER W. J. L. BRAND: Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh, EH9 3HJ, Scotland, UK

IAN GATLEY: National Optical Astronomy Observatories, 950 North Cherry Avenue, P.O. Box 26732, Tucson, AZ 85726-6732

TETSUO HASEGAWA and MASUO TANAKA: Institute of Astronomy, University of Tokyo, Osawa, Mitaka, Tokyo 181, Japan

MASAHIKO HAYASHI: Department of Astronomy, University of Tokyo, Yayoi, Bunkyo, Tokyo 113, Japan

IAN S. MCLEAN: Department of Astronomy, University of California at Los Angeles, 405 Hilgard Avenue, Los Angeles, CA 90024

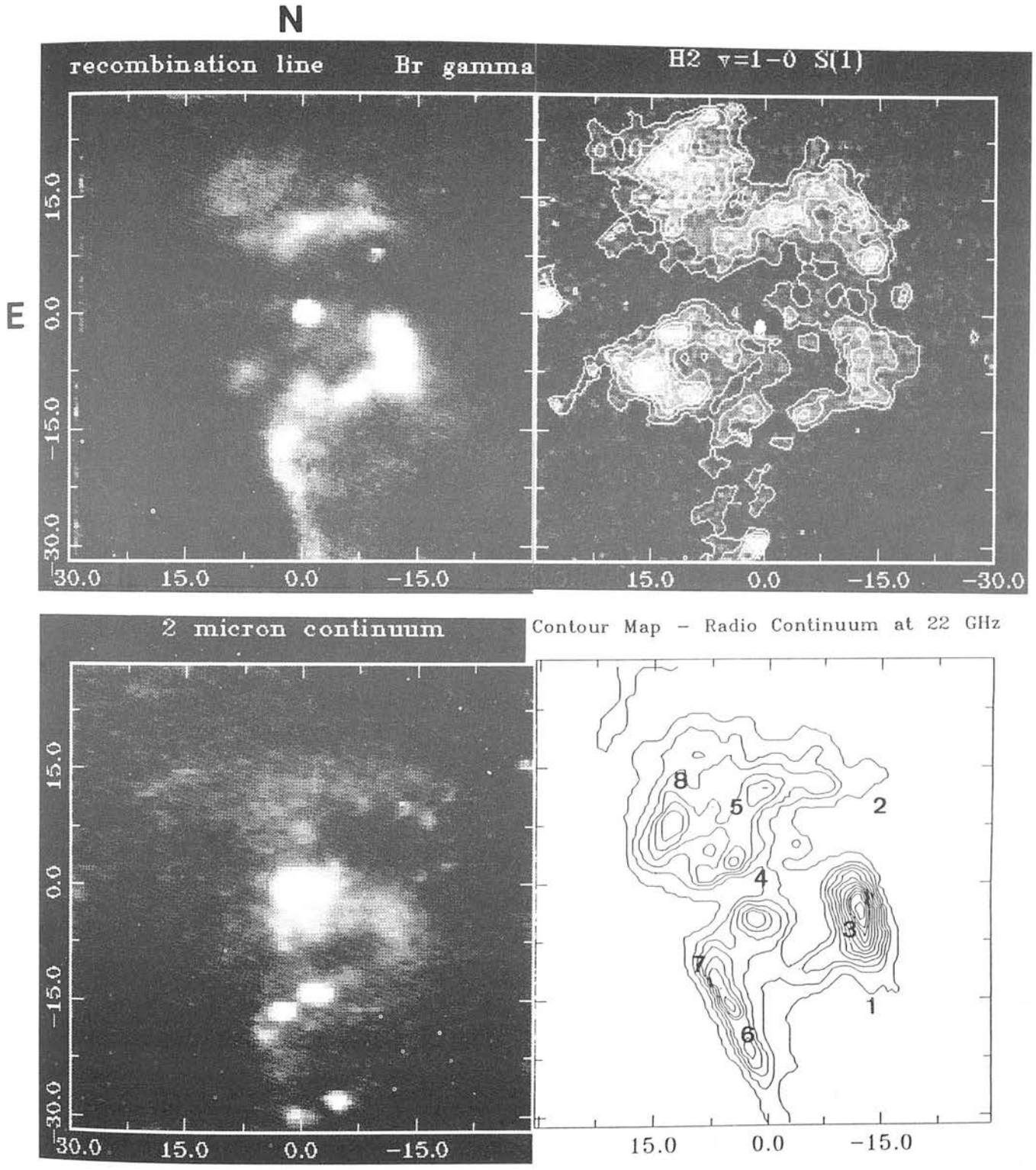


FIG. 1.—Positional offsets are in arcseconds relative to S106IR at $\alpha = 20^h25^m33^s8$, $\delta = +37°12'50''$ (epoch 1950). North is up and east is left. (*top left*) The distribution of Brγ emission ($\lambda = 2.166$ μm). The gray-scale representation ranges between 2×10^{-11} and 2×10^{-10} W cm^{-2} sr^{-1}. (*top right*) An image in the $v = 1$–0 $S(1)$ transition of H_2 emission ($\lambda = 2.122$ μm). The gray scale ranges between 6×10^{-13} and 2×10^{-11} W cm^{-2} sr^{-1}. Contours are for a running means of 3×3 pixels, with an interval of 2×10^{-12} W cm^{-2} sr^{-1} above a level of 6×10^{-12} W cm^{-2} sr^{-1}. (*Bottom left*) An image of the 2 μm continuum containing *off-line* data for top left frame. The contour interval is linear. (*Bottom right*) Radio continuum emission at 22 GHz reproduced from Felli *et al.* (1984). The contours are linear in 8% steps starting at 4% of the peak flux, corresponding to steps of 0.153 Jy per beam. The 20 μm sources detected by Gehrz *et al.* (1982) are also shown together with their numbers.

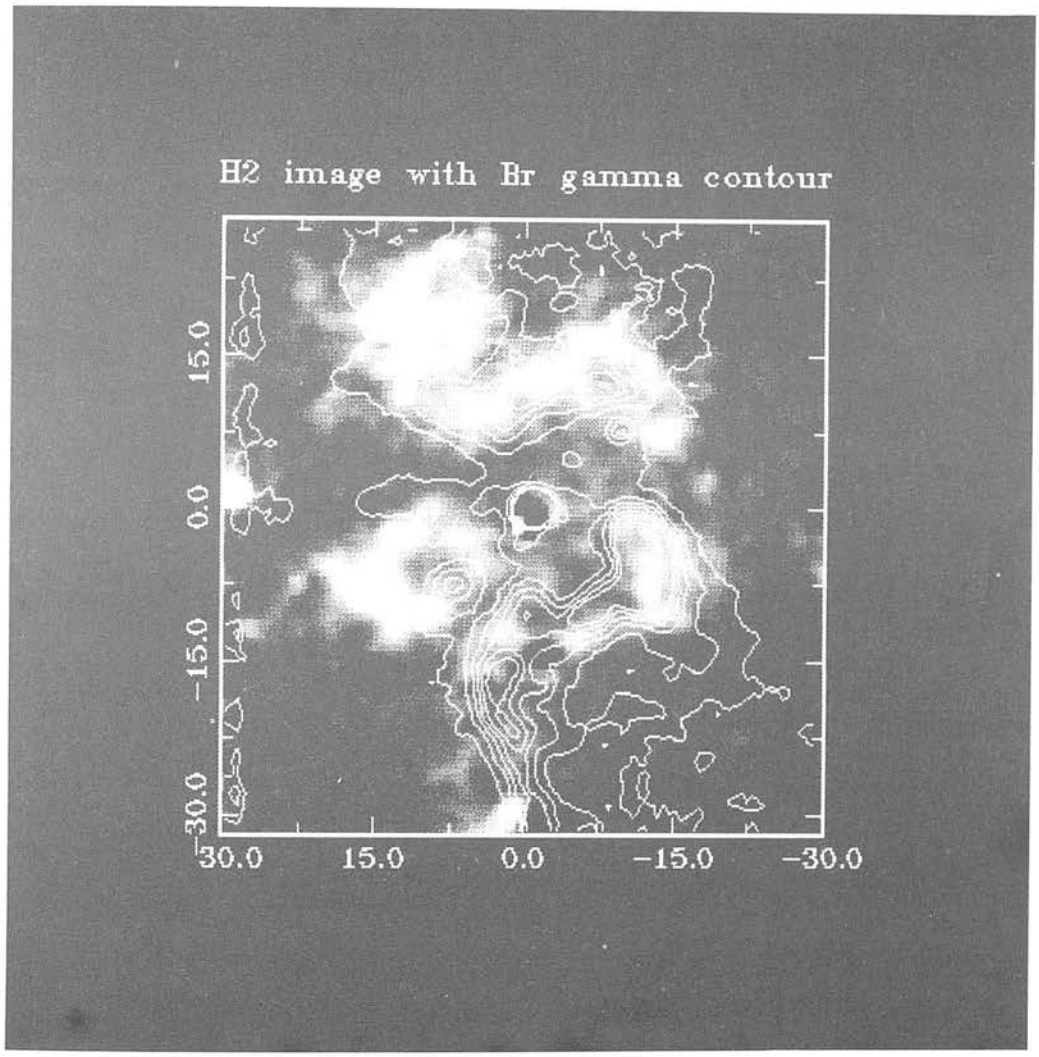


FIG. 2.—Brγ contours overlaid on the H_2 image. Both data sets are smoothed to 3 × 3 pixels. H_2 image displayed in the range 1×10^{-12} to 2×10^{-11} W cm^{-2} sr^{-1}, and Brγ contours spaced by 4×10^{-11} W cm^{-2} sr^{-1} above a level of 1×10^{-12}. Note the displacement between the peaks of the two species; the H_2 emission is shifted farther from S106IR than the adjacent ionized ridges.

Surprisingly high pressure shocks in the supernova remnant IC443

26

A. Moorhouse,[1,5] P.W.J.L. Brand,[1]
T.R. Geballe[2] and M.G. Burton[3,4]

[1] Department of Astronomy, Royal Observatory Edinburgh, Blackford Hill, Edinburgh EH9 3HJ.

[2] Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720, USA.

[3] N.A.S.A. Ames Research Center, MS 245-6, Moffett Field, CA 94035., USA.

[4] University of California at Irvine, Irvine, CA 92717, USA.

[5] Present address: Dept. of Mathematics, UMIST.

Submitted to Monthly notices

Summary

The intensities of several lines of molecular hydrogen have been measured from two regions of the supernova remnant/molecular cloud shock in IC443. The lines measured have upper state energies ranging from 7000 K to 23000 K. Their relative intensities differ in the two regions, but are consistent with those predicted from the post-shock regions of simple jump-type shocks of different pressure. The pressures so derived are far higher than the pressure in the supernova remnant itself, and a possible reason for this discrepancy is discussed.

1 Introduction

Observations of high velocity neutral atomic hydrogen in the supernova remnant (SNR) IC443 by (DeNoyer 1978) indicated that the remnant is interacting with a nearby molecular cloud (detected in mm CO lines, Cornett, Chin & Knapp 1977). Shock-excited line emission from OH and CO (DeNoyer 1979a,b) and molecular hydrogen (Treffers 1979) were subsequently discovered near to the remnant/cloud interface. The H_2 line emission has been mapped by Burton *et al.* (1988) and appears as a sinuous ridge roughly coincident with the other shocked species. Burton *et al.* argue that the SNR blast wave has punched a hole through the molecular cloud and the shock is occurring at the inner edge of this hole. This shock is driven by the high pressure of the tenuous hot material in the supernova remnant, which is at a greater pressure than the molecular cloud.

Shock excitation of the H_2 in IC443 is further supported by K band spectra observed by Burton *et al.* (1989) from several of the brighter points in their H_2 map, the v=1-0 & 2-1 transitions there having intensity ratios similar to those of other shocked sources (e.g. Garden *et al.* (1986), Lane & Bally (1986), Burton, Geballe & Brand (1989), Burton *et al.* (1989), Doyon & Nadeau (1988)). However more detailed information on the shock structure has not been available. Observations of a large number of H_2 lines in the Orion molecular outflow (Brand *et al.* 1988) suggest that jump-type shocks (J-shocks) dominate there, although most recent theoretical work has concentrated on continuous shock (C-shock) models (Draine 1980; Draine & Roberge 1982; Chernoff, Hollenbach & McKee 1982). Thus, it is of interest to try to determine the structure of shocks in other objects, such as IC443.

2 Observations

The observations presented here were all made at the United Kingdom Infrared Telescope. Several observations were made of two of the brightest positions in the molecular ring. The coordinates quoted in Table 1 are offsets from an arbitrary zero position with R.A. = $6^h\ 14^m$

43.0^s , Dec. = $22^o\ 23'\ 0.0''$ (1950).

On the nights of the 20^{th} and 21^{st} February 1989 the cooled grating spectrometer CGS2 was used at a resolving power of about 600 and an aperture size of 5″, chopping 90″ E-W. Flux calibration was achieved using the standard stars BS3314 (K=3.94) and BS1552 (K=4.07). Each spectrum was over sampled by a factor of three or four. Additional data for both positions are presented for simultaneous measurements of some of the bright 2 and 3 μm lines. Some of these data (the 3 μm lines) have already been published elsewhere (Burton *et al.* 1989), and were obtained on the nights of the 6^{th} and 7^{th} January 1987, again using CGS2 but at a lower resolving power of about 300, and a beam size 5″.

Much lower resolution spectra obtained using a circular variable filter (resolution $\approx 0.018\,\mu$m) in conjunction with the InSb photometer UKT9 were obtained between the 6^{th} and 9^{th} of January 1986. The beam size for these observations was 19″. Two spectra, one covering the K window and a second running from H to K were taken. The K window spectrum has been published by Burton *et al.* (1989) and that covering H and K by Burton (1987). Several of the line intensities obtained are inaccurate due to the difficulty of calibrating emission lines close to telluric absorption features, using the relatively low spectral resolution of the filter. Any lines which are close to such features have subsequently been excluded from the analysis. Integration times and further details are given in Table 1.

3 Results

The coadded spectra at the two locations in IC443 are shown in Figs. 1-3. The spectra at (844″W,635″N) were degraded badly by cirrus and should be regarded with some caution. The ten individual spectra at this position were scaled to the maximum observed intensity of the 1-0 S(0) line before coaddition. Weather conditions were good when the (20″W,20″S) spectra were measured.

The H_2 line intensities and derived column densities are given in Tables 2 and 3. Uncertainties are dominated by the limited accuracy of the observations. The line intensities were estimated by least squares fitting of Gaussian profiles to the data. The solid lines in Figs. 2 and 3 are the Gaussian fits to the emission lines. Column densities, N, were derived from the equation

$$I = \frac{h\nu A N e^{-\tau}}{4\pi} \qquad (1)$$

Where I is the observed specific intensity, A is the spontaneous transition probability (Turner, Kirby-Docken and Dalgarno 1977) and τ is the optical depth at frequency ν. As the line frequencies and transition probabilities are known to fractions of a percent, the random

errors in the derived column densities are determined by the accuracy of the observations. There is also a systematic percentage error of about 10 percent in the absolute column densities because of the uncertainty in the beam size.

It is apparent from Figs. 2 and 3 and from the tables that the intensities of the high excitation 3-2 lines relative to those of the stronger 2-1 and 1-0 lines are considerably different at the two locations in IC443. The ratio [3-2 S(3)]/[1-0 S(0)] is 0.13 ± 0.02 at (20″W,20″S), but 0.05 ± 0.02 at (844W,635N). Burton *et al.* (1989) found that the intensity ratio [1-0 O(7)]/[0-0 S(13)] was considerably different at these two positions; their measurements and derived column densities are also given in Tables 2 and 3.

Burton *et al.* (1988) argued that the extinction to IC443 was similar to that towards Peak 1 of the Orion molecular outflow. This was based on their observation that the relative strength of the total 1-0 Q-branch emission (observed at low spectral resolution) was similar in both sources. The extinction to the Orion molecular outflow is not well known, but the most recent estimates indicate a value of 0.8 magnitudes at K (Brand *et al.* 1988).

From the new observations presented here extinction estimates can be derived using combinations of lines from the same upper energy level. The ratios of the line intensities are then given by the molecular constants, any deviations from this being due to differential extinction at the different wavelengths involved; using equation 1 for two different lines from the same upper level, their ratio is given by

$$\frac{I_1}{I_2} = \frac{A_1\nu_1}{A_2\nu_2} e^{\tau_2 - \tau_1} \qquad (2)$$

We have observed several such line pairs. At both positions the combination of 1-0 S(1) and 1-0 O(5) can be used, and at (20″W,20″S) the S(1)/Q(3) and S(0)/Q(2). We assume an extinction law of $A_\lambda \propto \lambda^{-1.5}$, and the extinction estimates at K thus derived are shown in Table 4. There is some scatter in the results, but the best estimate for both positions is $A_k = 0.6 \pm 0.3$ magnitudes.

4 Discussion

IC443 is visible on sky survey plates as a bubble almost a degree across and contains tenuous X-ray emitting gas with a temperature of $\approx 2\times10^7$ K (Petre *et al.* 1983; Watson *et al.* 1983). The molecular cloud bissects this bubble along an axis NW-SE. An almost complete ring of 'shocked' H_2 emission has been mapped by Burton *et al.* (1988). They suggest that this ring provides strong evidence that the supernova explosion occurred close to, or possibly inside, the molecular cloud. The expansion of the remnant into the cloud has punched a hole through the cloud, and shocks on the inside edge of this hole are propagating into the cloud, visible as

the ring of H_2 emission. Other shock excited molecular line emission has also been observed in CO, OH, HCN and HCO^+ (White *et al.* 1987; De Noyer 1979a,b). High velocity atomic hydrogen emission (21cm) has also been detected by Braun & Strom (1986). All these shocked species are morphologically similar (Burton *et al.* 1988). The H_2 emission is broad (up to 100 km s^{-1}) and there are large changes in the peak velocity from different positions, with no obvious systematic trend.

There are basically two types of shocks which can excite H_2 emission. The simplest is a hydrodynamic jump (J-type) in which the high pressure region expands into the molecular cloud at a velocity greater than the local sound speed, accelerating, heating and compressing gas. The post-shock gas then cools radiatively and by dissociating molecules. As will be seen in Section 4.1 as long as the heating time in the shock front is sufficiently short compared to the cooling time of the gas, the gas can be assumed to have been heated instantaneously to a maximum temperature and the temperature distribution is then determined by the known cooling rate.

In the presence of slightly ionised, weakly magnetized gas, the ions can be driven ahead of the neutrals by magnetic forces and heat the gas gradually (while some cooling takes place) by frictional heating between the ions and neutrals. In this case there is a continuous change in parameters, and a C-shock is formed (Draine 1980, Draine, Roberge & Dalgarno 1983; Chernoff, Hollenbach & McKee 1982). The temperature structure is harder to calculate, but the resulting emission resembles that of a constant temperature slab, as the heating due to the ion-neutral drag is balanced by the cooling (e.g. Smith & Brand 1990).

The relative intensities of H_2 lines in the Orion molecular outflow suggest that the dominant shocks there are J-type (Brand *et al.* 1988). Further evidence for J-shocks in Orion comes from the constancy over the outflow of the ratio of two 3.8 μm H_2 lines (the 1-0 O(7) and 0-0 S(13)) which arise from widely separated energy levels and as such are very sensitive to density and temperature variations (Brand *et al.* 1989). This constancy presents difficulties for the C-shock models as the temperature in a C-shock is a strong function of shock velocity (e.g. Draine 1980), and is highly unlikely to be constant across the whole outflow. In J-type shocks a constant ratio simply implies that the cooling function is the same in all places.

Interestingly the S(13) to O(7) ratio has been found to differ in other shocked sources; in the four sources Burton *et al.* (1989) it ranged from 0.8 to 2.3. This variation initially seemed to pose a difficulty for the J-type shocks, but could possibly be explained as different velocity C-type shocks leading to different temperatures. However in the next section we show that in a more realistic J-type model the cooling function is pressure dependent, and this dependency can explain the observed variations.

This model is then applied to the present data, for locations where the S(13)/O(7) ratio was measured to be 0.84 and 1.9 by Burton *et al.* (1989), and is shown to provide a very good fit simultaneously to both the 2 μm and the 3.8 μm data.

4.1 Emission from J-shocks

In a J-shock the heating processes in the shock front occur on a time scale much faster than the subsequent cooling times. The temperatures, density and velocity just behind the shock front are simply given by the solution of the Rankine-Hugoniot conditions (e.g. Dyson & Williams). For a strong shock (Mach no. $\gg$ 1) the maximum temperature is

$$T_{max} = \frac{2(\gamma-1)}{(\gamma+1)^2}\frac{\mu m_h V_s^2}{k} \tag{3}$$

The molecular gas reaches this temperature and subsequently cools, through both dissociative and radiative processes. The column of molecules in a particular energy level integrated through this cooling layer is

$$N = \int_0^\infty n_j dx = \int_0^\infty n_j v dt \tag{4}$$

And as the mass flux, nv = constant = $n_o v_s$,

$$N = n_o v_s \int_0^\infty \frac{n_j}{n} dt \tag{5}$$

where n_o is the preshock density and v_s is the shock velocity. If the molecules are thermalized, (n_j/n) is given by the Boltzmann formula. Conservation of energy in the flow, neglecting bulk energy, implies

$$\frac{dT}{dt} = -\frac{\gamma-1}{\gamma}\Lambda(n,T) \tag{6}$$

where $\Lambda(n,T)$ is the total cooling rate per molecule. Thus

$$N = n_o v_s \int_0^{T_{max}} \frac{e^{-T_j/T}}{Q(T)\Lambda(n,T)} dT \tag{7}$$

where Q(T) is the partition function, which for H_2 is given approximately by

$$Q(T) = \frac{T_{rot}}{T}\frac{1}{1-e^{-T_v/T}} \tag{8}$$

where T_v = 6900 K is the vibrational partition energy and T_{rot} = 77 K. Equation 7 can be solved when the cooling processes are known. For purely radiative H_2 cooling at thermalized densities $\Lambda(n,T) \propto$ n T^s, s=4.7 . The solution of equation 7 for power law cooling is then simply

$$N \propto (T_j + T_v)^{-s} - T_j^{-s} \tag{9}$$

This simple calculation implies that the relative column densities of different energy levels will be the same in every source, regardless of shock velocity, providing that the maximum temperature is greater than that in the region where most of the emission is produced and the density is high enough to thermalize H_2.

The above calculation explains why the excitation temperatures in shocks derived from the low rotational 2-1 and 1-0 H_2 lines are always close to 2000 K. However, this result breaks down for high energy levels (> 12000 K), because a large fraction of the emission from such levels is produced in higher temperature gas where dissociative cooling is significant. The combined effects of radiative and dissociational cooling ($\propto$ n and n^2 respectively) produce a density dependence in intensity ratios involving high excitation lines which translates into a pressure dependence, as the cooling region is very close to isobaric. Increasing the pressure will allow more cooling through dissociation (higher density) and thus will cool the hot gas faster, leading to less line emission from high energy levels. Conversely a lower pressure in the shocked gas will lead to a slower cooling of the hottest gas and thus relatively stronger emission from lines from the high energy levels.

The model we have applied to the present data set is the same as used by Brand *et al.* (1988) to describe the Orion outflow. The cooling processes included are dissociation and line radiation of H_2 and rotational radiation of CO. There are only two free parameters, the absolute intensity of the shocked lines (dependent on the angle of sight through the post shock layer and beam filling factor) and the pressure. The first of these is fixed by forcing the absolute intensity to fit the bright 1-0 S(1) line, and the pressure is then varied to fit the relative intensities of the other lines to the 1-0 S(1). The results are displayed in Figs. 4 and 5. This simple J-shock is consistent with both the new 2 μm data and the 3 μm data of Burton *et al.* (1989), thus explaining the 'odd' S(13) to O(7) ratios.

The hot gas in the SNR has been presumed to be driving the shock wave into the molecular cloud. In the following discussion on the pressure in the gas we will actually be refering to the Pressure, P divided by k, the Boltzmanns constant. From X-ray and optical observations the pressure of the SNR is in the range 10^7—10^9 K cm^{-3} (Burton *et al.* 1988, and references therein). However, the pressures derived for the shocked H_2 gas are larger: for the (20″W 20″S) position the pressure needed is 2 $\times 10^{10}$ K cm^{-3} and at the (844″W,635″S) position the pressure is even greater at 6.0×10^{11} K cm^{-3} . In Section 4.2 we will examine possible mechanisms that could produce an apparent pressure enhancement. It is worth noting that even if a collection of many C-type shocks could be made to fit the data, the pressure needed

to explain the observed highly excited lines would still be high. For H_2 to be in LTE requires densities in excess of 10^6 cm^{-3}, and this critical density is even greater for highly rotationally excited molecules (i.e. 0-0 S(13) which is a transition from J=15) which require many collisions to populate the level. Significant populations of these high energy levels are only obtained at the highest temperatures (2000-4000 K in C-shocks). The pressure in the shocked gas is thus greater than 10^9 K cm^{-3}.

Far infrared observations of OI (63 μm) and CO (J=22-21, 119 μm) have recently been obtained at the (20″W,20″S) position (Burton *et al.* 1990). The authors combined these new observations with previous measurements of H_2 and Br_γ lines, and fitted a variety of shock models, with the constraint of forcing the driving pressure to be equal to that in the supernova remnant. Over 50 different model calculations were made, covering situations ranging from soft C-shocks to partially dissociative J-shocks. No single model could simultaneously explain the H_2, CO and OI lines, except a high pressure J-shock similar to the calculation we have presented here. However, on theoretical grounds they argue that such a shock is not to be expected. The shock for the given parameters should be C-type and secondly, the oxygen chemistry was artificially suppressed so as to reduce cooling from H_2O. There is no theoretical basis for either of these conditions.

At the densities we are considering the rotational energy levels of CO are in LTE, and so the rotational CO line emission can be calculated in a similar method to the H_2 emission. To perform this calculation equation 7 is multiplied by the CO fraction and the partition function used is that appropriate to CO. In order to make predictions concerning the strength of the CO emission it has to be borne in mind that the total H_2 emission in partially dissociative shocks is velocity dependent. Higher shock velocities lead to more dissociation and subsequently less H_2 emission (although the ratios remain the same). The CO emission is not as sensitive to the shock velocity, as CO rapidly reforms after dissociation. The model column density of the 1-0 S(1) line was normalized to the observed column density for the purpose of fitting to the column density ratios. This correction factor represents physical details of the shocked gas, such as the beam filling factor or the line of sight to the shock plane, and for the reasons stated above depends upon the velocity. This (unknown) factor is reflected as an uncertainty in predicting absolute CO column densities from the models fitted to the H_2 emission. For a shock appropriate to the (20″W 20″S) position (i.e. $n \times T = 2 \times 10^{10}$) and a shock velocity in the range 13-20 km s^{-1} the CO J=22-21 line intensity is $\leq 4 \times 10^{-12}$ W cm^{-2} sr^{-1}, consistent with the upper limit of 2.5×10^{-11} W cm^{-2} sr^{-1} of Burton *et al.* (1990).

4.2 Pressure in the shocked gas

As was pointed out in the last section, the pressures implied by the apparent observed J-shocks in IC443 are greater than that in the supernova bubble which is the principle driving force of the molecular cloud shock. In the rest of this section we describe a situation that could explain this apparent paradox.

The pressure in the post-shock gas is approximately equal to the pressure in the supernova bubble. Conservation of momentum across the shock front gives for a strong shock

$$P \simeq \frac{2}{\gamma+1}\mu m_h n_o v_s^2 \tag{10}$$

The gas behind the shock cools and condenses. Its final density (probably limited by magnetic field) may be several hundred times the preshock density, and then P $\approx \mu m_h n_o \, v_s^2$. This cool high density layer has a ram pressure (ρv^2) which is higher than the thermal pressure by the compression ratio.

If this compressed layer overruns a dense clump in the pre-shock medium, then a shock is reflected back into the layer, and a shock propogates into the clump. On scales large enough that the cooling zone behind the shock can be treated together with the shock as a single discontinuity, the resultant interaction is shown in Fig. (6)

Conservation of mass and momentum across the two shocks gives, assuming that the shocks are isothermal

$$\frac{P_{final}}{P_{SNR}} = \frac{\rho_{clump}}{\rho_{amb}} \frac{1}{[1 + M_o^{-1}(\rho_{clump}/\rho_{amb})^{0.5}]^2} \tag{11}$$

Where P_{final} is the final pressure, P_{SNR} is the pressure in the supernova remnant, ρ_{clump} and ρ_{amb} are the densities of the clump and ambient material, and M_o is the initial Mach number in the ambient material. The shock into the clump has a velocity of

$$V = \frac{v_s}{[1 + M_o^{-1}(\rho_{clump}/\rho_{amb})^{0.5}]} \tag{12}$$

And the corresponding velocity of the reflected shock is

$$V = v_s - \frac{v_s}{[1 + M_o^{-1}(\rho_{clump}/\rho_{amb})^{0.5}]} \tag{13}$$

For example, take the case of an ambient density of 3×10^3 cm^{-3} and a clump of 3×10^5 cm^{-3}, and assume the final compression ratio to be one hundred. Using equation 10, and a pressure of 10^9 K cm^{-3}, the initial shock velocity into the ambient material will be 41 km s^{-1}, fully dissociating the H_2. The initial shock into the dense material, driven by the pressure of the SNR, is only 3 km s^{-1}, far too small to excite H_2 emission. Thus the cooled and

compressed layer, following the shock in the dilute gas, (which because it has been compressed is moving at very close to the shock velocity) is travelling at around 38 km s^{-1} towards the dense material. At this point the densities of the clump and layer are nearly equal, it thus follows that two shocks form with the same velocity, one at 19 km s^{-1} into the clump and another slowing down the compressed layer to the same velocity. What is the pressure of this shocked gas? To answer this we use equation 10 again, the density is higher than the ambient density by one hundred times, however the shock velocity is down by a factor of $\approx$ two, so that the overall pressure increase is twenty five times the pressure driving the blast wave.

The densities do not have to be equal for this process to work. Note that faster shocks will be driven into the lighter gas, but the pressure will still be high. Further, it is not neccesarily the dense clump shock that is producing the H_2 emission we see. It is perfectly feasible, given enough time, that H_2 has reformed following the initial dissociative shock in the dilute gas (e.g. Duley & Williams 1986). Then the reflected shock into this reformed material could be the one that we see.

The shocks that we see are therefore the clumps that are in the right density range such that shocks of between 10 and 21 km s^{-1} are driven through them (if the shocks are actually C-type these limits could be changed). For the pressures we observe these limits imply densities of between 2–8$\times 10^5$ cm^{-3} dominate the emission at (20″W,20″S) wheras at the other position shocks into clumps of intial densities 5–20$\times 10^6$ cm^{-3} are emitting there.

5 Conclusions

Spectra of several H_2 emission lines from two locations in the SN remnant/molecular cloud shock of IC443 have been presented. The ratios of the observed lines obtained are shown to be consistent with a simple hydrodynamic J-type shock, in which the major coolants are radiative and dissociational cooling from H_2, and radiative cooling from CO. The line ratios from a simple J-type shock are shown to be pressure sensitive and variations in the shock pressure explain the observed variations in the ratios of the 1-0 O(7) and 0-0 S(13) lines, compared with the value of unity observed in the Orion molecular outflow. The pressure derived from the data is higher by about two orders of magnitude than the pressure derived from X-ray observations. A situation which may be able to explain this difference is the existence of density clumps colliding with the cool dense layer following a radiative shock front.

Acknowledgements

We thank the staff of UKIRT for their friendly and helpful assistance in obtaining the observations. Special thanks go to Rosemary Chapman and the CGS2 upgrade team without whom the most recent observations would have been impossible. AM was supported by an SERC studentship, and MGB held a National Research Council NASA Research Associateship at Ames Research Center during this work.

References

Brand, P.W.J.L., Moorhouse, A., Burton M.G., Geballe, T.R., Bird, M., Wade, R., 1988. *Astrophys. J.*, **334**, L103.

Brand, P.W.J.L., Toner, M.P., Geballe, T.R., Webster, A.S., Williams, P.M., & Burton, M.G., 1989. *Mon. Not. R. astr. Soc.*, **236**, 929.

Braun, R., & Strom, R.G., 1986. *Astron. Astrophys. Suppl.*, **63**, 345.

Burton, M.G., 1987. *Q. Jl. R. astr. Soc.*, **28**, 269.

Burton, M.G., Brand, P.W.J.L., Geballe, T.R. & Webster, A.S., 1989. *Mon. Not. R. astr. Soc.*, **236**, 409.

Burton, M.G., Geballe, T.R., Brand, P.W.J.L., & Webster, A.S., 1988. *Mon. Not. R. astr. Soc.*, **231**, 617.

Burton, M.G., Geballe, T.R., Brand, P.W.J.L., 1989. *Mon. Not. R. astr. Soc.*, **238**, 1513.

Burton, M.G., Hollenbach, D.J., Haas, M.R., & Erickson E.F., 1990, *Astrophys. J.*, submitted.

Chernoff D.F., Hollenbach, D.J. & McKee C.F., 1982. *Astrophys. J.*, **259**, L97.

Cornett, Chin & Knapp

Dabrowski, I., 1984. *Canadian J. Phys.*, **62**, 1639.

DeNoyer, L.K., 1978. *Mon. Not. R. astr. Soc.*, **183**, 187.

DeNoyer, L.K., 1979a. *Astrophys. J.*, **228**, L41.

DeNoyer, L.K., 1979b. *Astrophys. J.*, **232**, L165.

Doyon, R. & Nadeau, D., 1988. *Astrophys. J.*, **334**, 883.

Draine, B.T., 1980. *Astrophys. J.*, **241**, 1021.

Draine, B.T. & Roberge W.G., 1982. *Astrophys. J.*, **259**, L91.

Draine, B.T. & Roberge W.G. & Dalgarno A., 1983. *Astrophys. J.*, **264**, 485.

Duley W.W. & Williams, D.A., 1986. *Mon. Not. R. astr. Soc.*, **223**, 177.

Garden, R., Geballe, T.R., Gatley, I., & Nadeau, D., 1986. *Mon. Not. R. astr. Soc.*, 220, 203.

Lane, A.P., & Bally, J., 1986. *Astrophys. J.*, **310**, 820.

Petre, R., Canizares, C.R., Winkler, P.F., Seward, F.D., Willingale, R., Rolf, D., & Woods, N., 1983. In: *Supernova Remnants and their X-ray Emission*, p. 289, eds Danziger, J. & Gorenstein, P., Reidel, Dordrecht, Holland.

Smith, M.D., & Brand, P.W.J.L., 1990. *Mon. Not. R. astr. Soc.*, in press.

Treffers, R.R., 1979. *Astrophys. J.*, **233**, L17.

Turner, J., Kirby-Docken, K., & Dalgarno, A., 1977. *Astrophys. J. Suppl.*, **35**, 281.

Watson, M.G., Willingale, R., Pye, J.P., Rolf, D.P., Wood, N., Thomas, N., & Seward, F.D., 1983. In: *Supernova Remnants and their X-ray Emission*, p. 289, eds Danziger, J. & Gorenstein, P., Reidel, Dordrecht, Holland.

White, G.J., Rainey, R., Hayashi, S.S., & Kaifu, N., 1987. *Astron. Astrophys.*, **173**, 337.

Figure Captions

Figure 1 Spectra of the H_2 emission lines at the (20″W,20″S) position. Labels indicate the lines visible in the spectra, error bars are 1 sigma.

Figure 2 A deeper spectrum than Fig. 1., detecting the 3-2 S(3), 1-0 S(0), 2-1 S(1) and 3-2 S(2) lines (in wavelength order. The solid line is the Gaussian fit to the data points. Flux density is that in the 5″ beam.

Figure 3 Same as in Fig. 2 except for the (844″W, 635″N) data. Note that the 3-2 S(2) line is not detected.

Figure 4 Result of the J-shock fit to the column densities of the (20″W, 20″S) position. The column densities are normalized to a Boltzmann distribution of 2000 K and plotted against the energy of the upper level, so that in this plot a constant temperature is a straight line. The solid curve is the model fit for a pressure of 2×10^{10} K cm^{-3}. Crosses are the CGS2 data from Feb 1989; open circles, Feb 1987; squares the K data of Jan 1986 and the triangles the H–K data of Jan 1986. Arrows through data points are 3 sigma upper limits.

Figure 5 As Fig 4, except for the (844″W,635″N) position and a pressure of 6×10^{11} K cm^{-3}.

Figure 6 The situation that develops when a high density medium is shocked by a radiative cooling layer. The solid lines represent the density and the dotted lines the pressure. The contact discontinuity between the low and high density gas is indicated by the dashed line. Initially (1.) the shock front propogates through the dilute gas, with the cool compressed layer following behind. Finally the compressed layer (2.), moving at the shock velocity, runs into the dense gas. Two fast shocks develop, one into the dense gas and a reflected shock into the compressed layer. The pressure at this point is close to the ram pressure of the compressed layer.

Table 1: Observing log

Position	Date observed	Wavelength range μm	Total integration time (s)[1]	Calibrating star
20″W 20″S	20 Feb 89	1.7—2.5	3000	BS3314
20″W 20″S	21 Feb 89	2.2—2.25	3600	BS1552
22″W 23″S	6 Jan 87	2.1—3.8	10000	BS2821
25″W 22″S	6-8 Jan 86	K band	1500	BS3188
25″W 22″S	6-8 Jan 86	H—K bands	2000	BS3188
844″W 635″N	21 Feb 89	2.2—2.25	4800	BS3314
844″W 635″N	7 Jan 87	2.1—3.8	5000	BS2821
846″W 637″N	7 Jan 86	K band	1500	BS3188

1. Includes time spent on/off source.

Table 2: **Intensities and column densities for the (20″W 20″S) position**

Line	Wavelength[1] (microns)	Energy[1] of upper level (K)	Intensity (10^{-17} W m^{-2} beam^{-1})	Column density N_j/g_j $10^{20} m^{-2}$
February 1989				
1-0 S(7)	1.7474	12817	5.8 (0.4)	6.4 (0.4)
1-0 S(1)	2.1218	6951	28.0 (0.8)	87.8 (2.5)
2-1 S(2)	2.1535	13150	1.03 (0.09)	4.74 (0.41)
3-2 S(3)	2.2014	19084	0.81 (0.12)	1.03 (0.15)
1-0 S(0)	2.2235	6471	6.05 (0.42)	114.5 (7.9)
2-1 S(1)	2.2474	12550	2.73 (0.18)	6.32 (0.42)
3-2 S(2)	2.2872	18386	0.16 (0.08)	0.78 (0.29)
1-0 Q(1)	2.4060	6149	23.5 (2.0)	159.9 (13.6)
1-0 Q(2)	2.4135	6471	7.2 (0.4)	123.5 (12.0)
1-0 Q(3)	2.4235	6951	20.3 (0.8)	90.7 (3.6)
1-0 Q(4)	2.4375	7584	5.6 (0.3)	61.6 (3.3)
February 1987				
1-0 S(1)	2.1218	6951	31.66 (1.28)	99.3 (4.0)
1-0 O(5)	3.2350	6951	18.19 (2.40)	144.4 (19.0)
1-0 O(7)	3.8077	8365	5.38 (0.76)	63.1 (8.9)
0-0 S(13)	3.8464	17444	9.80 (0.76)	2.69 (0.21)
0-0 S(15)	3.736	21294	3.77 (1.54)	0.60 (0.25)
January 1986				
1-0 S(1)	2.1218	6951	322.9 (3.15)	70.1 (0.7)
1-0 S(0)	2.2235	6471	70.06 (3.15)	91.8 (4.1)
2-1 S(0)	2.3556	12094	5.16 (3.15)	4.9 (3.0)
2-1 S(1)	2.2474	12550	27.70 (3.15)	4.4 (0.5)
2-1 S(2)	2.1535	13150	12.58 (3.15)	4.0 (1.0)
January 1986				
1-0 S(1)	2.1218	6951	321.0 (2.5)	69.7 (0.5)
1-0 S(0)	2.2235	6471	71.5 (2.5)	93.7 (3.3)
1-0 S(6)	1.7880	11521	41.7 (5.5)	9.2 (1.2)
1-0 S(7)	1.7474	12817	72.2 (5.5)	5.5 (0.4)
1-0 S(8)	1.7147	14220	8.3 (5.5)	2.1 (1.4)
1-0 S(9)	1.6877	15721	18.6 (5.5)	2.0 (0.6)
1-0 S(10)	1.6665	17310	<5.5	<2.6[2]
1-0 S(11)	1.6504	18979	<5.5	<1.6[2]
2-1 S(0)	2.3556	12095	9.3 (2.5)	8.9 (2.4)
2-1 S(1)	2.2477	12550	26.7 (2.5)	4.3 (0.4)
2-1 S(2)	2.1535	13150	12.7 (2.5)	4.0 (0.8)

1. Calculated from energy levels of Dabrowski (1984).

2. 3 sigma upper limits

Table 3: **Intensities and column densities for (844″W 635″N) position**

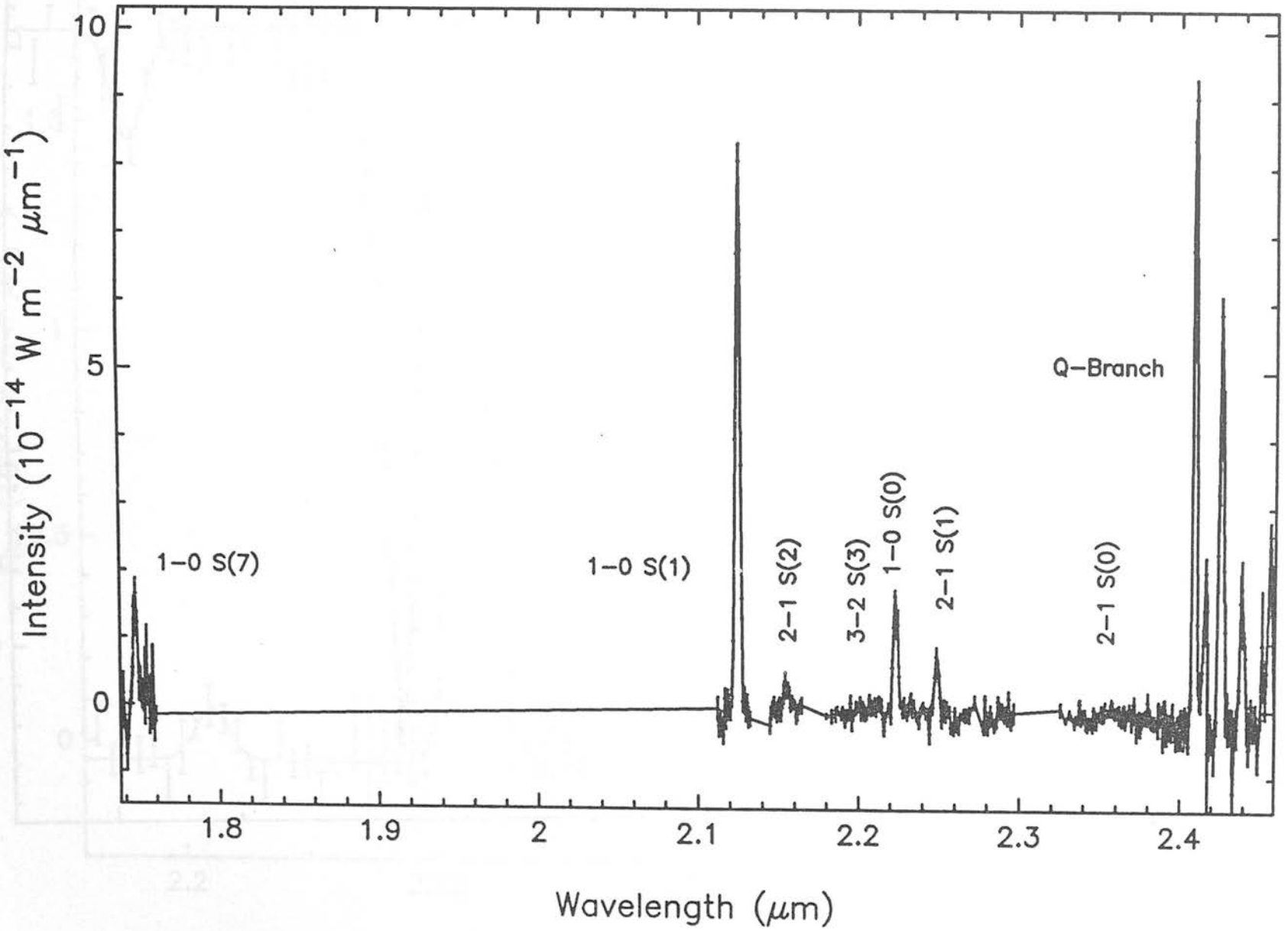
Line	Wavelength[1] (microns)	Energy[1] of upper level (K)	Intensity (10^{-17} W m^{-2} beam^{-1})	Column density N_j/g_j $10^{20}m^{-2}$
February 1989				
3-2 S(3)	2.2014	19084	0.30 (0.08)	0.38 (0.1)
1-0 S(0)	2.2235	6471	5.78 (0.3)	109.4 (5.7)
2-1 S(1)	2.2474	12550	1.58 (0.14)	3.7 (0.3)
3-2 S(2)	2.2872	18000	< 0.1[2]	<0.5
February 1987				
1-0 S(1)	2.1218	6951	34.98 (1.19)	109.7 (3.7)
1-0 O(5)	3.2350	6951	17.81 (1.66)	141.3 (13.2)
1-0 O(7)	3.8075	8365	5.98 (1.26)	70.1 (14.8)
0-0 S(13)	3.8464	17443	4.81 (1.26)	1.32 (0.35)
0-0 S(15)	3.736	21294	4.11 (1.26)	0.65 (0.20)
January 1986				
1-0 S(1)	2.1218	6951	227.7 (2.3)	49.4 (0.5)
1-0 S(0)	2.2235	6471	51.4 (2.3)	67.4 (3.1)
2-1 S(0)	2.3556	12094	3.0 (2.3)	2.9 (2.2)
2-1 S(1)	2.2474	12550	18.5 (2.3)	3.0 (0.4)
2-1 S(2)	2.1535	13150	7.0 (2.3)	2.2 (0.7)

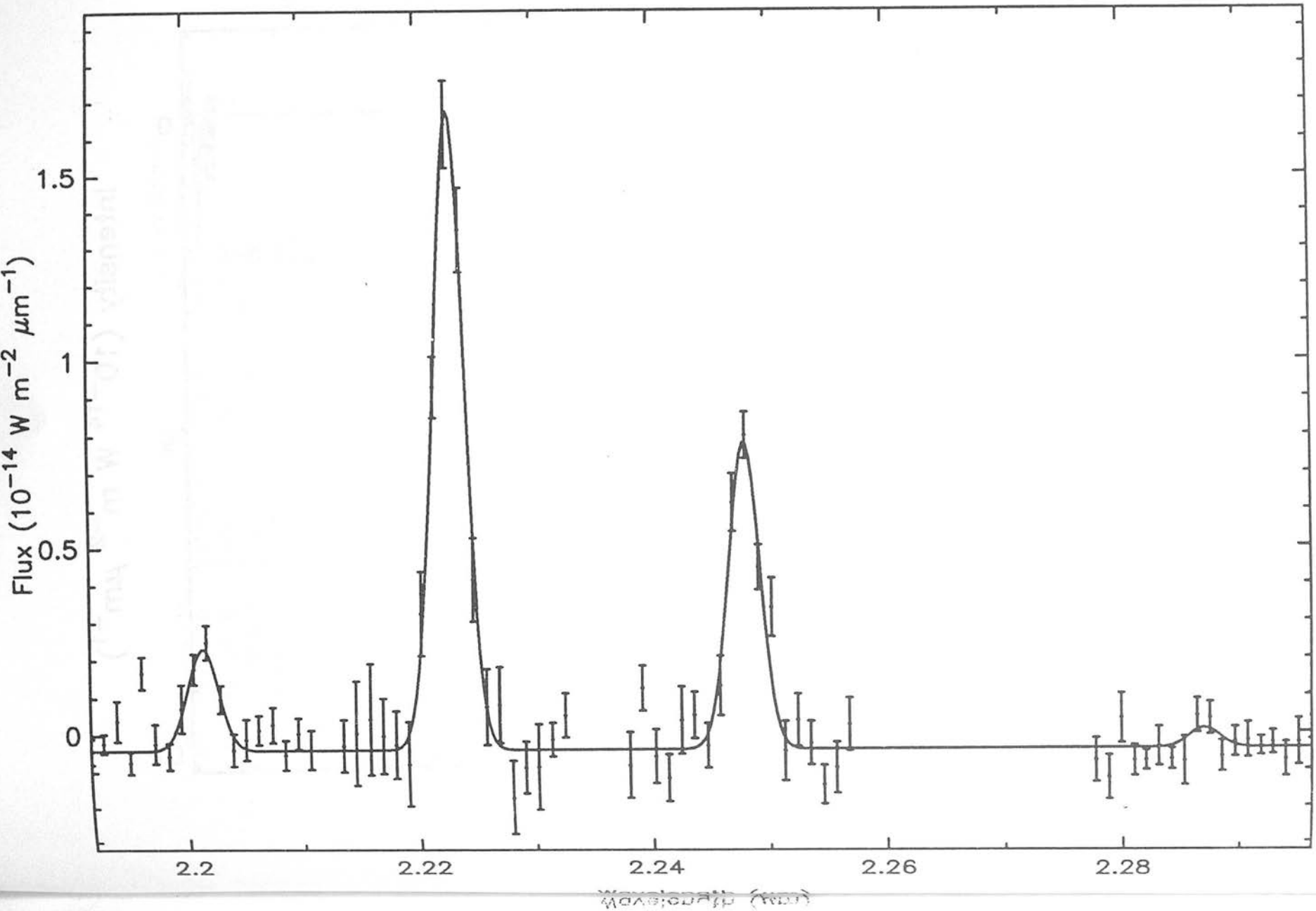
1. Calculated from energy levels of Dabrowski (1984).
2. 3 sigma upper limits

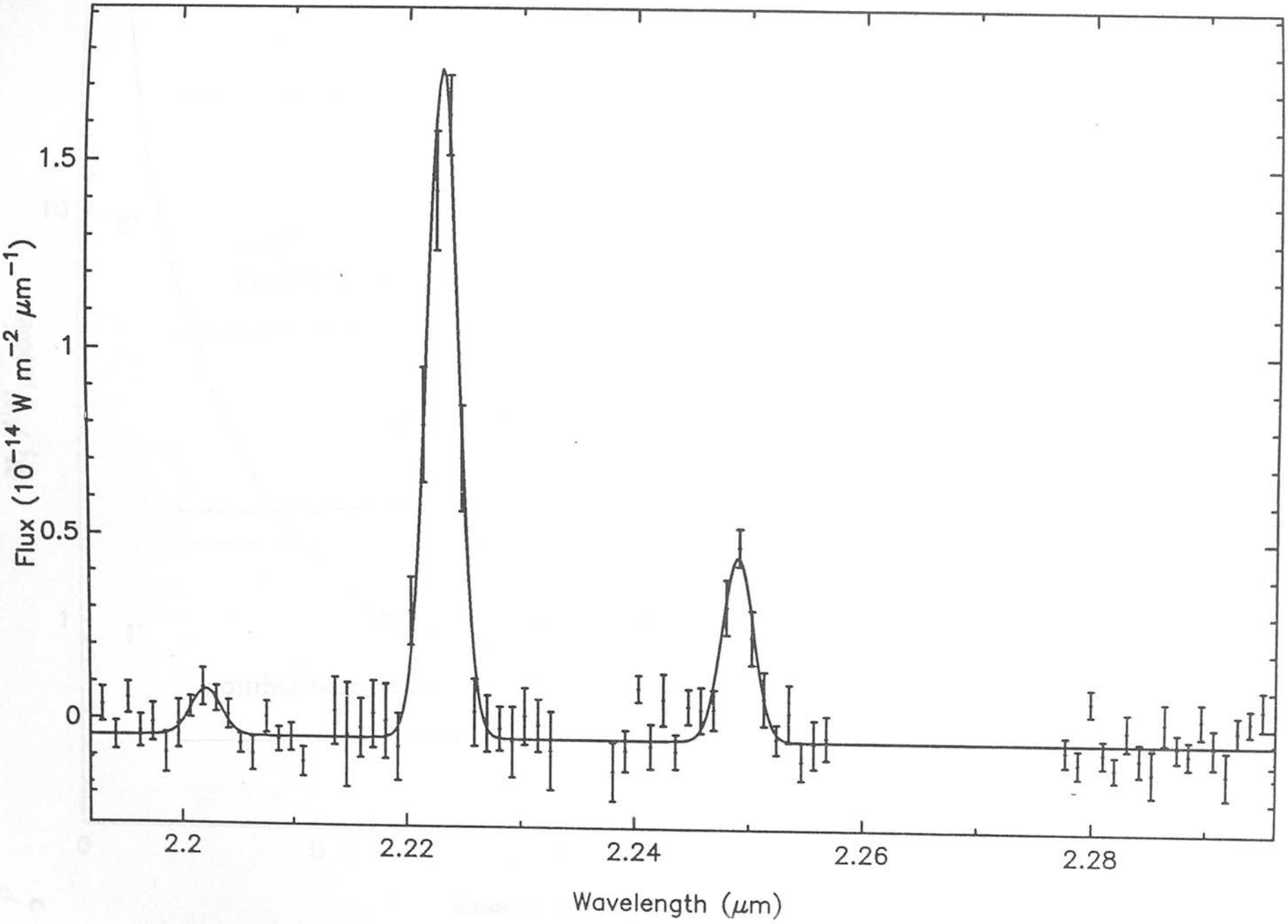
Table 4: **Extinction estimates**

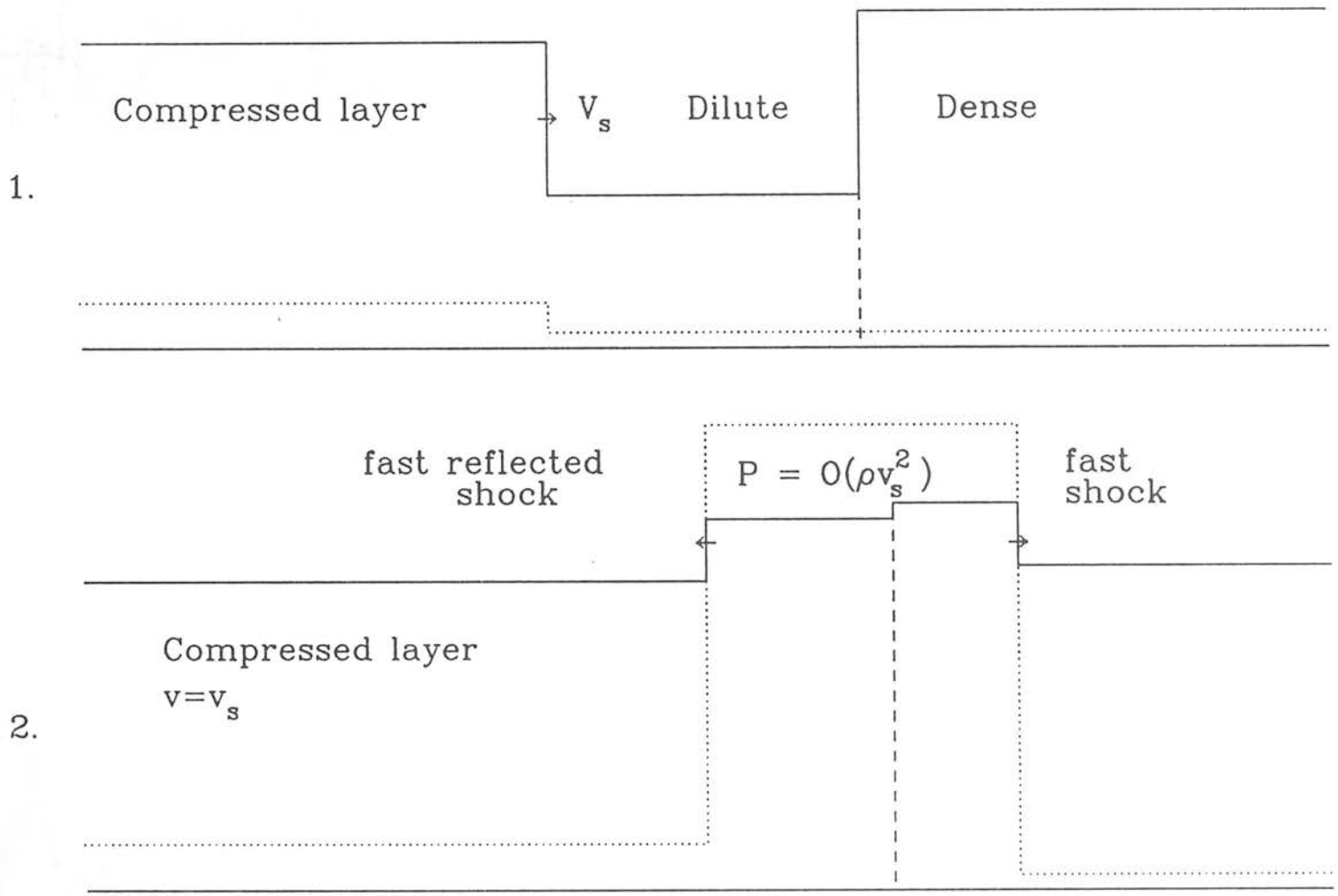
position	Line pair	ratio	A_k (magnitudes)[1]
20″W 20″S	1-0 O(5) / 1-0 S(1)	1.45 (0.20)	0.86 (0.3)
	1-0 Q(3) / 1-0 S(1)	1.03 (0.05)	0.20 (0.30)
	1-0 Q(2) / 1-0 S(0)	1.08 (0.13)	0.76 (1.1)
844″W 635″N	1-0 O(5) / 1-0 S(1)	1.29 (0.13)	0.59 (0.3)

1. Assuming A$\propto \lambda^{-1.5}$


10
5
0
Intensity (10^{-14} W m^{-2} μm^{-1})
1–0 S(7)
1–0 S(1)
2–1 S(2)
3–2 S(3)
1–0 S(0)
2–1 S(1)
2–1 S(0)
Q–Branch
1.8
1.9
2
2.1
2.2
2.3
2.4
Wavelength (μm)


Flux (10^{-14} W m^{-2} μm^{-1})
1.5
1
0.5
0
2.2
2.22
2.24
2.26
2.28


1.5
1
0.5
0
Flux (10^{-14} W m^{-2} μm^{-1})
2.2
2.22
2.24
2.26
2.28
Wavelength (μm)


1.
Compressed layer
V_s
Dilute
Dense
fast reflected shock
$P = O(\rho v_s^2)$
fast shock
2.
Compressed layer
$v=v_s$

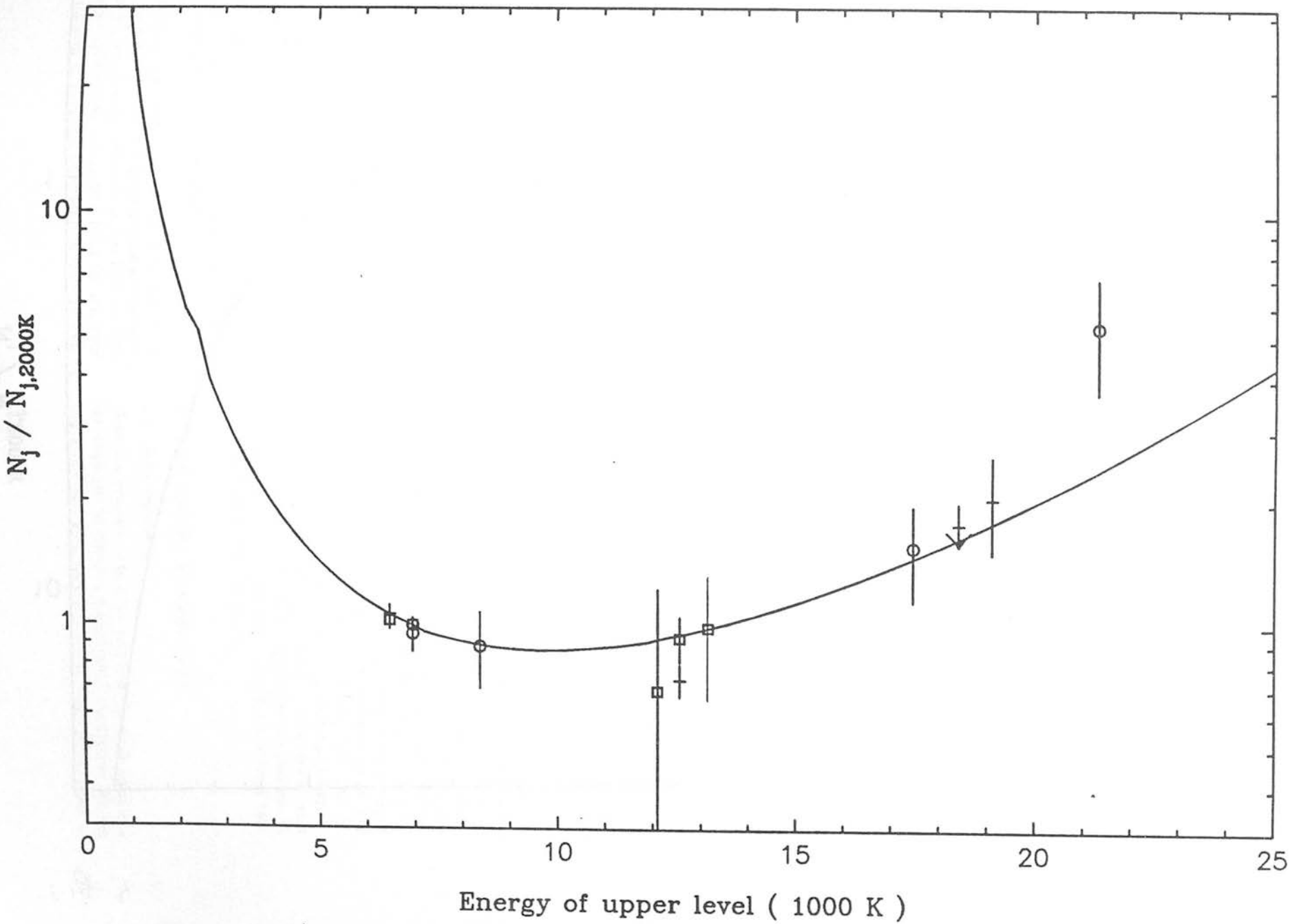
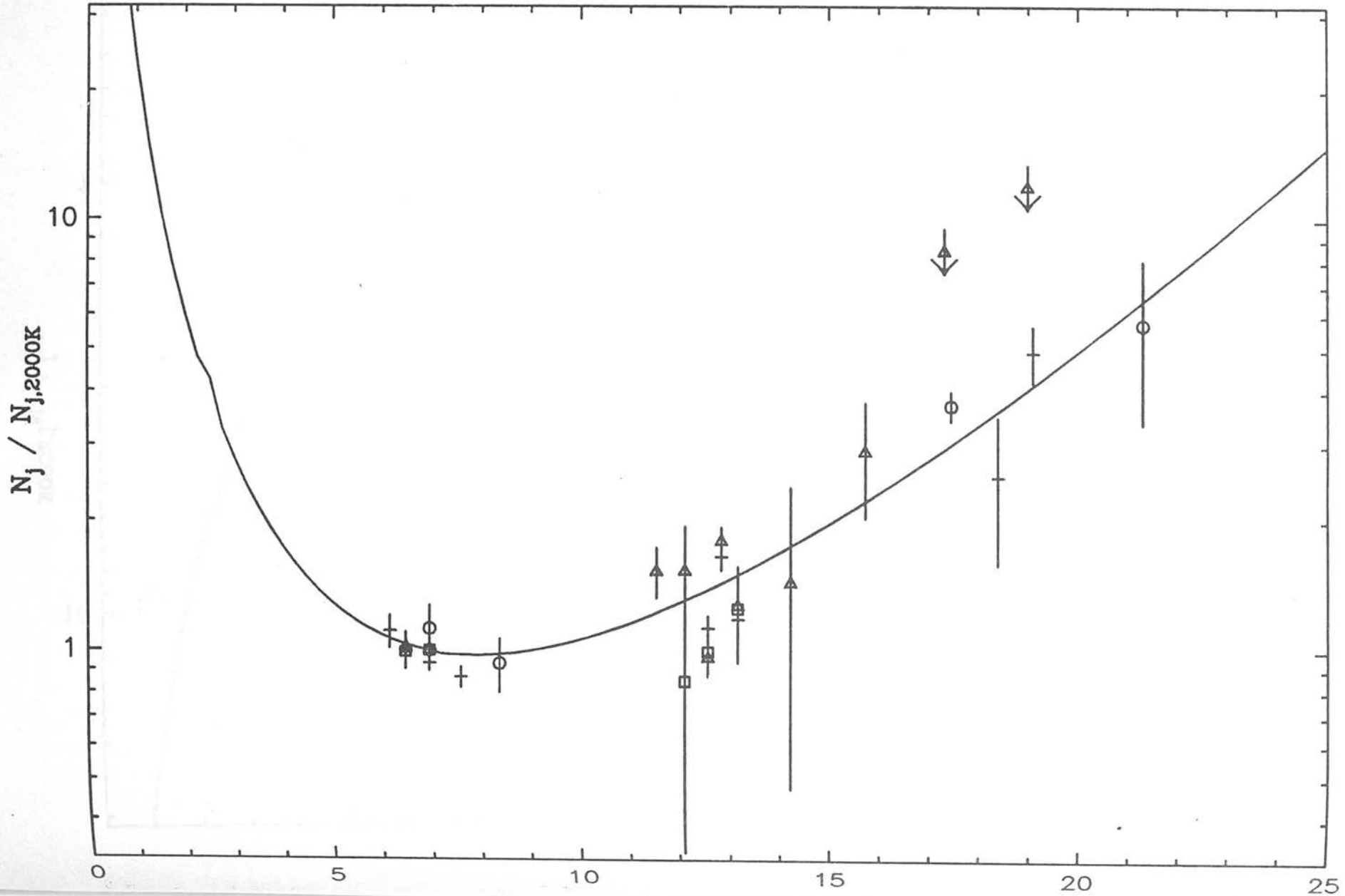

10
1
N_j / $N_{j,2000K}$
0
5
10
15
20
25
Energy of upper level (1000 K)

Fig 4



Mon. Not. R. astr. Soc. (1990) **242**, 495–504

Cool C-shocks and high-velocity flows in molecular clouds

M. D. Smith and P. W. J. L. Brand
Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ

Accepted 1989 July 21. Received 1989 June 9

SUMMARY

C-shocks can be driven through dense clouds when the neutrals and magnetic field interact weakly due to a paucity of ions. We develop a method for calculating C-shock properties with the aim of interpreting the observed high-velocity molecular hydrogen. A high Mach number approximation, corresponding to low temperatures, is employed. Under strong cooling conditions the flow is continuous even though a subsonic region may be present downstream. Analytic expressions for the maximum temperature, dissociation fraction, self-ionization level and J-shock transition are derived. This allows a wide range of velocities, densities, magnetic fields and ionization fractions to be explored and the parameter ranges for C-shocks to be displayed. The maximum speed of C-shocks in dense regions remains in the range 30–50 km s^{-1} with the exceptions of high density ($n \geqslant 10^7$ cm^{-3}) and low H_2O abundances or ionization fractions near or above 10^{-6}, for which the maximum breakdown speed is significantly smaller.

1 INTRODUCTION

High-velocity molecular outflows are associated with a very early stage of stellar evolution. Observations of molecular hydrogen lines indicate that shock-excited high-velocity gas is present in large quantities (see Burton *et al.* 1989). Finding the origin of this material will aid our understanding of the star formation process. Could this gas be accelerated locally by high-velocity shocks or is it smoothly accelerated from close to a young star before being excited by relatively weak shocks? Shocks in dense clouds can be quite distinct from those in the diffuse interstellar medium due to the extremely low levels of ionization. These C-shocks are continuous transitions in which the heating occurs over an extended region. The molecular hydrogen need not be destroyed because the gas can remain cool up to shock velocities ~ 50 km s^{-1} (e.g. Draine 1980; Draine, Roberge & Dalgarno 1983, henceforth DRD). In contrast, the usual jump shocks (J-shocks) are limited to less than 25 km s^{-1} due to the large 'instantaneous' heating (Kwan 1977). Yet J-shocks have been shown to produce line ratios which match those found in a number of observations (Brand *et al.* 1989). Here, in a programme aimed at calculating line profiles and line ratios from C-shocks in various configurations (e.g. bow shocks), we develop a method to determine which physical states lead to C-shocks.

In a C-shock the ions and magnetic field act as a separate fluid from the neutrals. The two fluids are not independent but interact via ion-neutral collisions. A few ions are thus responsible for transferring momentum from the magnetic field to the neutrals. If the flow is supersonic but sub-Alfvénic, the momentum transfer (i.e. drag) begins ahead of the neutral shock front. If the magnetic field is sufficiently strong, a continuous C-shock flow pattern arises with cooling in the extended drag region holding down the temperature and keeping the flow supersonic. The flows may also include subsonic regions and weak 'J-type' transitions (Chernoff 1987).

DRD have presented a highly detailed study of C-shocks and established the importance of a number of physical mechanisms. The extreme complexity of the numerical calculations, however, tends to inhibit the extraction of the essential relationships between observable parameters. Here, our approach is to set up the framework from which simple general predictions can be drawn. A high Mach number approximation is utilized to determine critical values for C-shock behaviour (Section 3). This implies strong cooling – a condition well satisfied over a wide parameter range. The temperature is then determined locally and can be expressed analytically. However, the dynamics and thermodynamics are not decoupled throughout a shock and we must find the range of validity of the approximation (Section 3). This clears the way for obtaining C-shock 'breakdown' conditions in dense clouds (Section 4).

2 THE MODEL FRAMEWORK

2.1 Basic equations

Here, the processes which are crucial to C-shock structures are employed to obtain a mathematical description. We are concerned with constructing a simple but reliable model to

predict the fundamental observational properties. The important physical mechanisms and approximations have been extracted mainly from DRD, Draine (1980) and Chernoff (1987).

We will ignore the momenta of ions and electrons. The total rate of change of internal energy of ions and electrons, $G_i + G_e$, is taken to be zero. This requires the ionization fraction χ to be low and their temperatures, T_i and T_e, are then determined separately from the formula $kT_{i,e} = kT_n + \frac{1}{2}(\gamma-1)\, m_n(v_n - v_i)^2$, in agreement with numerical results based on detailed physics (Draine 1980; Chernoff 1987). Here T_n and m_n are the neutral temperature and mass, v_n and v_i are the neutral and ion velocities and γ is the specific heat ratio. We assume a single neutral mass m_n and ion mass m_i. These will be taken to represent average masses.

The basic equations for a one-dimensional flow are given in Appendix A. The velocity v_n is specifically the component parallel to the direction of motion of the disturbance and B_0 is the effective magnetic field component normal to v_n in the undisturbed medium.

The following parameters will be used extensively

$$r = v_n/v_s = n_{n0}/n_n, \tag{1}$$

$$q = v_i/v_s = n_{i0}/n_i, \tag{2}$$

$$t = kT_n/(m_n v_s^2) = T_n/T_0, \tag{3}$$

where v_s is the shock speed, n_{n0} and n_{i0} are the (undisturbed) upstream densities, i.e. $\chi = n_{i0}/n_{n0}$, and ions and neutrals are conserved.

The neutral Alfvén Mach number M_A is defined through

$$M_A^2 = m_n n_{n0} v_s^2/(B_0^2/4\pi) \tag{4}$$

and we introduce the length scale $L = m_n v_s/(\mu_{ni}\langle\sigma v\rangle^{el} n_{io})$. L is generally not a shock constant (see equation 12). Equations (A1)–(A3) become

$$L\frac{d}{dx}\left(r + \frac{t}{r}\right) = \frac{q-r}{qr}, \tag{5}$$

$$L\frac{dq}{dx} = \frac{q^2(q-r)}{r} M_A^2 \tag{6}$$

and

$$L\frac{d}{dx}\left(\tfrac{1}{2}r^2 + \frac{\gamma t}{\gamma-1}\right) = \frac{q-r}{r} - \frac{\psi t}{qr} \tag{7}$$

representing momentum conservation for neutrals and ions and energy balance, where (exactly as defined by Chernoff)

$$\psi = \frac{m_n n_n \langle\sigma v\rangle^{I}}{\mu_{ni} n_i \langle\sigma v\rangle^{el}}. \tag{8}$$

Here $\langle\sigma v\rangle^{el}$ and $\langle\sigma v\rangle^{I}$ are the rate coefficients for elastic scattering and radiative cooling, respectively. Combining equations (6)–(8) and eliminating x yields

$$\frac{dr}{dq} = (\gamma-1)\frac{(r-q)[\gamma r/(\gamma-1) - q] - \psi t}{rq^3(r-q)(1-\gamma t/r^2) M_A^2} \tag{9}$$

and (6) and (7), with $r = 1$ at $q = 1$, integrates to give

$$r + \frac{t}{r} + \frac{1}{2M_A^2 q^2} = 1 + \frac{1}{2M_A^2}. \tag{10}$$

Chernoff (1987) has described and displayed solutions to these two equations. The two extreme cases of energy-conserving flow and zero-temperature flow were explained in detail. For other examples the parameter ψ was taken to be of the form $\psi_0(4q)^{\alpha}(4r)^{\beta}(t/t_s)^{\gamma^*}/t_s$, where $t_s = 3m_n v_s^2/(16kT_0)$ and $\gamma = 5/3$. The nature of the shock is then determined by M_A and ψ_0. The numerical solutions cover a wide range in t but were restricted to $M_A < 20$ and $\psi_0 < 3$. Here we only investigate solutions for $t \ll 1$ with no implicit restrictions on M_A and ψ_0. This approximation will be shown to be very accurate for many situations with t typically less than 0.02.

2.2 Drag and cooling

The momentum transfer rate coefficient generally contains several terms, involving the drift velocity, the relative importance of grains and the fraction of ions (Draine 1980). Two distinct cases for ion–neutral coupling will be considered here: $\langle\sigma v\rangle^{el} = \text{constant} = \sigma_0 v_0$ (model RCO) and $\langle\sigma v\rangle^{el} = \sigma_0(v_n - v_i) = \sigma_0 v_s(r-q)$ (model RC1), where $\sigma_0 = 10^{-15}\,\text{cm}^2$ is assumed and $v_0 \sim 24\,\text{km s}^{-1}$. Together these are written as

$$\langle\sigma v\rangle^{el} = \sigma_0 v_0 (v_s/v_0)^e (r-q)^e \tag{11}$$

with $e = 0, 1$. For dust, equation (11) will still be used with $e = 1$ and the effect of dust drag will be measured by adapting χ (see Draine 1980, equation 40). That is: $\chi = \chi_i + \chi_f$ where χ_i is the ion contribution and χ_f is the ion fraction which would produce the same force on the neutrals as the dust. In DRD $\chi_f \sim 2\times10^{-7}$.

An appropriate shock length scale is now definable

$$L_0 = L(r-q)^e = (m_n/\mu_{ni})(v_s/v_0)^{1-e}/(n_{i0}\sigma_0). \tag{12}$$

Cooling of the time-independent separable form $\Gamma = n_n^{1+f}\Lambda(T_n)$ is assumed. Under LTE, above the appropriate critical density, n_{cr}, of the dominating cooling molecule, $f = 0$. Below n_{cr} we take $f = 1$.

The flow pattern is governed by ψ_0 and M_A where the ratio of drag to cooling time-scales takes the form

$$\psi = \frac{q\Lambda(tT_0)}{t^f(r-q)^e\Lambda(T_0)}\,\psi_0. \tag{13}$$

ψ_0 is a shock constant:

$$\psi_0 = \frac{n_{n0}^f \Lambda(T_0)}{n_{i0}\mu_{ni} v_s^2 (v_s/v_0)^e \sigma_0 v_0}. \tag{14}$$

In Section 4 we show that ψ_0/M_A is of order $(10^{-3}/\chi)(v_s/10\,\text{km s}^{-1})^{2.6}$ for typical dense cloud conditions with pure H_2 cooling and is of order $(10^{-5}/\chi)(v_s/10\,\text{km s}^{-1})^{-1}$ for H_2O cooling. Thus under very low ionization conditions ($\chi < 10^{-6}$) we have $\psi_0 \gg M_A$. To further orient the reader it should be noted that $M_A = v_s/(v_A\cos\theta)$ where v_A, the Alfvén speed, is typically a few km s^{-1} where measured (Myers & Goodman 1988) and $\cos\theta$ is the field component normal to the shock.

3 THE HIGH MACH NUMBER MODEL

3.1 The temperature

The temperature tT_s, neutral velocity nv_s and ion velocity qv_s are related through two equations (9 and 10). Spatial distributions can be found through either equation (5) or (6). The flow is determined by M_A and ψ_0, where ψ_0 is a ratio of ion-neutral drag time to cooling time. C-shocks will tend to occur for ψ_0 large. Then the drag is not efficient and so the neutrals are not tied to the magnetic field. C-shocks also tend to occur for M_A small. M_A^2 is the ratio of shock thrust to magnetic pressure. For low ionization fractions the ion-magnetosonic Mach number $\sim(n_{i0}/n_{n0})^{1/2}M_A=\lambda^{1/2}M_A$ is extremely small. Hence the subsonic field-ion fluid will act to cushion the neutral thrust through the drag force. The cushioning is most effective for low thrust and high field. A C-shock then occurs rather than a J-shock.

The initial and final states are both supersonic since there is no heating to stabilize the downstream temperature (the final velocity is not zero because of magnetic field support). A continuous supersonic transition of the neutral parameters will occur if the denominator in equation (9) remains positive, i.e. there is no singular point. The sufficient requirement is that the Mach number $M=r/(\gamma t)^{1/2}$ remains greater than unity throughout. Shock solutions are possible if there is a point at which $M=1$ and $(r-q)[\gamma r/(\gamma-1)-q]=\psi t$, i.e. the numerator of equation (9) is also zero (Chernoff 1987).

Here we take $M\gg 1$. Then equation (10) yields

$$r=1+(1-1/q^2)/(2M_A^2);\ dr/dq=1/(M_A^2q^3). \tag{15}$$

Equation (9) then becomes

$$\psi t=(r-q)^2, \tag{16}$$

independent of γ as expected when the thermodynamics are decoupled from the dynamics. Equation (16) can be derived from physical considerations, simply equating the drag heating rate to the cooling rate. Combining it with equation (13) gives

$$\frac{\Lambda(tT_0)}{\Lambda(T_0)}=\frac{r^f(r-q)^{2+e}}{q\psi_0} \tag{17}$$

which is equivalent to t^α when $\Lambda(T)\propto T^\alpha$, i.e. a power-law cooling function. The maximum temperature $T_{max}=t_{max}T_0$ can be found from equations (16) and (17) provided that $\Lambda(T)$ is a monotonically increasing function of T. We actually find the maximum value of Λ from which T_{max} can be found for any cooling function. The maximum occurs at

$$r_m=\frac{2\alpha_c}{2\alpha_c+1/2}\left[1-\frac{(e+2)(2\alpha_c-1)(2\alpha_c+1/2)^{1/2}}{8M_A\alpha_c^2}+0\left(\frac{1}{M_A^2}\right)\right], \tag{18}$$

$$q_m=\frac{(2\alpha_c+1/2)^{1/2}}{M_A}\left[1-\frac{(e+2)(2\alpha_c-1)(2\alpha_c+1/2)^{1/2}}{2M_A\alpha_c}+0\left(\frac{1}{M_A^2}\right)\right]. \tag{19}$$

for M_A large. The critical index α_c is defined as

$$\alpha_c=1+(e+f)/2. \tag{20}$$

and lies in the range $1<\alpha_c<2$. At this stage the ion velocity has greatly decreased while the neutral velocity change is generally small. Hence the density is still close to the pre-shock value. Substitution into equation (17) yields

$$\frac{\Lambda(T_{max})}{\Lambda(T_0)}=\left[\frac{M_A}{\psi_0}\right]\frac{(2\alpha_c)^{2\alpha_c}}{(2\alpha_c+1/2)^{2\alpha_c+1/2}}\times\left[1-\frac{(e+2)(2\alpha_c+1/2)^{3/2}}{2M_A\alpha_c}+0(1/M_A^2)\right]. \tag{21}$$

It is important to retain the $1/M_A$ terms for the purpose of estimating the degree of molecular dissociation, which is highly sensitive to the maximum temperature.

3.2 The downstream analysis

Here we find the conditions for J-shocks to occur within strongly cooled flows and determine the properties of warm subsonic sections. The above formulae were derived under the condition $M^2=r^2/(\gamma t)\gg 1$. In fact M can then be shown to possess a minimum value, M_*, for a given C-shock. If M_* is large the above analysis is accurate. As long as M remains large across the region which dominates the emission features the analysis is still usable. However, if M_* does approach unity a discontinuous transition to a subsonic flow might ensue. We first determine the conditions under which a flow, assumed to be a C-shock, stays supersonic throughout.

In Appendix B we show that the minimum M_* occurs after the maximum temperature but while the gas is still hot for $\alpha<\alpha_c$. For $\alpha>\alpha_c$ the minimum is located far downstream in the cooled gas. In comparison to the maximum temperature we find for $\alpha<\alpha_c$ and $\alpha>\alpha_c$, respectively

$$\left[\frac{t_*}{t_{max}}\right]^\alpha\sim\left[\frac{\alpha_c-\alpha}{\alpha}\right]^{2\alpha_c}\left[\frac{\alpha_c+1/4}{\alpha_c+1/4-\alpha}\right]^{2\alpha_c+1/2} \tag{22}$$

$$\left[\frac{t_*}{t_{max}}\right]^\alpha\sim\frac{2^{1/2}(2\alpha_c+1/2)^{2\alpha_c+1/2}(e+2)^{e+2}(2\alpha-f)^f}{[2^{5/2}\alpha_c(\alpha-\alpha_c]^{2\alpha_c}M_A^{2\alpha_c}} \tag{23}$$

to zero order. These two results are consistent with Figs 1 and 2. In Fig. 2 it is seen that M becomes small well after the major changes in the shock structure have taken place.

The minimum Mach number can be enumerated on substituting for M_A and ψ_0. The results are displayed in Fig. 3. Notice that the high Mach number approximation is very accurate for the whole shock layer provided $\alpha<2$. For a steep cooling function and typical dense cloud parameters $M_*\sim 0(1)$. The approximation is then invalid in the cooled downstream gas. Here we argue that the approximate solutions are still accurate for large α at least for the high-temperature region of a flow. Provided t_{max}/t_* is large and since the neutrals have already been greatly decelerated, the discontinuity is a minor disturbance to the downstream boundary condition. Moreover, steep cooling functions are generally not realisable over a wide temperature range as has been assumed. It is clear that the H_2 cooling cannot dominate once the temperature has dropped somewhat below 1000 K,

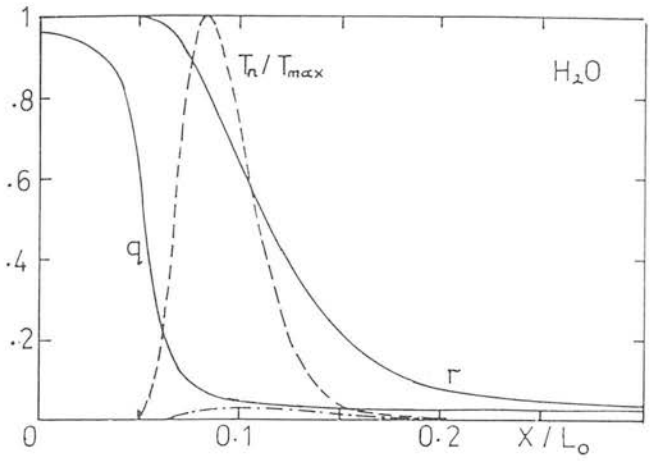


Figure 1. Structure of a 40-km s^{-1} shock in a dense molecular cloud with $n_6 = 1$, $B_{-3} = 1$, $\chi_8 = 10$, $\xi = 3.5$ and $f = 1$ and a flat cooling function ($\alpha = 1.5$). Here $T_{max} = 4800$ K and $L_0 = m_n/(\mu_n n_n \sigma_0) \sim 10^{16}$ cm. Note that $1/M^2$ (the dot-dash line) peaks near T_{max} and remains extremely small throughout.

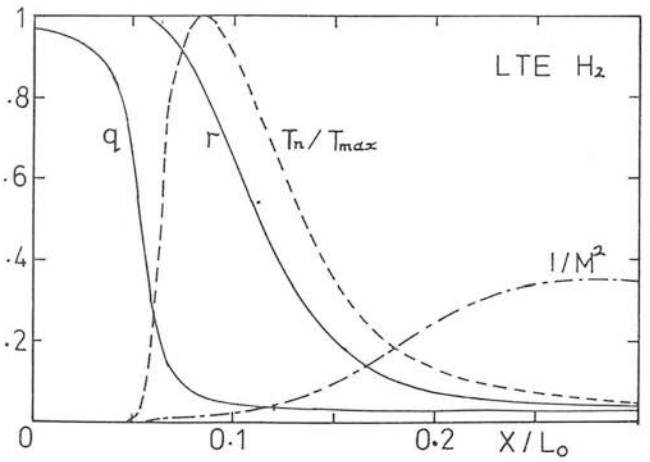


Figure 2. As Fig. 1 but with a steep cooling function ($\alpha = 3.3$), $f = 0$, and $T_{max} = 3230$ K. The pre-shock medium has $q = r = 1$.

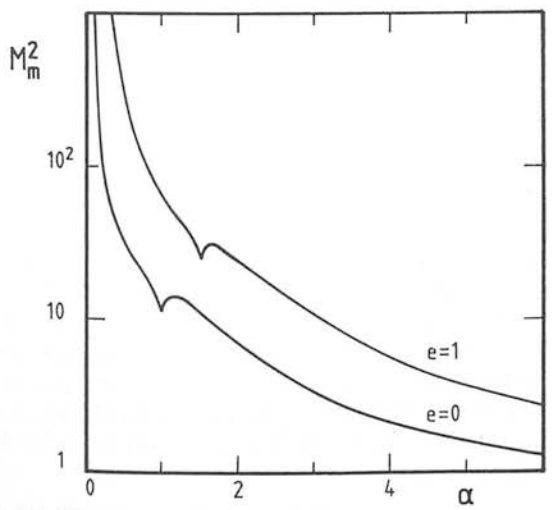


Figure 3. The minimum Mach number in the C-shock region for a range of α, from equations (B3) and (B5), taking n_6, B_{-3}, v_{10}, χ_8 and $\xi = 1$, $f = 0$.

whereas the minimum Mach number is reached at a temperature $T_* = 150\,[\chi_8 B^2_{-3} v_{10}/\varepsilon]^{0.21}$ for $\alpha = 4.7$ (see Section 4). CO cooling will cut in before this temperature has been reached in the cooling downstream gas, reducing t and increasing M. CO cooling takes the form $\Lambda = 2.25 \times 10^{-15} T_3^2\,[x(\mathrm{CO})/(5 \times 10^{-5})]$ for $n > n_{cr} = 2 \times 10^6 T^{3/4}$ cm^{-3} (McKee *et al.* 1982). This latter condition will generally be easily satisfied in the region far downstream. CO cooling dominates (on comparison to the cooling rates in Section 4.1) for $T < 870\,[x(\mathrm{CO})/(5 \times 10^{-5})]^{1/2.7}$. Employing this as the cooling function yields $M_*^2 = 21\,[B_{-3}/(\gamma n_6 v_{10} \chi 8^{1/2})]$. Moreover, H_2O cooling will further improve the accuracy even with a minute fraction of the oxygen abundance since it (a) has a flatter cooling function and (b) has a strong density dependence making it effective in the compressed downstream region.

These arguments, however, are not general, applying only to the indicated values. A weak or oblique magnetic field, for example, may lead to a significant shock in the flow which will alter the emission characteristics. Furthermore, while C-shocks are limited by dissociation and ionization above 25 km s^{-1}, another limit is provided by the transition to J-shocks at lower velocities. Such low-velocity shocks must be included in our analysis since we wish to apply our results to bow shocks in which the field will be highly oblique over a wide region.

3.3 The transition to subsonic flow

We look at the downstream region by 'magnifying' this region in the following manner. If the flow does not remain supersonic throughout there must exist a transition region with critical points C and C'. The flow becomes subsonic by passing through C or near C either continuously or via a shock (Chernoff 1987). Further downstream it again becomes supersonic by passing through C'. As demonstrated below these points generally lie near the final state at

$$(q_f, r_f) = \left(\frac{1}{M_A\sqrt{2}} + \frac{1}{4M_A^2} + 0(1/M_A^3),\ \frac{1}{M_A\sqrt{2}} + \frac{1}{4M_A^2} + 0(1/M_A^3) \right) \tag{24}$$

for our regime of interest ($\psi_0 \gg M_A$). Taking

$$(q, r) = \left(\frac{1}{M_A\sqrt{2}} + \frac{1+Q}{4M_A^2},\ \frac{1+R}{M_A\sqrt{2}} \right) \tag{25}$$

gives

$$t = (1+R)(Q-R)/(2M_A^2) + 0(1/M_A^3). \tag{26}$$

Writing $dr/dq = (4M_A/\sqrt{2})\,dR/dQ$, equation (9) becomes

$$\frac{dR}{dQ} = (\gamma - 1)\,\frac{-\psi_0[(1+R)(Q-R)/(2M_A^2)]^{\alpha}(M_A\sqrt{2}/R)^{[\gamma(1+R)/(\gamma-1)-1]+c+f}}{1+R-\gamma(Q-R)}. \tag{27}$$

From the denominator we obtain the $M = 1$ line

$$Q = (1+\gamma)\,R/\gamma + 1/\gamma \tag{28}$$

and, from the numerator, the heating = cooling ($h=c$) line

$$Q = R + \frac{1}{1+R}\left(\frac{(\gamma R+1)\,R^{1+e+f}(M_A\sqrt{2})^{2\alpha-1-e-f}}{(\gamma-1)\,\psi_0}\right)^{1/\alpha}. \qquad (29)$$

The points of intersection of equations (28) and (29) are simply C and C' (see Chernoff 1987 for a full discussion). A shock will occur if C is a spiral point; a smooth transition to subsonic flow occurs if C is a node.

The value of dR/dQ at C is found by taking the limit as C is approached. This yields

$$\left.\frac{dR}{dQ}\right|_c = \frac{\gamma}{\gamma+1} + \frac{\gamma-1}{\gamma+1}\left(\frac{\gamma}{\gamma-1}R+\frac{1}{\gamma-1}\right)\left(\frac{(\gamma-1)\alpha}{1+R}+\frac{1+e+f}{R}\right). \qquad (30)$$

At C the $h=c$ line is steeper than the $M=1$ line. Consequently for a flow trajectory, with r and q both monotonically decreasing, to intersect C we require $dR/dQ|_c$ to exceed the gradient of the $h=c$ line (otherwise C is a spiral point). C is given (equation 29) by the parameters α and $K=\psi_0/(M_A\sqrt{2})^{1+2(\alpha-\alpha_c)}$, the same factors which determine M_*. For K large the high M approximation has already shown that the flow remains supersonic. For K small M_* is small and subsonic regions are inevitable.

For R large the intersection occurs at low values of K: $K\sim\gamma^{\alpha+1}R^{-2(\alpha-\alpha_c)}/(\gamma-1)$. The magnification process holds for Q, $R\ll M_A$. Thus for the intersection to be analysed in this manner we require $K\gg\gamma^{\alpha}[\gamma/(\gamma-1)]\,M_A^{-2(\alpha-\alpha_c)}$ which is rewritten simply as $\psi_0\gg M_A$. This is the strong cooling regime which is easily satisfied for all low-ionization states (see the rough values in Section 2.2). For $\psi_0\sim M_A$ the intersection is then at values of r of order unity and strong J-shocks are then possible, in agreement with the numerical results of Chernoff (1987).

For $R\gg 1$ the gradient of the $h=c$ line at C is

$$\left.\frac{dR}{dQ}\right|_{h=c} \sim \frac{1}{1+(2\alpha_c-\alpha)/(\gamma\alpha)}. \qquad (31)$$

After some manipulation we find

$$(dR/dQ)_c/(dR/dQ)_{h=c} = [2\alpha_c+(\gamma-1)\alpha]^2/[(\gamma+1)\alpha]. \qquad (32)$$

The first result is that for $\gamma>1$ there exist continuous transitions for a range of high α values. Secondly, solving equation (32), shocks arise only if $(\gamma+1)/(\gamma-1)>8\alpha_c$. Hence shocks can arise only for $\gamma<9/7$ (since $\alpha_c\geqslant 1$). Since $\gamma>9/7$ in molecular clouds we conclude that the transition to subsonic flow is always continuous for $\psi_0\gg M_A$. The explanation is that the ion–neutral interaction is sufficient to broaden the transition to subsonic flow. Over this transition, however, R can change quite steeply and the temperature can still rise considerably. A second peak in the temperature profile will still result and will now be discussed.

3.4 The downstream temperature peak

For α near 3 the temperature reaches a maximum value $\sim(1+R)^2/(2M_A^2)$ very close to C. For general α this is not clear but from equation (26) we know the temperature in the subsonic region to be $t_2<Q(1+R)/(2M_A^2)$. For K small this yields $t_2(1+\gamma)(\gamma-1)^2/[\gamma^{2\alpha+3}K^{1/(\gamma-\gamma_c)}]$. It can be shown to take the value $t_2/t_{max}<0.029\,(M_A/\psi_0)^{\alpha_c/[\alpha(\alpha-\alpha_c)]}$ relative to the first maximum provided $\gamma\geqslant 7/5$ and $1<\alpha_c<2$. Thus the temperature of the second peak is low. Nevertheless, the column of gas through this peak could be large since the density is relatively high at this stage. It can be shown using equation (6) that the second peak is of length $\Delta x\sim L_0\Delta Q(1+R)(M_A\sqrt{2})^e/(2M_A^2R^{1+e})$. Given the rise in density, the column N is then $\sim 0[n_{n0}L_0M^{e-1}(1+R)^{1-e}]$ which takes the highest value for $e=0$: $N\sim 0\{n_{n0}L_0[\psi_0/(M_A\sqrt{2})]^{1/2(\alpha-\alpha_c)}\}$. Since, for example, $\alpha\sim 4.7$ for low-temperature H_2 cooling, moderately high columns, 10–100 $n_{n0}L_0$, are possible. Although this is a relatively large column compared to that near $T_{max}[\sim 0(n_{n0}L_0/M_A)$, see Section 4], the emission depends exponentially on the temperature. For this reason the secondary peak is, observationally, always completely negligible.

3.5 The upstream transition to decoupled flow

The thermodynamics are not decoupled at the beginning of the shock front (i.e. upstream). Here, large thermal pressure gradients act with the drag to slow down the neutrals. From Figs 1 and 2 it is evident that the thermal pressure ($\propto t/r$) rises rapidly from the zero pressure of the initial cold state. If this phase is extended it could negate our method of determining the maximum temperature as well as having a profound effect on the H_2 line profiles. The importance of this phase is found by putting $\varepsilon=1/M^2\ll 1$ into equation (10). This yields

$$\frac{(r^2-t)}{r^2}\frac{dr}{dq}+\frac{1}{r}\frac{dt}{dq}\sim\frac{1}{M_A^2q^3} \qquad (33)$$

and, to first order in equation (9),

$$\psi t\sim(r-q)^2+(r-q)\left[\frac{t}{r}+\frac{M_A^2q^3}{(\gamma-1)}\frac{dt}{dq}\right]. \qquad (34)$$

By using equation (17) for t, from which we derive the expression

$$\frac{dt}{dr}=\frac{t}{\alpha}\left(\frac{f}{r}+\frac{2+e}{r-q}-\frac{2+e+r-q}{r-q}\frac{dq}{dr}\right), \qquad (35)$$

the terms in equation (34) can be compared. The solution $\psi t=(r-q)^2$ is then valid provided $t\ll r(r-q)$ and $M_A^2q^3\,dt/dq\ll r-q$. Using equation (35) the latter condition becomes $t\ll(r-q)^2/(M_A^2q^3)$. If, as at t_{max}, $r\sim 0(1)$ and $q\sim 0(1/M_A)$ then both conditions are clearly consistent with the high Mach number approximation. Using equation (10) the second condition becomes $\psi_0\gg r^f(r-q)^{2+e-2\alpha}(M_A^2q^3)/q$ which breaks down as $q\to 1$ for $\alpha>1+e/2$. For this reason the approximation can be used to determine T_{max} but is not reliable in the early stage of the shock. Numerical results for planar shocks (Smith & Brand 1989) bear out this conclusion.

4 C-SHOCK BREAKDOWN

4.1 Cooling functions and C-shock temperatures

Quantities are expressed in the following units

$$T_n = 1000 T_3 \text{ K}, \; n_n = 10^6 n_6 \text{ cm}^{-3}, \; B = 10^{-3} B_{-3} \text{ G}, \tag{36}$$

$$\chi = 10^{-8} \chi_{-8} \quad \text{and} \quad v_s = 10 v_{10} \text{ km s}^{-1}.$$

Note that n_n will be taken as $0.5 n_H$ for a fully molecular gas where n_H is the atomic hydrogen density.

H_2 cooling is approximated by

$$\Gamma = 3.3 \times 10^{-15} \xi T_3^{\alpha} n_6^{1+f} \text{ erg s}^{-1} \text{ cm}^{-3}, \tag{37}$$

with $f = 0$, $\alpha = 4.7$ and $\xi = 1$ appropriate for LTE cooling at densities $n_n > n_{cr} \sim 3 \times 10^6$ cm^{-3} and $T_n < 2000$ K (Burton 1986), and $\alpha = 3.3$ and $\xi = 1$ being a close fit in the range $1000 \text{ K} < T_n < 5500$ K. Non-LTE cooling is approximated from Hollenbach & McKee (1979) with $\xi = 0.36$, $\alpha = 3.3$ and $f = 1$ in the latter temperature range, valid for $n_6 < 2.8$.

The molecule H_2O may dominate the cooling if the oxygen is sufficiently abundant, efficiently converted (DRD) and not photodissociated. From Hollenbach & McKee (1979)and McKee *et al.* (1982) we extract

$$\Gamma(H_2O) \sim 11.6 \times 10^{-15} [x(H_2O)/(4.25 \times 10^{-4})] \times T_3^{3/2} n_6^2 \text{ erg s}^{-1} \text{ cm}^{-3}, \tag{38}$$

where x is the abundance relative to n_H (e.g. DRD) and, since $n_{cr} \sim 10^{11}$ cm^{-3} for H_2O, $f = 1$. Thus $\alpha = 1.5$ and taking $x = 4.25 \times 10^{-4}$ gives $\xi = 3.5$ for the indicated abundance.

The two flow parameters become

$$\psi_0 = \frac{41 \, \xi (2.4)^e (24.2)^{\alpha} v_{10}^{2\alpha - 2 - e}}{(\mu_{ni}/m_n) \chi_8 n_6^{1-f}}, \tag{39}$$

$$M_A = \frac{6.5 \, n_6^{1/2} v_{10}}{B_{-3}}. \tag{40}$$

The maximum temperature is now given (using equations 21 and 39 with $\mu_{ni} = m_n$, i.e. the ions are, on average, comparatively heavy)

$$\left[\frac{T_{max}}{10^3 \text{ K}}\right]^{\alpha} = \left[1 - \frac{(e+2)(2\alpha_c + 1/2)^{3/2}}{2 M_A \alpha_c}\right] \times \frac{0.16 (2\alpha_c)^{2\alpha_c} \chi_8 n_6^{3/2-f}}{(2.4)^e (2\alpha_c + 1/2)^{2\alpha_c + 1/2} \xi B_{-3}} v_{10}^{3+e}. \tag{41}$$

Note that the temperature increases with the density and the ionization fraction since these give a higher collision rate. Most significant is the strong dependence on the shock velocity. The maximum temperature is not proportional to v^2 as in a J-shock but depends sensitively on the cooling mechanism.

The actual temperatures reached under the possible cooling laws can now be seen. For H_2O cooling with $e = 1$ and $\alpha_c = 2$

$$\left(\frac{T_{max}}{10^3 \text{ K}}\right)^{3/2} = 14.3[1 - 0.27(26/M_A)] \times \left(\frac{n_6^{1/2}}{B_{-3}}\right)\left(\frac{\chi}{10^{-7}}\right)\left(\frac{4.25 \times 10^{-4}}{x(H_2O)}\right)\left(\frac{v}{40 \text{ km s}^{-1}}\right)^4, \tag{42}$$

where $M_A/26 = (n_6^{1/2}/B_{-3})(v/40 \text{ km s}^{-1})$ and LTE H_2 cooling ($\alpha_c = 1.5$) gives

$$\left(\frac{T_{max}}{10^3 \text{ K}}\right)^{3.3} = 56.3[1 - 0.25(26/M_A)] \times \left(\frac{n_6^{3/2}}{B_{-3}}\right)\left(\frac{\chi}{10^{-7}}\right)\left(\frac{v}{40 \text{ km s}^{-1}}\right)^4. \tag{43}$$

Low-density non-LTE H_2 cooling gives

$$\left(\frac{T_{max}}{10^3 \text{ K}}\right)^{3.3} = 1.4 \times 10^2 [1 - 0.27(26/M_A)] \times \left(\frac{n_6^{1/2}}{B_{-3}}\right)\left(\frac{\chi}{10^{-7}}\right)\left(\frac{v}{40 \text{ km s}^{-1}}\right)^4. \tag{44}$$

Fig. 4 displays some values for separate LTE H_2 and H_2O cooling derived from equations (42) and (43). If we fix the H_2O fraction at the indicated value we can add together the cooling functions of H_2 and H_2O provided $n_n < 3 \times 10^6$ cm^{-3}. Then, since $f = 1$ for both cooling processes, we can still employ equation (21). The result is shown in Fig. 4. Reasonable agreement with DRD (fig. 9) is found on taking $n_6 = 0.5$ and $\chi = 2$. Note that for this case the two

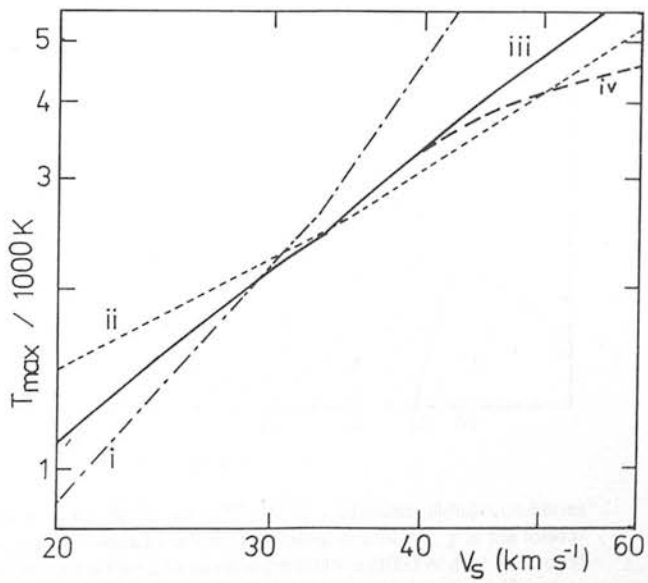


Figure 4. The maximum temperature reached in a C-shock with $\chi_8 = 10$ for a range of shock velocities with (i) H_2O cooling ($n_6^{1/2} = B_{-3}$) and (ii) LTE H_2 cooling H_2 ($n_6^{3/2} = B_{-3}$). The combined non-LTE $H_2 + H_2O$ case with $\chi_8 = 20$ and $n_6 = 0.5$ is shown as line (iii). The correspondence with a detailed numerical calculation from DRD is even closer if dissociative cooling is included (line iv).

cooling mechanisms both contribute significantly near 4000 K.

4.2 Dissociation of molecular hydrogen

C-shocks can accelerate molecular hydrogen to higher velocities than J-shocks without dissociation. DRD found a breakdown velocity $v_s = v_b \sim 40$–$45\ \mathrm{km\ s^{-1}}$ for specific values of $n^{1/2}/B$ and high H_2O conversion at $T > 500$ K. To test the generality of these conclusions we can calculate the total dissociation through a shock using the model developed here.

An approximation for collision-induced dissociative cooling for a pure molecular gas has been calculated from the results presented by Roberge & Dalgarno (1982) and Lepp & Shull (1983). We take the dissociation rate per unit volume as $n_n^2\langle\sigma v\rangle_d$ where the rate coefficient is

$$\langle\sigma v\rangle_d = 8.7\times10^{-9}\kappa(n,T)\exp(-54700/T)\ \mathrm{cm^3\ s^{-1}},\quad T<4000\ \mathrm{K} \qquad (45a)$$

$$\langle\sigma v\rangle_d = 5.3\times10^{-9}\kappa(n,T)\exp(-52700/T)\ \mathrm{cm^3\ s^{-1}},\quad T>4000\ \mathrm{K}. \qquad (45b)$$

Here radiative stabilization is expressed by $\kappa(n,T)$ with $\kappa=1$ for $n \geqslant 10^7\ \mathrm{cm^{-3}}$ and is a strong function of density below 4000 K (Lepp & Shull 1983). The total dissociation fraction per unit length is $n_n\langle\sigma v\rangle_d/v_n$. Integrating, the total dissociative fraction f (provided $f \ll 1$) becomes

$$f = \frac{n_{n0}L_0}{v_s}\int \frac{\langle\sigma v\rangle_d}{r^2}\frac{dx}{L_0} \qquad (46)$$

on using equation (1). Putting $s = x/L_0$ and substituting for L_0 yields

$$f = \frac{(m_n/\mu_{ni})}{\chi_i\sigma_0(v_s/v_0)^e v_0}\int_0^\infty \frac{\langle\sigma v\rangle_d}{r^2}\,ds. \qquad (47)$$

Thus f is inversely proportional to the ionization fraction; it is not independent of the density, however, which appears in the rate coefficient. On substitution, with $e=1$ and $\mu_{ni}=m_n$, we obtain $f = 5\times10^{16}(\chi_8 v_{10})^{-1}\int[\langle\sigma v\rangle_d/r^2]\,ds$.

The characteristic length scales can be derived from L_0. The length scale over which the neutrals are warm is $\Delta L_n = 2L_0/M_A$ (from equation 5) and over which the ions are decelerated $\Delta L_i = L_0/M_A^2$ (equation 6). The latter corresponds to the magnetic precursor length scale (Draine 1980).

At this stage we could resort to numerical calculations. Several good approximations, however, allow us to obtain an analytical result. We can take equation (45b) for the dissociation rate since it remains 25 per cent accurate above 3000 K, the critical dissociation temperature range. The exponential temperature dependence of $\langle\sigma v\rangle_d$ allows us to simply substitute T_{max} for T, set $r=0.8$ and integrate over an interval $\Delta s \sim 0.12\ B_{-3}/(v_{10}n_6^{1/2})$ (somewhat smaller than ΔL_n, see Figs 1 and 2). Hence

$$f \sim 4.8\times10^7[B_{-3}/\chi_8 v_{10}^2 n_6^{1/2})]\exp(-52700/T_{max}). \qquad (48)$$

It is clear that f depends strongly on the velocity through T_{max} and only weakly on the other parameters.

A dissociation temperature T_d is obtained from equation (48). If T_{max} exceeds T_d then the dissociation fraction is high. Making the requirement that $f<0.3$ gives $T_d \sim 4090$ K and $T_d \sim 3560$ K for $\chi_8 v_{10}^2 n_6^{1/2} B_{-3} = 400$ and 60, respectively. The former values correspond well with DRD. We shall take $T_d = 4000$ K since this is accurate at high shock speeds and the highest temperatures are in fact reached when $\chi_8 v_{10}^2 n_6^{1/2}/B_{-3}$ is large. Moreover the temperature range 3500–4000 K is sufficiently narrow and the $T_{max}-v_s$ relation sufficiently steep (Fig. 4) for this to have a minimal effect on the resulting breakdown velocity (less than 6 per cent in v_b).

The breakdown velocity for H_2O cooling is found by substituting T_d for T_{max} in equation (41)

$$v_b/(40\ \mathrm{km\ s^{-1}}) = 0.94\{[x(H_2O)/4.25\times10^{-4}](B_{-3}/n_6^{1/2})(10^{-7}/\chi)\}^{1/4}. \qquad (49)$$

We set $M_A = 26$ where it appears explicitly in sub-section 4.1 to allow these limits to be displayed in Figs 5 and 6 without recourse to covering each parameter individually. We also take the high-velocity value of $e=1$ here but use the value given by equation (41) in the figures. Note that if n_6/B_{-3}^2 is fixed then the breakdown velocity is only a function of χ. This accounts for the more or less constant breakdown velocity obtained in previous calculations (DRD; Chernoff & McKee 1989).

A higher density with a fixed magnetic field gives a lower breakdown velocity but with a small dependence; $v_b \propto n^{-1/8}$. A velocity $v_b = 88\ \mathrm{km\ s^{-1}}$ is reached by taking the perhaps extreme parameters $B_{-3}/(\chi_8 n_6^{1/2}) = 3$. However such speeds cannot be reached since self-ionization will provide a more stringent condition (see below).

Non-LTE H_2 cooling produces a similar result

$$v_b/(40\ \mathrm{km\ s^{-1}}) = 1.06[(B_{-3}/n_6^{1/2})(10^{-7}/\chi)]^{1/4}. \qquad (50)$$

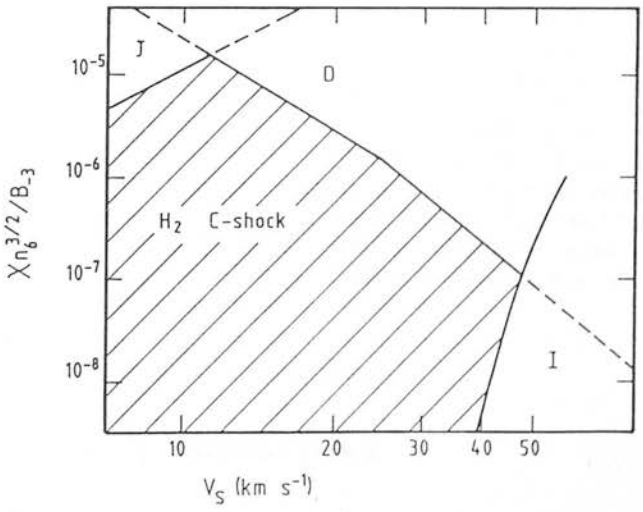


Figure 5. The dense ($n > 10^6\ \mathrm{cm^{-3}}$) molecular cloud conditions which allow C-shocks when H_2 cooling dominates. χ is the ionization level and is taken to include the relative effect of dust on neutral deceleration. J, D and I are the J-shock, dissociation and self-ionization limits. Self-ionization here limits high-velocity shocks according to the ionization fraction only. Hence the I limit displayed here is appropriate for $n_6^{3/2}/B_{-3} = 1$ and must be shifted for other values. This limit on the shock velocity is, however, only weakly dependent on this value for $M_A \gg 1$.

Figs corrected from erratum MNRAS 224, 384

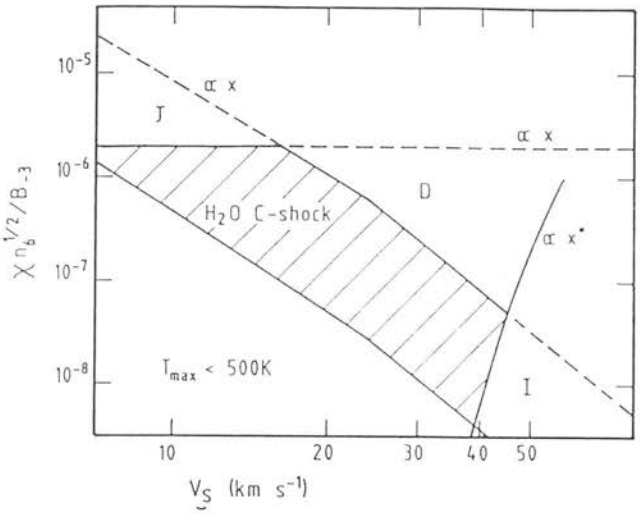


Figure 6. The C-shock regime for H_2O-dominant cooling. The positions of lines relative to the explicit abundance shown of 4.25×10^{-4} are indicated. The H_2O formation rate rises rapidly once $T>500$ K, thus providing the lower boundary.

for $n_6<2$. This implies that we do not have to rely on the oxygen abundance and its conversion to H_2O to predict the maximum shock speed. This is due to the two forms of cooling having comparable cooling rates at the crucial dissociation temperature of 4000 K. Taking both cooling processes together and fixing x, however, only increases v_b by a factor of ~ 1.2. At densities below 10^6 the dissociation rate coefficient is much lower and again self-ionization is the breakdown process.

At high densities, LTE H_2 cooling yields

$$v_b/(40\ \mathrm{km\ s^{-1}}) = 1.2[(B_{-3}/n_6^{1/2})(10^{-7}/\chi)/n_6]^{1/4}. \quad (51)$$

Consequently, in high-density regions molecular hydrogen is dissociated by relatively slow shocks (in the absence of H_2O cooling) with $v_b \propto n^{-3/8}$. With H_2O depleted by a factor of 10 from the value of 4.25×10^{-4} and $n_6>27(B_{-3}/n_6^{1/2})(10^{-7}/\chi)$ the breakdown velocity is just 21 km s^{-1}.

Dissociation will also cool the gas. The cooling rate is $\sim\langle\sigma v\rangle_d n^2(4.48\ \mathrm{eV})$ and can be compared to the cooling rates given in Section 4.1. It can provide a moderate additional cooling at densities near 10^6 and reduce temperatures from 4000 K by ~ 200 K. At higher densities dissociative cooling can exceed H_2 cooling and be comparable to H_2O cooling. We will now see, however, that the breakdown velocity will then be determined by self-ionization rather than dissociation.

4.3 Dissociation and ionization by streaming ions

Dissociation of molecular hydrogen can also occur through collisions with the streaming ions. In Appendix C we evaluate the dissociation fraction through a C-shock and find $f\sim 0.8\,\gamma_1[(8E_0)/(m_n v_s^2)]^{1/2}(\chi_i/\chi)=3.3\gamma_1 v_{10}^{-1}(\chi_i/\chi)$ where γ_1 is the fraction of those 'orbiting' collisions with centre-of-mass energy exceeding the threshold value $E_0=4.48$ eV which leads to dissociation (see DRD). This result assumes $m_n(v_n-v_i)^2/(2E_0)\gg 1$ but, provided it exceeds unity, f remains accurate to within a factor of 2. This implies streaming speeds $v_n-v_i>21$ km s^{-1}. This converts to shock speeds >24 km s^{-1}, within our regime of interest. Hence ion-impact dissociation occurs over a broad shock region as compared to neutral–neutral dissociation.

In high-density regions, DRD take $\chi_i<0.1\chi$ giving a maximum possible dissociated fraction of order 10 per cent. Complete ion impact dissociation is conceivable in lower density regions ($n_n\sim10^4$ cm^{-3}) where $\chi_i\sim\chi$. Even low values of γ_1 will produce a high fraction of H atoms compared to the pre-shock values of $n(\mathrm{H})<10^{-5}n_n$ in high-density cores (DRD). We shall assume that γ_1 is indeed low.

Ionization via the streaming ions places a definite limit to the shock velocity almost independent of the temperature reached by the neutrals. In Appendix D we demonstrate that recombination will balance locally the ionization if the streaming velocity remains small:

$$v_{10}(r-q)/(10\ \mathrm{km\ s^{-1}}) < 55.6 - 16.1\,T_4^{1/2} \times[1-0.76\log(\chi_i/10^{-7})], \quad (52)$$

where $T_4\sim T_n+0.044v_{10}^2(r-q)^2$, valid for $\chi_i\ll 2\times10^{-6}$. Using equation (15) the maximum value for $r-q$ is $\sim 1-(\frac{3}{5})M_A^{-2/3}\sim 0.83$ for $M_A=26$ which is likely to be in the upstream coupled zone (sub-section 3.5). At T_{max} $r-q\sim 2\alpha_c/(2\alpha_c+\frac{1}{2})\gtrsim\frac{4}{5}$. At high velocities the extra ions produced by streaming lead to further ionization as well as raising the dissociation rate by increasing the temperature. A runaway process ensues, as proposed by DRD). So we estimate self-ionization to occur for velocities exceeding 44–51 km s^{-1} for χ_i in the range 3×10^{-7}–3×10^{-8}. In Section 5 we employ (52) with $T_4=1$ and $r-q=0.83$.

4.4 J-shock limit

The temperature remains small and the flows are continuous when $\psi_0\gtrsim M_A$. Then we find $t_{max}\sim(0.24\,M_A/\psi_0)^{1/\alpha}$ from equation (21). If t_{max} is of order unity then the structure approximates to that of a J-shock with drag and cooling taking place in separated regions. The subsonic flow section is no longer confined to the far-downstream region but moves towards $r=(\gamma-1)/(\gamma+1)$ (Chernoff 1987). Here we divide C- and J-shocks by the critical value $t_{max}=0.2$ but note that there may be a considerable range in which the C-shock/J-shock transition occurs and, moreover, the solution may not be unique. Using this value in equation (21) yields the C-shock condition

$$\frac{\chi_8 n_6^{3/2-f} v_{10}^{3+e-2\alpha}}{\xi B_{-3}} < 6.3(2.4)^e(4.8)^\alpha \quad (53)$$

which is utilized in Section 5.

5 DISCUSSION

We have developed a simple but powerful method of analysing cool C-shocks. Radiative cooling keeps the temperature low and the thermal pressure is divorced from the dynamics. These 'cool' flows possess many of the dynamical characteristics of the 'cold' flow discussed by Chernoff (1987). Neutral and ion velocities are thus obtained independent of the cooling function. Consequently, to predict the emission

from a given energy level, we only need to know the cooling function over the temperature range relevant to the energy level. This method is shown to be accurate after an initial upstream phase in which the thermodynamics and dynamics are coupled.

Three limits on the parameters for which C-shocks occur have been found. The rapid dissociation of molecular hydrogen at high temperatures limits the shock speed. The limit of $v_b \sim 40$ km s^{-1} is found for strong H_2O and/or non-LTE H_2 cooling but can be greatly exceeded for high field strengths and low ionizations. High-energy collisions with the streaming ions, however, efficiently produce new ions which increase the heating as well as produce further ions. This yields a breakdown velocity that depends on the level at which the process becomes catastrophic. The high-velocity dependence of the ionization level indicates $v_b \sim 40$–50 km s^{-1}. This results partly from the fact that the streaming velocity is a large fraction of the shock velocity when M_A is large. Sufficiently strong magnetic fields (e.g. $B_{-3}/n_6^{1/2} \sim 10$) can yield values of M_A close to unity. This regime will be explored elsewhere.

The accumulated data for all limits, as derived in Section 4, are presented in Figs 5 and 6 for H_2- and H_2O-dominated cooling, respectively. Note that the ordinate in the H_2 involves an extra density factor. The figures display C-shock regimes as a function of $n_6^{1/2}/B_{-3}$, the H_2O abundance, x, relative to the value of 4.25×10^{-4}, as well as ionization parameter and shock velocity. They demonstrate that in dense molecular clouds (i) H_2 C-shocks with $\chi n > 10$ do not occur if $n_6^{1/2}/B_{-3}$ is near unity, (ii) high-velocity H_2 C-shocks require an efficient recombination process to avoid rapid dissociation (e.g. the H_2O abundance must exceed $\sim 10\chi_i$), (iii) the H_2O C-shock regime is limited by the requirement that H_2O formation is high and (iv) H_2O C-shocks with $\chi > 2 \times 10^{-6}$ do not occur.

The results presented here will be employed to calculate the observable properties of planar and bow shocks in forthcoming papers. It will be seen that there is a complex relationship between shock velocity, molecular-hydrogen linewidths and line ratios.

REFERENCES

Brand, P. W. J. L., Toner, M. P., Geballe, T. R., Webster, A. S., Williams, P. M. & Burton, M. G., 1989. *Mon. Not. R. astr. Soc.*, **236**, 929.

Burton, M. G., Brand, P. W. J. L., Geballe, T. R. & Webster, A. S., 1989. *Mon. Not. R. astr. Soc.*, **236**, 409.

Chernoff, D. F., 1987. *Astrophys. J.*, **312**, 143.

Chernoff, D. F. & McKee, C. F., 1989. In: *Molecular Astrophysics*, ed. Hartquist, T., Cambridge University Press, Cambridge, in press.

Draine, B. T., 1980. *Astrophys. J.*, **241**, 1021.

Draine, B. T., Roberge, W. G. & Dalgarno, A., 1983. *Astrophys. J.*, **264**, 485.

Hollenbach, D. & McKee, C. F., 1979. *Astrophys. J. Suppl.*, **41**, 555.

Kwan, J., 1977. *Astrophys. J.*, **216**, 713.

Lepp, S. & Shull, J. M., 1983. *Astrophys. J.*, **270**, 578.

McKee, C. F., Storey, W. V., Watson, D. M. & Green, S., 1982. *Astrophys. J.*, **259**, 647.

Myers, P. C. & Goodman, A. A., 1988. *Astrophys. J.*, **326**, L27.

Pineau des Forêts, G., Flower, D. R., Hartquist, T. W. & Dalgarno, A., 1986. *Mon. Not. R. astr. Soc.*, **220**, 801.

Roberge, W. & Dalgarno, A., 1982. *Astrophys. J.*, **255**, 176.

APPENDIX A

The conservation laws for momentum and energy flux are (Chernoff's set of equations 20–23 rewritten for general γ)

$$\frac{d}{dx}(n_n m_n v_n^2 + n_n k T_n) = F, \tag{A1}$$

$$\frac{d}{dx}\left[\left(\frac{v_s}{v_i}\right)^2 \frac{B_0^2}{8\pi}\right] = -F, \tag{A2}$$

$$\frac{d}{dx}\left(\tfrac{1}{2} n_n m_n v_n^3 + \frac{\gamma}{\gamma-1} n_n k T_n v_n\right) = F v_n + \Gamma + G \tag{A3}$$

and

$$\frac{d}{dx}\left[v_i \left(\frac{v_s}{v_i}\right)^2 \frac{B_0^2}{4\pi}\right] = -F v_i, \tag{A4}$$

where

$$F = n_n n_i \langle \sigma v \rangle^{cl} \mu_{ni} (v_i - v_n), \tag{A5}$$

$$G = \frac{2\mu_{ni}}{m_n + m_i} n_n n_i \langle \sigma v \rangle^{cl} [k(T_i - T_n)/(\gamma - 1) + m_i (v_i - v_n)^2/2] \tag{A6}$$

and

$$\Gamma = -n_n^2 \langle \sigma v \rangle^{I} k T_n. \tag{A7}$$

Here $\mu_{ni} = m_n m_i/(m_n + m_i)$, v_s is the shock speed, F and G are the momentum and energy transfer rates between neutrals and ions and $\langle \sigma v \rangle^{I}$ and $\langle \sigma v \rangle^{cl}$ are the cooling and drag rate coefficients (defined in the main text). Equations (A2) and (A4) are identical (since B_0 and v_s are constants).

APPENDIX B

With coordinates q and r the initial state is $(1, 1)$ and $(q, r) = (q_*, r_*)$ at $M = M_*$. Using equations (15) and (17) and setting $dM/dq = 0$ yields

$$r_* = \frac{\alpha_c - \alpha}{\alpha_c + 1/4 - \alpha}\left[1 - \frac{(e+2)(\alpha_c - 1/2 - \alpha)(\alpha_c - 1/4 - \alpha)^{1/2}}{2^{3/2} M_A (\alpha_c - \alpha)} + 0\left(\frac{1}{M_A^2}\right)\right], \tag{B1}$$

$$q_* = \frac{(2\alpha_c + 1/2 - 2\alpha)^{1/2}}{A_A} + 0\left(\frac{1}{M_A^2}\right). \tag{B2}$$

for $\alpha < \alpha_c$. This approaches the position of the maximum temperature for $\alpha = 0$ and is located downstream of t_{max} for $\alpha > 0$ (Fig. 1). This gives

$$M_*^{2\alpha} = \frac{\psi_0}{M_A}\left[\frac{(2\alpha_c + 1/2 - 2\alpha)^{2\alpha_c + 1/2 - 2\alpha}}{\gamma^{\alpha}(2\alpha_c - 2\alpha)^{2\alpha_c - 2\alpha}}\right]. \tag{B3}$$

For $\alpha > \alpha_c$ the shock is qualitatively different, the minimum M now occurring after the neutrals have been greatly

decelerated (Fig. 2)

$$r_* = \frac{(2\alpha - f)}{2^{1/2}(2\alpha - 2\alpha_c)\, M_A} + 0(1/M_A^2), \quad q_* = \frac{1}{2^{1/2} M_A} + 0(1/M_A^2), \tag{B4}$$

$$M_*^{2\alpha} = \frac{\psi_0}{M_A^{1+2(\alpha-\alpha_c)}} \times \left[\frac{(2\alpha - f)^{2\alpha - f}}{2^{1/2}\gamma^{\alpha}(2\alpha_c - f)^{2\alpha_c - f}[2(2\alpha - 2\alpha_c)]^{2\alpha - 2\alpha_c}} + 0(M_A^{-1}) \right]. \tag{B5}$$

APPENDIX C

The rate coefficient for ion impact dissociation is (DRD, equation 44)

$$\langle \sigma v \rangle = \frac{\sigma_0 x_0}{s} \left(\frac{2kT_r}{\pi m_r} \right)^{1/2} \int_{s_0}^{\infty} [e^{-(x-s)^2} - e^{-(x+s)^2}]\, x\, dx \tag{C1}$$

on taking separate Maxwellian velocity distributions for ions and neutrals, where $\sigma_i = 7.2 \times 10^{-16} \gamma_1$ and γ_1 is a reaction fraction, $x_0^2 = E_0/kT_r$, $s^2 = m_r v^2/(2kT_r)$, $m_r = m_i m_n/(m_i + m_n)$, $T_r = (m_i T_n + m_n T_i)/(m_i + m_n)$ and $v = v_i$ is the streaming velocity. Here $E_0 = 4.48$ eV and we take $m_r = m_n$.

The quantity $s^2/x_0^2 = m_n v_s^2 (r-q)^2/(2E_0)$ exceeds unity in regimes of high-velocity shocks where $r - q \gg 0$. This gives a high rate coefficient taking $s - x_0 \gg 1$ and integration yields $\langle \sigma v \rangle \approx \sigma_i (8E_0/m_n)^{1/2}$. The dissociated fraction in the shock is given as in equation (46) which yields (for $e = 1$)

$$f = \frac{\chi_i}{\sigma_0 v_s \chi} \int_{r_1}^{r_2} \frac{\langle \sigma v \rangle}{(r-q)^2}\, dr \sim 0.8\, \gamma_1 (\chi_i/\chi)[8E_0/(m_n v_s^2)]^{1/2} \tag{C2}$$

to within a factor of 2 where the integration limits are set by $s > x_0$. Note that the dissociated fraction decreases with increasing velocity. This arises from the high streaming speeds – once collision energies are above the threshold value, a further increase of v_s only decreases the shock width.

APPENDIX D

Radiative stabilization rapidly decreases the dissociation rate at densities below $\sim 10^5$ cm^{-3} (Roberge & Dalgarno 1982; Lepp & Shull 1983). Then breakdown occurs through self-ionization (DRD). Self-ionization may also be crucial to limiting velocities of C-shocks under those high-density conditions for which we have found dissociation to be less effective. The most important ionization mechanism found by DRD was collisional ionization by streaming ions. This ionization rate is balanced locally by recombination. As in Appendix A we take the rate coefficient from DRD (their equation 29) for the reaction $H_2 + I^+ = H_2^+ + I^+ + e$ with a cross-section $7.7 \times 10^{-19}\ (2I_H/E_0)$ cm^2 where I_H is the ionization potential of the ion and now $E_0 = 14.7$ eV:

$$\langle \sigma v \rangle = \frac{\sigma}{s x_0^2} \left(\frac{2kT_r}{\pi m_r} \right)^{1/2} \int_{x_0}^{\infty} [e^{-(x-s)^2} - e^{-(x+s)^2}] x^2 (x^2 - x\ 2_0)\, dx. \tag{D1}$$

Here $x_0^2 = 35.8/T_4$ and $s_2 = 1.21\, m_r v_{10}^2 (r-q)^2/(m_n T_4)$ where $T_4 = T_r/10^4$ K. We take an approximation to the integrand of (D1) for $2x_0(x_0 - s) \gg 1$: $x_0^3 \exp[-(x_0 - s)^2]/[2(x_0 - s)^2]$. With $\gamma = \frac{2}{5} T_r = T_n + (m_r/m_i) m_n v^2 (r-q)^2/(5\ \mathrm{K})$.

The reaction $H_2^+ + H_2 \rightarrow H_3^+ + H$ is rapid through the dominance of the H_2. $H_3^+ + H_2O \rightarrow H_3O^+ + H_2$ follows once $T > 500$ K since H_2O is then also abundant. The H_3O ions then combine with electrons: $H_3O^+ + e \rightarrow OH + H_2$ with a rate coefficient $\langle \sigma v \rangle_r = 5.7 \times 10^{-7} (300\ \mathrm{K}/T)^{1/2}$ (Pineau des Forêts *et al.* 1986) or with H atoms (produced by limited dissociation) to give H_2O^+. The latter has a similar recombination rate coefficient.

The balance $n n_i \langle \sigma v \rangle = n_i^2 \langle \sigma v \rangle_r$ then requires an ionization fraction $\chi_i = \langle \sigma v \rangle / \langle \sigma v \rangle_r = 6.4 \times 10^{-7} T_4^{\frac{1}{2}} \exp[-35.8(1 - 0.18U)^2/T_4]/[U(1 - 0.18U)^2]$ where $U = v_{10}(r-q)$. Fixing $T_4 = 1$, $m_i = 10 m_n$ and $U = 4$ in all but the exponential term then yields a limit to the streaming velocity for a presumed ionization fraction (equation 52) valid for $\chi_i \ll 2 \times 10^{-6}$. A solution for high ionization levels can only be obtained by a full numerical integration.

Mon. Not. R. astr. Soc. (1990) **243**, 498–503

Signatures of C-shocks in molecular clouds

M. D. Smith and P. W. J. L. Brand
Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ

Accepted 1989 October 13. Received 1989 July 31

SUMMARY

We attempt to interpret the molecular hydrogen emission lines from dense interstellar clouds by planar C-shocks. The lines are predicted to be narrow if H_2O dominates the cooling, the full width being less than 20 km s^{-1} even for shock velocities near the breakdown level. In contrast, the lines are broad if H_2 cooling is dominant (assuming the shock plane is face-on). The profiles are single peaked, double peaked, triangular or flat topped according to the upper energy level and the shock parameters. For double peaked profiles the maxima are separated by ~ 30 km s^{-1} for a shock velocity of 40 km s^{-1}. Excitation temperatures and column densities are obtained. The line ratio data accumulated for OMC-1 cannot be interpreted by a single planar C-shock.

1 INTRODUCTION

Shocks driven through dense clouds in the interstellar medium produce wide emission line profiles from molecular hydrogen. The full widths correspond to velocities in the range 20–140 km s^{-1} (e.g. Nadeau & Geballe 1979; Brand *et al.* 1989a; Burton, Geballe & Brand 1989). Yet the classical J-shocks in which the heating and cooling layers are adjacent may accelerate molecular hydrogen to velocities of only about 25 km s^{-1} without wholesale dissociation occcurring (Kwan 1977). A dynamical model in which low-velocity shocks propagate through pre-accelerated fast-moving clouds is then invoked. Such emission may also rise from a C-shock where an extended heating region overlaps with the cooling layer (Draine 1980). This inhibits the rise to temperatures at which dissociation takes place allowing velocities of up to 40–50 km s^{-1}. It is then possible that a C-shock signature will not be lost within the complexity of a dynamical configuration and may thus be identified. Here we explore the possibilities of observing planar C-shocks by using a simple model (Smith & Brand 1990, hereafter Paper I) to predict profile widths and shapes, line intensities and ratios for a wide range of conditions. Planar shocks may prove particularly relevant to shock interactions of supernova remnants with molecular clouds. This work is also a step towards our final goal of predicting H_2 line properties from bow shocks.

In Paper I we determined the parameter regimes for which C-shocks occur. A simple treatment of the major processes as determined by detailed studies (Draine 1980; Draine, Roberge & Dalgarno 1983) allowed the MHD flow properties to be extracted and understood for a wide range of basic molecular cloud quantities. The shock breakdown velocity, above which dissociation or self-ionization is overwhelming, was found. Its maximum possible value lies in the narrow range 40–50 km s^{-1} (see also Chernoff & McKee 1989), although the means of breakdown depends on the chosen pre-shock parameters. We shall mainly investigate shocks with velocities of 40 km s^{-1}, offering a good possibility of reasonable velocity resolution in the near future. It has already been shown that, for a narrow range of velocities near this value, some of the hydrogen line intensities in OMC-1 can be reproduced (Chernoff, Hollenbach & McKee 1982; Draine & Roberge 1982). More recent data (Brand *et al.* 1988), however, have demonstrated that these particular C-shock models are inconsistent with the column densities derived from high upper energy levels.

In this study we determine C-shock signatures as a function of emission line upper energy, velocity, cooling mechanism (H_2O or H_2), degree of ionization, magnetic field and density. The only effect of the dust included in this work is the drag on the neutrals. This is expressed within the ionization fraction (Paper I). Ions and dust are assumed to be frozen to the magnetic field. First we establish how the line profiles are extracted from the neutral molecular hydrogen flow patterns as calculated in Paper I.

2 THE METHOD

A one-dimensional steady fluid pattern moving with a shock speed v_s through a stationary medium is taken. The distance is denoted by x. The medium is initially of density n_{no}, ionized fraction $\chi = n_{io}/n_{no}$ and H_2O abundance relative to hydrogen atoms of ε. The magnetic field B_0 is the component in the shock plane, an additional angle dependence being probable (Wardle & Draine 1987).

The emission line properties of H_2 can be derived once we know the H_2 density n_n and temperature T_n as functions of x. The derivation is relatively simple if the levels are populated with a Boltzmann distribution at the local kinetic temperature. This limits us to high densities ($n_n > 10^6$ cm^{-3}). The column density N_j of level j with energy kT_j and space density n_j is $N_j = \int n_j\,dx = n_{no}\int (n_j/n_n)(n_n/n_{no})\,dx$ which becomes

$$N_j/g_j = n_{no}\int (n_n/n_{no})\,[\exp(-T_j/T)]/Q(T)\,dx, \qquad (1)$$

where g_j is the statistical weight and $Q(T)$ is the molecule partition function which we take as $Q(T) = (T/44\ \mathrm{K})\,[1-\exp(-6000\ \mathrm{K}/T)]^{-1}$ for H_2.

In Paper I we found that we can follow the neutral density ($n_n = rn_{no}$) and the ion density ($n_i = qn_{io}$) through a shock

$$L(dq/dx) \sim q^2(r-q)\,M_A^2/r, \qquad (2)$$

where M_A is the pre-shock Alfvén Mach number given by $M_A^2 = m_n n_{no} v_s^2/(B_0^2/4\pi) = 42.3\ n_6 v_{10}^2/B_{-3}^2$. We shall write $T_n = 1000 T_3$ K, $n_n = 10^6 n_6$ cm^{-3}, $B = 10^{-3} B_{-3}$ G, $\chi = 10^{-8}\chi_{-8}$ and $v_s = 10 v_{10}$ km s^{-1}. L depends on the ion-neutral drag coefficient which is presumed to be $\propto (v_s/v_0)^e (r-q)^e$ with $v_0 = 24$ km s^{-1} and $0 < e < 1$. Thus we define

$$L_0 = (m_n/\mu_{ni})(v_s/v_0)^{1-e}/(n_{io}\sigma_0) = (r-q)^e L \qquad (3)$$

as the length scale with μ_{ni} the reduced mass ($\sim m_n$ for $m_i \gg m_n$) and $\sigma_0 \sim 10^{-15}$ cm^2.

These flows are cool in the sense that $t = T_n/T_0 = kT_n/(m_n v_s^2)$ remains small. With t typically < 0.01 the flows approximate to that of a cold flow with some important differences (Chernoff 1987; Paper 1). A flow either remains supersonic throughout or possesses a subsonic section far enough downstream to have a negligible influence (provided $\chi \ll 10^{-5}$). In the upstream flow, where the neutrals are beginning to decelerate, the dynamics and thermodynamics are inseparable. This region, although important for profile details, has a minimal effect on the column densities and line ratios. As a result, the specific heat ratio γ has a small but not ignorable influence on line profiles and the full equations (Paper I, equations 9 and 10)

$$\frac{dr}{dq} = (\gamma-1)\,\frac{(r-q)[\gamma r/(\gamma-1) - q] - \Psi t}{rq^3(r-q)(1-\gamma t/r^2)\,M_A^2}, \qquad (4)$$

$$r + \frac{t}{r} + \frac{1}{2M_A^2 q^2} = 1 + \frac{1}{2M_A^2} \qquad (5)$$

must be integrated. Here Ψ is a ratio of drag to cooling timescales. Together with equation (2) we can solve numerically for $q(x)$, $r(x)$ and $t(x)$.

The cooling per unit volume, $\Gamma = \Sigma n_n^{1+f}\Lambda(T)$, comes in through Ψ. If we take $\varepsilon = 4.25\times10^{-4}$ then H_2O cooling will dominate (the exception being shocks which reach a temperature close to 4000 K for which H_2 contributes significantly). Then $\Gamma(H_2O) \sim 11.6\times10^{-15} n_6^2 T_3^{3/2}$ erg s^{-1} cm^{-3}. For $\varepsilon < 5\times10^{-5}$ only H_2 cooling will be important above 500 K, and $\Gamma = 3.3\times10^{-15} n_6 T_3^{3.3}$ is the best power-law fit for 1000 K $< T <$ 5500 K (Paper 1). We here limit our analysis to power laws since we wish to promote a physical understanding of how flow patterns are related to line profiles.

Paper I demonstrated that a shock constant is then $\Psi_0 = tr^f(r-q)^e\Lambda(T_0)\Psi/[q\Lambda(tT_0)]$. Substitution into equation (1) yields

$$N_j/g_j = \frac{44 n_{no} L_0}{T_0 M_A^2}\int_0^1 e^{-T_j/(tT_0)}\,[1-e^{-6000\ \mathrm{K}/(tT_0)}]\,\frac{dq}{tq^2(r-q)^{1+e}}. \qquad (6)$$

We write $\Delta N_j/g_j = 44 n_{no} L_0 C_c(\Delta v)$ where ΔN_j is the column density within the velocity interval Δv. Hence line profiles are simply given by plotting ΔN_j against v. A rather narrow Δv of 1 km s^{-1} is taken to emphasize the profile details. It is straightforward to deduce the prospects of observing features at lower resolution.

A Boltzmann slab factor is defined as

$$N_{jB}/g_j = \frac{44 n_{no} L_0}{T_B}\,e^{-T_j/T_B}\,[1-e^{-6000\ \mathrm{K}/T_B}], \qquad (7)$$

the equivalent column density of a slab of width L_0, temperature T_B and constant density n_{no}. For a given shock we fix $n_{no}L_0$ so that $N_{jB} = N_j$ at $T_j = 7000$ K. We choose 7000 K since this corresponds closely to the 1–0 S(1) line from $T_j = 6956$, the most precisely observed line. The 'excitation temperature' is the value of T_B which reproduces a measured line ratio.

3 RESULTS

Numerical solutions have been obtained. We concentrate mainly on H_2 cooling since, as will be seen, the profiles can be broad. Results for H_2O cooling are presented in Subsection 3.6.

The diagrams are drawn from the view of a planar shock propagating along the line-of-sight through a stationary quiescent cloud ($v = 0$). If the observer is at an angle θ to the shock direction, then all velocities are reduced by a factor cos θ. We take $\gamma = 7/5$ unless otherwise specified.

3.1 Shock velocity

A profile varies in shape and width as the shock velocity changes. As shown in Fig. 1 for $T_j = 7000$ K, the profile is

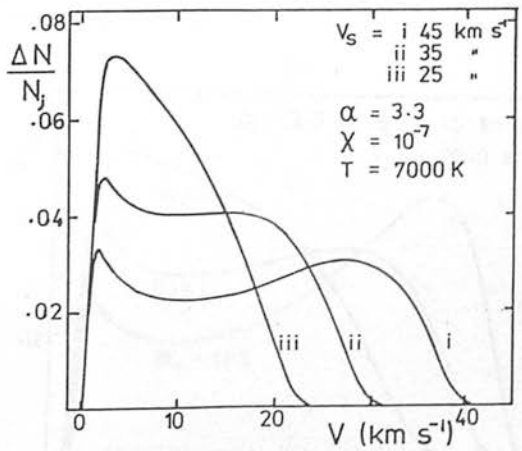


Figure 1. H_2 line profiles from LTE H_2 cooling for various shock velocities keeping $n_6/B_{-3} = 1$. The profiles are normalized to constant area.

double peaked for $v_s = 45$ km s^{-1} and single peaked for 25 km s^{-1}. In essence, this is because relatively high temperatures are reached at the higher shock velocity with $T_{max} \propto v^{1.21}$ (see Paper I, equation 43). Then much emission is produced before and after the temperature has reached the maximum. At low v_s the region near the peak temperature dominates a profile since this is the only location that is sufficiently hot.

The full width corresponds to about 90 per cent of the shock velocity. The widths will appear broad even when observed with low signal-to-noise ratios since the lines are 'shouldered' rather than 'winged'. The wide profiles are a result of the mechanism of cooling. The cooling rate per unit volume is proportional to the density for H_2 cooling above the critical density. The increase in density through the shock therefore leads to only a modest increase in the cooling rate. The temperature is slow to decrease. This can be contrasted with the cooling rate $\propto n^2$ per unit volume for H_2O (see Subsection 3.6).

3.2 The cooling function

The profile shape depends on the cooling function. In Fig. 2 we take $\Lambda = 3.3 \times 10^{-15} n_6 T^{\alpha}$ with a steep cooling law ($\alpha = 4.7$, suitable for H_2 cooling below 2000 K) and a shallow cooling law ($\alpha = 2.0$, corresponding to CO-like cooling at densities above $2 \times 10^6 T^{3/4}$ cm^{-3} although a high CO abundance and a larger constant in Λ would be necessary to dominate H_2).

The flatter cooling amplifies the two bumps whereas the profile loses the second peak for steep cooling. Clearly the larger index (with the cooling rate at 1000 K fixed) produces rapid cooling at higher temperatures and so less emission in the downstream region.

3.3 The ionization fraction and H_2 density

The neutrals get hotter when there are more ions available since the collision rate increases. As expected this produces a more prominent second high-velocity bump in the profile (Fig. 3) with a deeper depression between the bumps. The peaks are separated by 30 km s^{-1}. A single low-velocity peak and a triangular profile are found for low ionization.

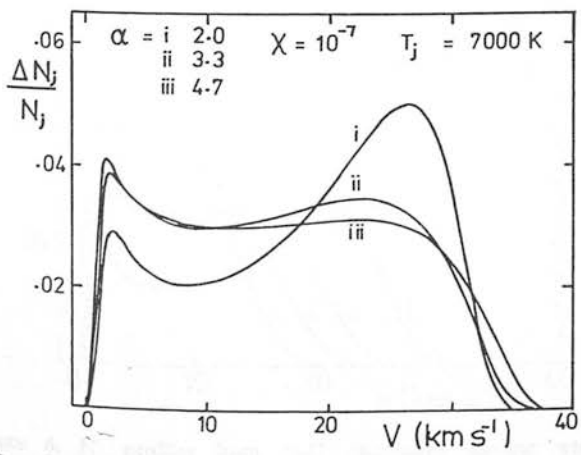


Figure 2. H_2 profiles for various cooling laws with $f = 0$ (i.e. above appropriate critical densities).

Ionization and density variations are equivalent here since χ enters the calculations through Ψ_0 and $\Psi_0 \propto (\chi n_{no})^{-1}$. Fig. 3 thus corresponds to fixing $B^2_{-3} = n_6$, i.e. M_A is held constant.

The specific heat ratio has a small effect on the profile. Taking $\gamma = 5/3$ (Fig. 3) somewhat increases the low-velocity peak.

3.4 The Mach number

Increasing the Mach number while holding the shock speed constant implies that $n_6/B^2_{-3} > 1$. Here Ψ_0, which is density dependent, is also constant. Thus field changes are being analysed. Fig. 4 demonstrates that the magnetic field can have a profound effect on the profile. In fact it can be seen that the field increase or decrease has almost precisely the opposite effect to the ionization change. This is understood from the manner in which these quantities affect T_{max} which is $\propto (\chi/B)^{0.3}$ (Paper I).

3.5 The upper energy level

Lines from high-energy levels reverse the characteristics of the $T_j = 7000$ K profiles (Fig. 5). The depression at

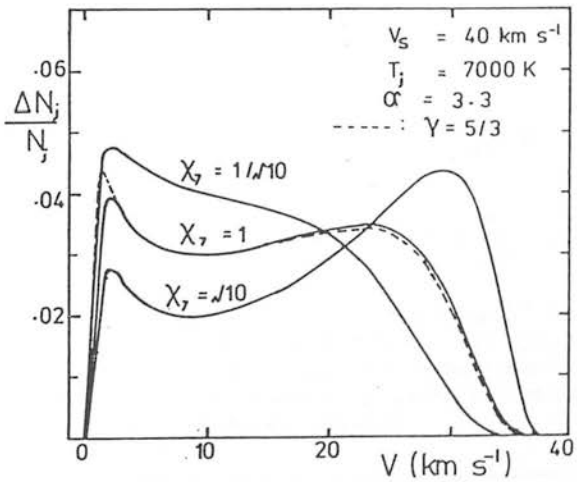


Figure 3. H_2 profiles from H_2 cooling for various ionization fractions keeping $n_6/B_{-3} = 1$.

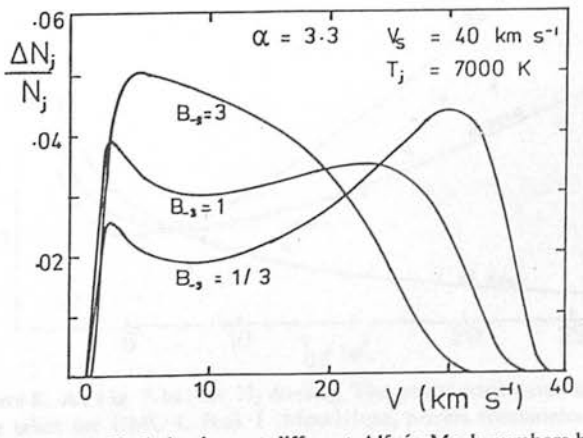


Figure 4. Profile behaviour at different Alfvén Mach numbers but the same shock velocity.

~ 10 km s^{-1} becomes a strong broad peak at $T_j > 12\,000$ K. The high-velocity peak disappears and the fall-off becomes approximately linear. Moreover, low-intensity wings become apparent. Thus we predict the lines to be narrower. From the previous discussion it is evident that the broad peak arises from the hot gas near T_{max} and the gas is too cold in the downstream region for emission from high energy levels.

3.6 H_2O results

Fig. 6 demonstrates that H_2O cooling has all the attributes of H_2 cooling except that the profiles are narrow. This is a crucial point since, although C-shocks have extended the velocity to which molecular hydrogen can be accelerated locally from ~ 25 to 45 km s^{-1}, the line widths are still less than 26 km s^{-1}. This is caused by the n^2 cooling of H_2O, as noted above, which rapidly reduces the temperature in the downstream region. This can be seen directly by comparing Figs 1 and 2 of Paper I.

4 LINE INTENSITIES AND RATIOS

Data on the line intensities due to rotational/vibration transitions of molecular hydrogen can be expressed in terms of the column densities of the upper energy levels when in LTE. A convenient means of displaying the columns in each upper energy level is to compare results with that of a slab at 2000 K. This is done in Figs 7 and 8 for H_2O and H_2, respectively. In both cases a wide range of line ratios are found with excitation temperatures from above $T_j = 6000$ K in the range 1750–3100 K being displayed in the diagrams. This is sufficient to cover the observed range of values from 1800 K to ~ 3100 K obtained by Burton *et al.* (1989) for four sources. The ratios can be interpreted by the C-shock model because of the free parameters involved. This wide range in line ratios is, however, difficult to reconcile with the measured values in the Orion molecular outflow. The excitation temperature derived from the 1 – 0 O(7) and 0 – 0 S(13) lines ($T_j = 8366$, 17 450) for the molecular outflow OMC-1 remains constant throughout the source (Brand *et al.* 1989b).

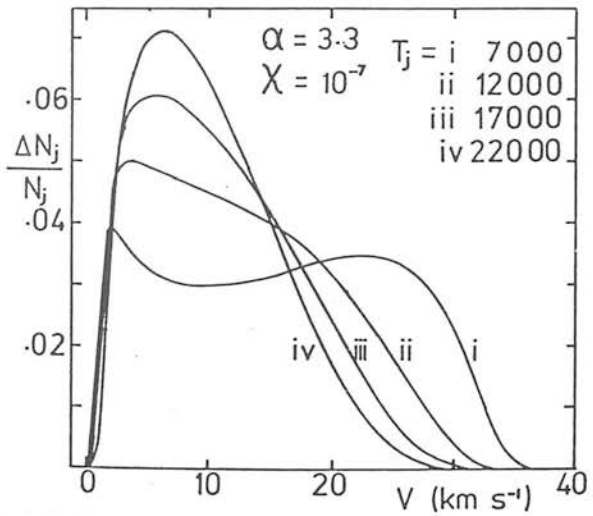


Figure 5. H_2 line profiles from a range of upper energy levels holding $n_6/B_{-3} = 1$.

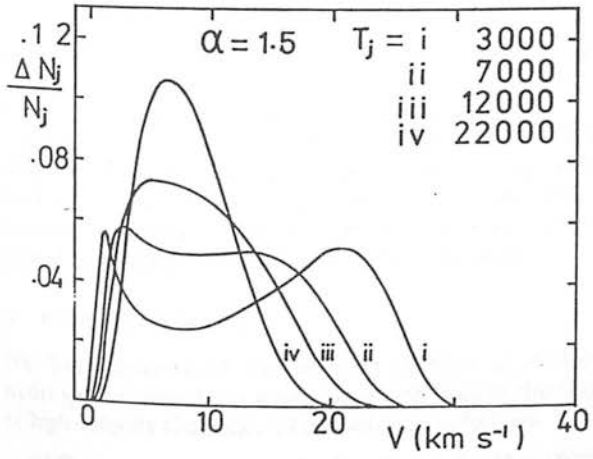


Figure 6. H_2 profiles from H_2O dominated cooling with $n_6/B_{-3} = 1$.

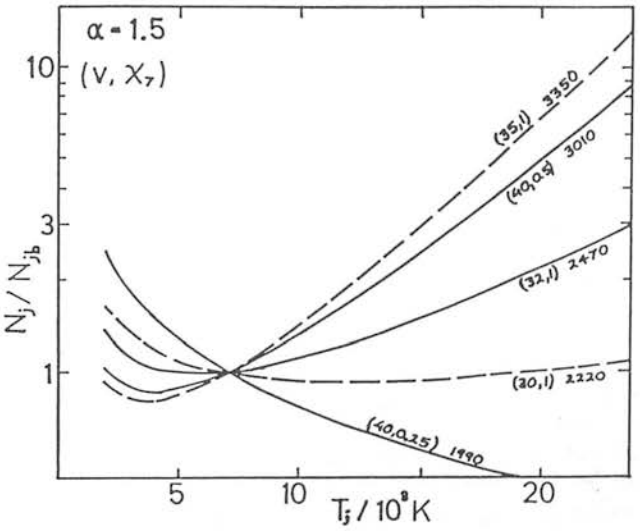


Figure 7. The column densities for energy levels T_j for H_2O cooling, normalized to a 2000 K slab of size such that N_j/N_{jB} at $T_j = 7000$ K. The numbers in brackets are the velocity in km s^{-1} and the ionization fraction is in units of 10^{-7}. The maximum temperature attained in a C-shock is also indicated.

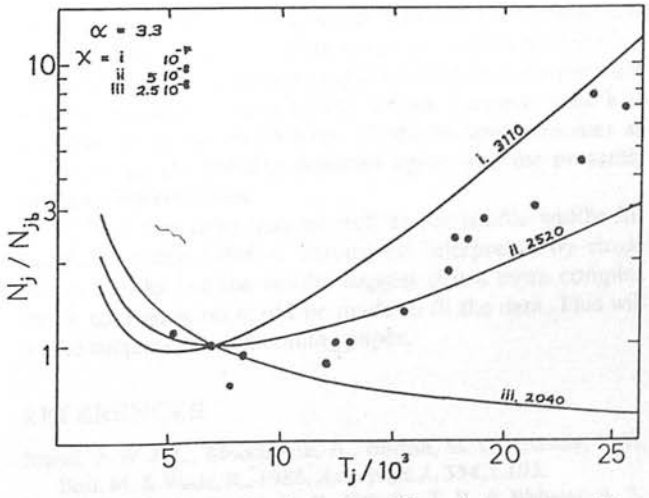


Figure 8. As Fig. 7 but for H_2 cooling. The points correspond to data taken for OMC-1, Peak 1 (Moorhouse, private communication).

A single C-shock cannot fit the data for peak 1 of OMC-1 as displayed in Fig. 8. It is evident that an appropriate combination of planar shocks would provide a fit and it will be interesting to determine if a bow shock provides this.

The predicted column densities are derived by taking $L_0 = 1/(n_{no}\sigma_0)$, giving $N_j/g_j = 44\Sigma C_c/(\chi\sigma_0)$. Note that the density is not explicitly involved, coming in only through χ, the ion fraction. For all the results presented here the column density at $T_j = 7000$ K remains in the narrow range $4.1-17\times10^{17}$ cm^{-2}, with the lower limit obtained from the 25 km s^{-1} example. This is sufficient to account for the columns in OMC-1 ($\leqslant 3\times10^{18}$ cm^{-2}, Brand *et al.* 1988) and NGC 2071 ($4-13\times10^{17}$ cm^{-2}, Burton *et al.* 1989a) as measured by the 1 − 0 S(1) line.

We now demonstrate an approximate method for calculating column densities which shows that the C-shock predictions are very stringent. We write $Q(T)\propto T$, accurate to better than 77 per cent since $T\leqslant 4000$ K. Equation (6) is approximated by two terms. First, from the peak in the temperature profile we substitute $T = T_{max}$ and integrate over a small interval Δr. Secondly, for the cooling downstream flow we use the fact that q is $\sim O(1/M_A)$ and take $e = 0$ (to avoid incomplete Gamma functions). We also write $(t/t_{max})^{\alpha} = r^{2+e+f}$ as an approximation to equation (17) of Paper I, valid for all but the upstream flow. Consistent with this, we ignore the contribution of the first narrow peak (where it is present). This yields

$$N_j/g_j = 44\eta n_{no} L_0 (T_{max}/T_j + \varepsilon)\exp(-T_j/T_{max})/(T_{max}M_A), \quad (8)$$

where ξ is of order 0.1 and η is of order unity. The first term in brackets represents the downstream cooling flow and the second term is the contribution from the peak. With $\xi = 0.1$, determined by fitting a single H_2 curve from Fig. 8, we find that equation (8) very accurately follows all the curves in Fig. 8 for $T_j > 4000$ K and is only a factor of < 1.3 too low at $T_j = 2500$ K. For the H_2O results the column densities relative to the 2000 K slab are also matched to within 10 per cent for $T_j > 5000$ K and to within a factor of 1.3 at $T_j = 2500$ K. Some solutions are shown in Fig. 9. With the definition of a local excitation temperature as $T_{exc} = dT_j/d(\log N_j)$, we find

$$1/T_{exc} = 1/T_{max} + (1 + \xi T_j/T_{max})/T_j. \quad (9)$$

The success of equation (8) indicates that line ratios are a function of the maximum shock temperature. A given line ratio can thus be interpreted by the ranges of individual parameters which yield this temperature. Fig. 9 can thus be used to find the maximum temperature and equation (42) or (43) of Paper I then yields the cloud parameters.

It is evident that a curve is rigidly defined by a single line ratio (i.e. the corresponding columns for two separate T_j levels). This leaves little scope for model fitting once the maximum temperature has thus been fixed. Line ratio data provide a stringent test for planar C-shock models.

5 CONCLUSIONS

We have determined the velocity profiles of molecular hydrogen rotational/vibrational emission lines in the context of high-velocity C-shocks. The results are as follows.

(i) First we assume the H_2O abundance of 4.25×10^{-4} per hydrogen atom (Draine *et al.* 1983) so that H_2O cooling dominates up to about 3500 K. The lines are very narrow, typically half of the shock velocity. At present it is not conceivable to resolve such narrow emission line profiles. A C-shock can still be used to accelerate molecular hydrogen to velocities ~ 45 km s^{-1} but multiple or curved shocks in the context of a dynamical model are necessary to produce wide H_2 lines.

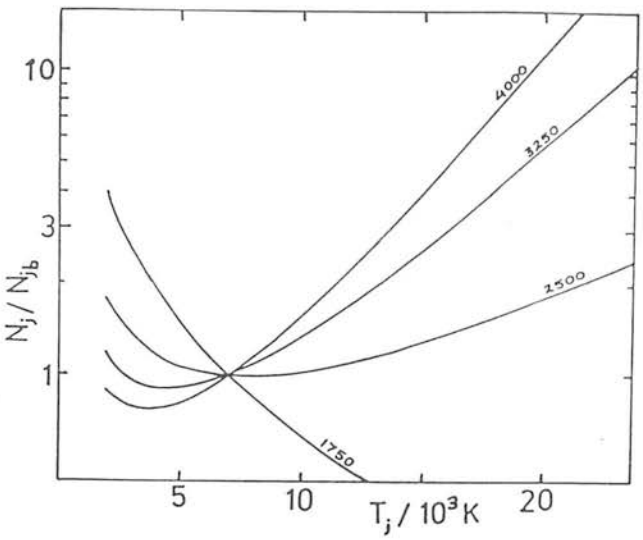


Figure 9. The simple model to predict column densities of the energy levels given the C-shock maximum temperature.

(ii) H_2 cooling produces wide emission lines, typically 90 per cent of the shock velocity, and are thus limited by the breakdown velocity to < 45 km s^{-1}. The lines from upper energy levels with $T_j < 7000$ K are double peaked with the peaks split by 30 km s^{-1} for a 40 km s^{-1} shock. The low-velocity peak is narrow and the high-velocity peak grows into a bump if the maximum shock temperature is increased by increasing the ionization or Alfvén Mach number. In contrast, the lines from $T_j > 12\,000$ K possess only a low-velocity peak and have a triangular appearance with an extended low-intensity wing. Such structure and structural differences may soon be measurable with some accuracy.

(iii) C-shocks produce an approximately constant excitation temperature from 7000 K upwards with a slow increase at high T_j. This temperature is closely related to the maximum temperature in the C-shock and can be anything less than ~ 3500 K. Thus a wide range of observed excitation temperatures can be interpreted. The excitation temperature will be expected to vary greatly across a source since it is sensitive to all the parameters. Predicted line intensities as measured by the column densities agree with the presently available observations.

(iv) The line ratio data as well as the profile widths for locations within OMC-1 cannot be interpreted by single planar shocks but the results suggest that a more complex shock configuration could be made to fit the data. This will be the subject of a forthcoming paper.

REFERENCES

Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. R., Bird, M. & Wade, R., 1988. *Astrophys. J.*, **334**, L103.

Brand, P. W. J. L., Toner, M. P., Geballe, T. R. & Webster, A. S., 1989a. *Mon. Not. R. astr. Soc.*, **237**, 1009.

Brand, P. W. J. L., Toner, M. P., Geballe, T. R. & Webster, A. S., Williams, P. M. & Burton, M. G., 1989b. *Mon. Not. R. astr. Soc.*, **236**, 929.

Burton, M. G., Geballe, T. R. & Brand, P. W. J. L., 1989a. *Mon. Not. R. astr. Soc.*, **238**, 1513.

Burton, M. G., Brand, P. W. J. L., Geballe, T. R. & Webster, A. S., 1989b. *Mon. Not. R. astr. Soc.*, **236**, 409.

Chernoff, D. F., 1987. *Astrophys. J.*, **312**, 143.

Chernoff, D. F. & McKee, C. F., 1989. In: *Molecular Astrophysics*, ed. Hartquist, T., Cambridge University Press, Cambridge, in press.

Chernoff, D. F., Hollenbach, D. J. & McKee, C. F., 1982. *Astrophys. J.*, **259**, L97.

Draine, B. T., 1980. *Astrophys. J.*, **241**, 1021.

Draine, B. T. & Roberge, W. G., 1982. *Astrophys. J.*, **259**, L91.

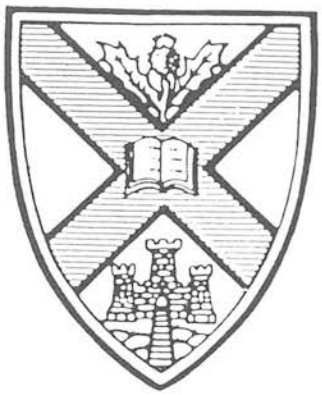
Draine, B. T., Roberge, W. G. & Dalgarno, A., 1983. *Astrophys. J.*, **264**, 485.

Kwan, J., 1977. *Astrophys. J.*, **216**, 713.

Nadeau, D. & Geballe, T. R., 1979. *Astrophys. J.*, **230**, L169.

Smith, M. D. & Brand, P. W. J. L., 1990. *Mon. Not. R. astr. Soc.*, **242**, 495 (Paper I).

Wardle, M. & Draine, B. T., 1987. *Astrophys. J.*, **321**, 321.



Edinburgh Astronomy Preprint

Number 3/90

February 1990

H_2 profiles of C-type bow shocks

M.D. Smith
P.W.J.L. Brand

University of Edinburgh Department of Astronomy
Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland
Telephone: 031-668 8100 Telex 72383

Accepted for publication in

Monthly Notices of the Royal Astronomical Society

H_2 PROFILES OF C-TYPE BOW SHOCKS

M. D. Smith & P. W. J. L. Brand

Department of Astronomy, University of Edinburgh, Royal Observatory,

Edinburgh EH9 3HJ, U.K.

SUMMARY

We present emission line profiles of molecular hydrogen from curved C-shocks within molecular clouds. Shock configurations arising from the supersonic motion of jets and bullets within a dense cloud are chosen. Bow shock speeds in the range v_w = 40 - 200 km s^{-1} are investigated. Breakdown through dissociation and self-ionization restricts the C-shock section to the bow tail. We find that profiles are essentially single-peaked and narrow with full widths (at 10% maximum intensity, deconvolved) of up to about 50, 40 and 30 km s^{-1} for cones, hemispherical caps and paraboloids, respectively. Exceptional field alignments can produce lines as wide as 75 km s^{-1} in the conical shock model. Line asymmetry is critically dependent on the orientation to the observer and, for low velocity bows, the magnetic field direction. The peak is never significantly shifted away from the radial component of the pre-shock velocity. These results are consistent with observed profiles from H_2 peaks associated with Herbig-Haro objects but *not* with the particularly broad H_2 lines in OMC-1 and DR21.

1 INTRODUCTION

Molecular hydrogen emission lines have been observed from collimated outflows associated with the early stages of stellar evolution (e.g. Zinnecker et al. 1989). The line ratios often implicate radiative shock waves as the means of heating this gas. The line widths, however, can far exceed those predicted by hydrodynamic shocks. Another type of radiative magnetohydrodynamic wave called a C-shock has an advantage in this respect (Draine 1980). A C-shock can accelerate molecular hydrogen to velocities approaching 50 km s^{-1}, whereas a standard J-shock is limited by molecular dissociation to 25 km s^{-1}. However, from the use of an approximate model (Smith & Brand 1989a, Paper 1) and the predictions for planar shocks (Smith & Brand 1989b, Paper 2) we found that the predicted profiles are still far too narrow to explain some of the intense H_2 sources (up to 140 km s^{-1} wide in OMC-1). Moreover the predicted line intensities are not consistent with the data. Nevertheless it was argued that a suitable combination of planar shocks could produce (i) line ratios to match present data and (ii) wider profiles. Here, in Paper 3, we concentrate on the latter problem and return to the line ratio question in a subsequent work (Smith & Brand 1989c, Paper 4). Are C-type bow shocks, driven by jets, bullets or clouds, responsible for the observed wide H_2 emission lines?

Bow-shock structures are often observed in optical emission lines. Recently it has been shown that H_2 line emission is associated with the optical emission and has a velocity structure with some similarities (Zinnecker et al. 1989, ZMGZ). Hence it has become urgent to calculate the H_2 properties of C-type bow-shocks. The progression from planar to bow shocks has been crucial to the interpretation of many Herbig-Haro (HH) objects (Hartigan 1988). H_2 emission is also apparent from some HH objects within jets and ZMGZ conclude that oblique shocks associated with a jet/ambient medium interaction are also important. The simplest shock front

configuration either for the stimulation of shocks in the cloud external to the jet or for co-moving internal molecular gas is that of a cone (e.g. Fig 7b of ZMGZ).

To find the H_2 emission properties we have first had to overcome a number of difficulties, as follows.

(1) For non-planar shocks we must apply detailed constraints to account for dissociation, self-ionization and the transition to a J-shock. As shown in Paper 1 we cannot take a fixed breakdown velocity of, say, 40 km s^{-1}. In the C-shock model the breakdown velocity is a function of density, ionization level, H_2O abundance, magnetic field strength and magnetic field direction. Here we take the pre-shock medium to be uniform. The breakdown condition is still a function of the angle between the field and the shock normal and so must be determined for each location on the bow surface.

(2) The magnetic field, also assumed uniform, will generally be oblique to the shock front. Transfering into the frame in which the electric field vanishes results in a modified set of equations (Wardle & Draine 1987). Wardle & Draine find that the transverse field component can be taken as $B_\perp = B(\sin \varsigma)^{1/2}$ where ς is the angle between the field and the shock normal. They also note that the dust grains apply a drag force on the neutrals transverse to the field-shock normal plane. This force is ignored here (requiring either a moderate ionized fraction (i.e. $> 10^{-7}$) to dominate the drag or a tight coupling of the grains to the field or neutrals).

(3) A network of physical and chemical processes have been included in C-shock calculations (Draine, Roberge & Dalgarno 1983). The neutrals and the ions/field are treated as weakly interacting interpenetrating fluids, a condition well-satisfied when the ionisation level is low. In Paper 1 we isolated the most important mechanisms and expressed their effect in simple forms. A summary of these results and the method employed to calculate H_2 profiles from a planar element was presented in Paper 2.

(4) Bow shocks are often approximated as parabolic and some HH objects do appear to have roughly this shape. More complicated shapes for HH objects are given by Choe, Böhm & Solf (1985), and these have been used in a simplified form by Hartigan, Raymond & Hartmann (1987). They are close to parabolic at the leading edge but bend faster in the tail. A hemispherical cap is introduced as a contrasting alternative, since the shock cross-section is small with a finite length tail. In theory, a shock shape depends on the (time-dependent) shape of the obstacle as well as the degree of compression and rapidity of cooling in the flow around it.

(5) Multi-dimensional fluid simulations are necessary if the cooling length for the warm H_2 is comparable to or greater than the obstacle size. Our numerical and analytical calculations presume that the gas has cooled locally. We must ensure that this planar approximation is valid at each location on the bow shock (sub-section 2.4).

(6) The calculation of H_2 line strengths when the populations are not in local thermodynamic equilibrium (LTE) is beyond the scope of the present work. Here we emphasize the flow properties and assume LTE, appropriate for densities above 10^6 cm^{-3}. Nevertheless we expect our general conclusions will hold down to possibly 10^4 cm^{-3} for the 1-0 S(1) line (Hollenbach & McKee 1979) since the breakdown velocities and the shock temperature structure remain quite similar to the high density case (Paper 1). H_2 and H_2O are considered. When taken separately the numerical results can be modelled by analytical approximations which allow us to explore the conditions for the existence of C-shocks (Paper 1) and understand otherwise hidden relationships in the numerical results (Paper 2). Here we shall explore the complete cooling function including dissociative cooling.

(7) Copious UV radiation is emitted from the front dissociative region of fast bow shocks - the speckled region in Fig. 1. This radiation could, in principle, radically alter the ionization state in the bow tail. The C-shock section in the tail may thus be converted into a J-shock. In sub-section 2.5

we determine the bow conditions for which C-shocks remain. We find that a substantial depletion of the trace metals with low ionization potentials is required. Such depletions are typical in clouds (Langer 1985).

We begin by outlining the method for calculating line profiles. Numerical and analytical details follow in Sections 3 and 4. A summary of the results and a comparison with observations are contained in the final two sections.

2 THE MODEL

2.1 Geometry

The parameters defining the bow shock, observer and magnetic field orientations are displayed in Fig. 1. Profiles are expected to be asymmetric through the latter two directions. For this reason numerical calculations are quite lengthy. For each run the angles Ψ, Π and Θ are fixed. The (x, y, z) coordinate system is used below.

The local angle of incidence of the flow to the shock surface is defined as α. The numerical calculations proceed by first dividing a bow shock into annuli of constant $d\alpha$. For each annulus an initial normal shock velocity $v_\perp$ is determined. An annulus is then divided into segments of azimuthal angle Φ. Each segment must be treated as an individual planar shock element according to the particular magnetic field orientation relative to the shock normal. The segment contribution to the line profile is binned into intervals of 1 km s^{-1} (with the constraint that the velocity parallel to the shock front, $v_\parallel$, is constant) in the observer's frame.

Taking a single shock element, the pre-shock flow has velocity $v_w(i_t \cos \alpha + i_n \sin \alpha)$ where $i_t = (\sin \alpha \cos \Phi, \sin \alpha \sin \Phi, \cos \alpha)$ and $i_n = (-\cos \alpha$

cos Φ, $-\cos\alpha \sin\Phi$, $\sin\alpha$). The neutral velocity vector in the C-shock is written as $v = v_w(i_t\cos\alpha + i_n r \sin\alpha)$. Here r proceeds from unity to a value near zero where the magnetic pressure inhibits further compression. The manner in which the temperature and density are determined as functions of r was presented in Paper 2.

The magnetic field component perpendicular to the shock is $|B \times i_n|$ where $B = B$ (sin Ψ, cos Ψ cos Π, cos Ψ sin Π). An 'effective' field component appropriate for C-shocks is approximated by $|B \times i_n|^{1-\tau}$, where Wardle & Draine (1987) suggest $\tau \sim 0.5$. We shall also consider $\tau = 0$ to determine the sensitivity to this parameter. The observer sees the emission at a velocity $i_\theta.v$, where i_θ = (sin θ, 0, cos θ).

2.2 Weighting and line strengths

H_2 luminosities are derived from the column densities of the upper energy levels when in LTE, as assumed here. The method to find the column density within a given velocity interval for a planar shock is given in Paper 2 (equations 4 - 6). Particularly important here is the dependence on the normal velocity, effective field component and cooling function. The pre-shock values of density n_o, ionisation fraction χ and field B are constants for each bow shock considered i.e. a homogeneous medium with a uniform magnetic field. The contribution from an element is weighted by a factor proportional to its area: $(R\, d\Phi)(dz/\cos\alpha)$.

A paraboloid is described by $z = R^2/(2L_p)$, where L_p is a constant and $dR/dz = \tan\alpha$. It follows that an element $d\alpha d\Phi$ has area $dA = L_p{}^2\cos\alpha/(\sin\alpha)^4 d\alpha d\Phi$. For a hemispherical cap we have $R^2 + z^2 = L_s{}^2$ and $dR/dz = \tan\alpha$. This yields $dA = L_s{}^2\cos\alpha\, d\alpha d\Phi$. The results for a cone are directly extractable from an individual annulus of a bow shock.

2.3 Cut-offs

Dissociation and ionization of H_2 provide strong upper limits to the normal shock velocities and so determine the locations at which H_2 emission occurs. A precise limit on the breakdown velocity does not in reality occur - there is a transition region where some H_2 emission still arises from the upstream/pre-curser region of a planar shock which is completely dissociated far downstream. Fortunately it is meaningful to define a breakdown velocity since the breakdown conditions are very sensitive to the velocity (Paper 1). For this reason these transitional velocities are ignored. Nevertheless it should be noted that the theoretical breakdown speeds are only definable to within a few km s^{-1}.

For dissociation we derive a formula from paper 1 to limit the velocity:

$$v_\perp < v_d \left[\frac{B}{10^{-3}G}\right]^{1/4}\left[\frac{10^6 cm^{-3}}{n}\right]^{3/8}\left[\frac{10^{-7}}{X}\right]^{1/4}\left[\frac{(n/10^6 cm^{-3})\epsilon(H_2O)}{4.25\ 10^{-4}}\right]^{p/4}. \qquad (1)$$

where p = 1 for H_2O dominant cooling (i.e. $n > 10^7$ cm^{-3} or H_2O abundance $\epsilon > 4\ 10^{-4}$) and p = 0 for LTE H_2 cooling. We take v_d = 40 km s^{-1} & 45 km s^{-1} for these cooling mechanisms, respectively, but also consider the sensitivity to this value (Section 3). Ionization fractions are kept well below 10^{-6} in this study, and so the J-shock transition limit is not encountered here (see Paper 1).

Equation (1) indicates that thermal collisions with neutrals does not dissociate H_2 even for velocities somewhat in excess of 50 km s^{-1} provided certain conditions are met, such as a high magnetic field. We then employ the ionization streaming limit to dissociation of 50 km s^{-1} (Paper 1).

2.4 The cooling length

If the cooling length L_c is much smaller than the radius of curvature of the bow shock, then each surface element can be treated as a planar shock. In Paper 1 (Section 4.2) we found that $L_c \sim 2/(\sigma_o n_o X M_A)$ where $\sigma_o = 10^{-15} cm^{-2}$ and M_A is the Alfvén Mach number. It follows that

$$L_c \sim 8\ 10^{14} \left[\frac{B}{10^{-3}G}\right]\left[\frac{10^6 cm^{-3}}{n}\right]^{3/2}\left[\frac{10^{-7}}{X}\right]\left[\frac{40\ km\ s^{-1}}{v}\right] cm. \qquad (2)$$

The typical size of observed HH bow structures are a few x 10^{16} cm. These are sufficiently large for the planar-element treatment to be valid for the canonical values but would not be for much higher fields or smaller bow-shock dimensions.

2.5 UV ionization

The effects of self-generated UV radiation on atomic gas in bow shocks have been explored by Raymond et al (1988). High preionization levels can be caused by the penetration of photons with energies exceeding 13.6eV. If the head of the bow shock is J-type and is producing such photons, a column density of $N_H \gtrsim 10^{18}\ cm^{-2}$ will shield the rest of the bow - presumed to be C-type - from the UV radiation. In these cicumstances, the parts of the J-type shock can be treated as planar elements. Hollenbach and McKee (1989) have shown for such shocks that the UV radiation upstream is totally quenched by incoming neutrals if $v_\perp < 120\ km\ s^{-1}$.

For hemispherical caps with $v_w > 160\ km\ s^{-1}$, however, the number of ionizing photons per arriving atom exceeds unity over a large fraction of the flow, and preionization occurs, forming a small HII region around the

J-shock. We estimate that recombinations within the HII region are sufficient to shield the UV from the C-shock region for $N_o > 3\ 10^{19}\varsigma\ cm^{-2}$ where ς is the fractional excess of ionizing photons per atom over the J-shock section. Given that the obstacle (bullet or cloud) can only shield the tail from radiation emitted downstream, there appears to be a narrow range of bows with dimensions near $10^{19}/n$ cm for which C-shocks are not permitted. Moreover these bows must have high velocities and possess cap-like morphologies.

The photoionization of trace metals with low ionization potentials such as Mg, Si and Na results from the highly-penetrating UV radiation below 13.6eV. The total observed abundance of such elements in quiescent clouds is $\lesssim 10^{-6}$ (Langer 1985), implying substantial depletion onto grains. If the depletion also exists in the environments discussed here then a C-shock remains viable even if all these metals are ionized. The ionization fraction is consistent with the C-shock criteria (i.e. $\chi < 10^{-5}$) found in Paper 1 although the breakdown velocity (and, hence, line widths) will be reduced (see equation 1).

The dissociated part of the bow will radiate large amounts of soft UV capable of complete ionization of the ionizable species. If they are *not* locked up in the grains the ionization level reaches ~ 10^{-4} and C-shocks disappear (with the possible exception of $n < 100\ cm^{-3}$ (Draine, Roberge & Dalgarno 1983)). Furthermore, this abundance implies a similar photoionization cross-section and dust absorption cross-section per hydrogen nucleus of ~ $10^{-21} cm^2$ (e.g. Draine, Roberge & Dalgarno 1983). Therefore dust is capable of shielding the metals only when $N_o > 10^{22}\ cm^{-2}$ and the metals are depleted. A very low ionization fraction of order 10^{-7} is possible when both conditions are satisfied. We conclude that the depletion of the ionizable metals is necessary to the present H_2 bow shock model.

3 RESULTS

We take the four combinations of either H_2 or H_2O dominating the cooling with either a paraboloidal or a hemispherical (cap) shock. We take velocities between 40 and 200 km s^{-1} in steps of 40 km s^{-1}. We assume the ambient gas to be stationary and the shock configuration to be moving with velocity $-v_w$ for display purposes. The full widths given in the tables are at 10% maximum intensity after the profile has been smoothed to 3 km s^{-1}. All figures are displayed with this resolution. The line shapes and strengths have a weak dependence on the specific heat ratio γ (Paper 2). We fix $\gamma = 7/5$. The profiles are normalised to the percentage of the total intensity per km s^{-1} (i.e. constant area) unless otherwise indicated.

3.1 Line widths

Tables 1 to 3 display a range in line widths of up to 41 km s^{-1}. The widest lines are generally produced by 1) bow shock motions $\geq$ 80 km s^{-1}, 2) H_2 cooling and 3) hemispherical caps (Fig. 2). Lines wider than 31 km s^{-1} are produced by caps with H_2 cooling.

A line width changes with the observer's angle (Tables 2 & 3). For v_w = 160 km s^{-1} the lines are widest when moving in the plane of the sky (θ = 90°, see Fig. 3). In contrast, for caps with v_w = 40 km s^{-1} the lines are widest when the motion is in the observer's line of sight (θ = 0°).

The magnetic field direction is chosen as ψ = 90° in Tables 1-3. A change to the field direction within the x-y plane results in a small change to a line profile (Table 4 and Fig. 4). Generally, the field direction is expected to have little bearing on the line width at high velocities (Fig. 5) since in such cases the emission comes from the highly oblique tail section of the shock configuration (dissociation occurs over the leading edge).

The effective magnetic field coefficient τ also has a limited effect on the results (Fig. 4). Unless stated we have kept $\tau = 0.5$. An increase in the breakdown velocity produces a similar increase in the line width (Table 4). Hence the uncertainty in the dissociation limit will have an insignificant bearing on our comparison with observed profiles (see also Fig. 5).

3.2 Line shapes, shifts and asymmetries

The line shapes are simple in comparison to the range of shapes found for planar shocks described in Paper 2. They are single peaked apart from an insignificant depression near 0 km s^{-1} in some examples. The lines have very low level extended wings which contain at most 5% of the total line luminosity outside the width defined by the 10% peak intensity level.

Line asymmetry is predominantly caused by the inclination of the flow to the observer. Large asymmetry is found when the motion is close to the line of sight, although the line is then generally narrow (Fig. 3, Tables 2 & 3) and the asymmetry difficult to detect. In Fig. 3 the lines have been shifted so that the pre-shock gas is stationary in the observer's frame. As explained in Section 4, profiles are dominated by the warm gas which is still moving close to the pre-shock velocity. Thus the asymmetric lines have large shifts if the bow shock is caused by a stationary obstacle: i.e. by ~ $v_w \cos \theta$, and, on the other hand, the asymmetric lines are not shifted in the bullet model.

Small asymmetries result from motion at ~ 30° to the plane of the sky, an asymmetric magnetic field, non-uniform parameter distributions or non-axially symmetric bow shocks. Therefore small observed asymmetries will be difficult to interpret. A range in field directions can produce high asymmetry. This will be discussed in Sub-section 3.4

Lines from higher upper energy levels are broader for paraboloids (Fig.

6) but narrower for caps (Fig. 7). This is due to the long tail of a paraboloid contributing greatly to the peak for low excitation lines; the tail is too cold to contribute to high excitation lines. A cap does not possess such a tail and, for the same reason as a planar shock (the hottest gas has only been marginally decelerated), produces a narrower line at a higher excitation energy. Note that, in contrast to planar shocks (Paper 2), the profile shapes are quite similar especially for high excitation lines. This is because the profile from each planar element is narrow and peaked near the pre-shock velocity and so the shock configuration alone begins to determine the profile shape.

3.3 Cones

We have investigated conical shocks with a uniform pre-shock field. Wider lines are possible (Table 5) - up to 75 km s^{-1} full width at zero intensity. But this can only be achieved for a small range of cone angles i.e. the Mach angle must be near to the maximum permitted value before molecular hydrogen dissociates (Fig. 8). The dissociated fraction actually depends on Φ through the magnetic field orientation, as is evident in equation (1), and so does not cut in abruptly at a precise cone angle α i.e. α_D is a weak function of Φ. This makes possible wider profiles when $\alpha \sim \alpha_D$ and $(\Psi, \eta) = (0, 0)$ for which cone emission is restricted to the locations of maximum and minimum radial velocity at $\Phi = 0^\circ$ and 180°. This effect is not significant for paraboloids or caps since then any particular annulus does not dominate the total intensity (e.g. Table 4). There is no inherent reason for the cone Mach angle to be at the most favourable for producing wide lines. A selection effect may operate: the hottest, most luminous, H_2 locations are also the widest.

Line widths and asymmetries are strong functions of the viewing angle

(Fig. 9). Note that for low velocities, v_w = 40 km s^{-1}, the lines are broadest when the cone is moving along the line of sight (θ = 0°). This case approaches that of a planar shock i.e. $\alpha \rightarrow 90°$. The width exceeds the value at θ = 90° (a local maximum - see Section 4) of 34 and 23 km s^{-1} for H_2 and H_2O cooling which are obtained for a Mach angle of 45°.

3.4 The total cooling function and line asymmetry

We add together H_2, H_2O and dissociative cooling:

$$\Gamma = 3.3\ 10^{-15} n_6 [T_3^{3.3} + 3.5\{x(H_2O)/4.25\ 10^{-4}\} n_6 T^{1.5} + 1.2\ 10^{7} n_6 \exp(-52.7/T_3)] \quad (3)$$

where T_3 = T/1000K (Paper 1).

The line properties are quite similar to those of H_2O taken alone (e.g. Table 1 and Fig. 2). This is true for all the configurations except cones with v_w = 80 - 160 km s^{-1}, where the line widths are intermediate between the H_2 and H_2O results (Table 5).

We have employed this case to study oblique magnetic fields in detail. High asymmetry can arise when the field is perpendicular or near to perpendicular over one large section of the configuration but not highly inclined over the opposite section (Fig. 9). These conditions are most likely in a slow moving cone. The asymmetry is present even if the motion is in the plane of the sky. To illustrate this we take a cone with v_w = 80 km s^{-1} which gives a breakdown angle of 38.7° for v_d = 50 km s^{-1}. For a cone Mach angle of 38° the magnetic field direction is varied. For an observer at θ = 90° the maximum asymmetry occurs for the field angle η = 90° since

then the difference in the field-shock angle between the locations of highest and lowest radial velocities (i.e. at Φ = 0° and 180° is greatest. Where the shock normal and field direction coincide there is no contribution, consistent with the dissociation condition.

The asymmetry results can be summarised by using the asymmetry index AI20 (Heckman et al. 1981). This is the difference in half width to the blue and red of the peak divided by the total width, all measured at 20% of the peak intensity. We find AI20 > 0.23 for 30° < Ψ < 65° and AI20 > 0.3 for 35° < Ψ < 50° with θ = 90° and η = 90°. Fixing Ψ = 45° we find AI20 > 0.23 for η > 30°. From these ranges we estimate that of order 20% of field directions produce AI20 > 0.23 even for motions in the plane of the sky. This of course assumes there is no preferential field direction. At other orientations strong asymmetries can be found (Fig. 10) with AI20 = 0.5 and 0.84 for θ = 45° and 135° respectively, on taking Ψ = 45° and η = 0°.

4 CALCULATION OF PROFILE WIDTHS AND SHAPES

4.1 Widths

One can calculate theoretical upper limits to profile widths in terms of the viewing angle and the shock velocity. This analysis is more complex than that for J-type atomic bow shocks. There, the most commonly observed lines in the shocked gas are mainly produced only after cooling by a factor of ~ 40 (Hartigan, Raymond & Hartmann 1987, hereafter HRH). This allows the emitting velocity to be set at $v_{\parallel}$. The full width of an emission line is then shown to be simply v_w, independent of orientation. In C-shocks, however, the H_2 lines peak when $v_{\perp}$ is still close to its pre-shock value and, moreover, an individual planar element can radiate greatly over a velocity interval exceeding 30 km s^{-1} (Paper 2).

4.1.1 Cones

The first task is to obtain an upper limit to a profile width from a cone or single annulus at an angle α to the axis, observed at an angle θ. We need only consider the top ($\Phi = 0^\circ$) and bottom ($\Phi = 180^\circ$) since these two locations will tally with the extremes in radial velocity. The planar-element emission is assumed to occur at all normal velocities ($rv_\perp$ with $0 < r < 1$) below the initial value $v_\perp = v_w \sin \alpha$. This component is limited to the breakdown velocity v_d which thus becomes a condition on α: $\alpha < \alpha_o = \sin^{-1}(v_d/v_w)$. The 'breakdown angle' α_o is so defined and will recur throughout this section.

First suppose that the observer is outside the cone defined by α: $\theta > \alpha$. Then the maximum blueshifted velocity, v_+, is found by putting $v_\perp = 0$: $v_+ = v_\parallel \cos(\theta - \alpha) = v_w \cos\alpha \cos(\theta - \alpha)$. The maximum redshifted value is similarly $v_- = -v_w \cos\alpha \cos(\theta + \alpha)$. The velocity width can be written:

$$v_+ + v_- = v_w \sin\theta \sin 2\alpha. \qquad (4)$$

The maximum width, v_{max}, clearly occurs at $\alpha = \alpha_o$ or $\alpha = \pi/4$, whichever is smaller. Therefore this maximum width occurs at $\alpha = 45^\circ$ for $v_w < v_d\sqrt{2}$ as found for the 40 km s^{-1} shock (Table 4). Thus

$$v_{max} = 2v_d \sin\theta \cos\alpha_o \quad \text{for } v_w > v_d\sqrt{2}, \qquad (5a)$$

$$v_{max} = v_w \sin\theta \quad \text{for } v_w < v_d\sqrt{2}. \qquad (5b)$$

These formulae satisfactorily explain the strong dependence on the observer's orientation for cones, with a maximum width at $\theta = 90^\circ$. The numerical values at 0.1% derived in Sub-section 3.3 reach 90% of the theoretical maximum width. The highest value of $2v_d$ is found for $v_w \gg v_d$. Taking $v_d = 45$ km s^{-1} then yields $v_{max} = 90$ km s^{-1}.

If the observer is inside the cone ($\alpha > \theta$), the maximum blueshift is $v_+ = v_w \cos\theta$ since $v_\perp$ is now also blueshifted. The maximum redshift is unchanged. The total width is $v_+ + v_- = v_w \sin\alpha (\sin\alpha \cos\theta + \cos\alpha \sin\theta) = v_w \sin\alpha \sin(\alpha+\theta)$. For α fixed, this has a maximum value at $\theta + \alpha = 90^\circ$. Since we require $\alpha > \theta$ this maximum is only reached for $\alpha > 45^\circ$ i.e. for $v_w < \sqrt{2}v_d$. The full width is then $v_{max} = v_w \cos\theta$. For $v_w < v_d$ α can approach 90° and the maximum width occurs for $\theta = 0^\circ$, as found in Section 3.3. This is of course just a planar shock. For $\alpha < 45^\circ$ the width monotonically increases with θ, reaching the value $v_w \sin\alpha \sin 2\alpha$ when $\theta = \alpha$. This then matches onto equation (2) for larger θ, and so only the single local maximum at $\theta = 90^\circ$ is present.

4.1.2 Bow shocks

The case of a bow shock in which $0 < \alpha < \alpha_o$ is more complex but the foundations have been laid in the above analysis. Now one must maximize v_+ and v_- separately since these may occur at different values of α. The maximum blueshift is now found to be $v_b = v_w(1 + \cos\theta)/2$ for $\theta < 2\alpha_o$ and is from the annulus with $\alpha = \theta/2$. For $\theta > 2\alpha_o$, $v_b = v_w \cos\alpha_o \cos(\theta-\alpha_o)$ and is from the annulus α_o. These values were found to be dominated by the $\theta > \alpha$ case of Sub-section 4.1.1 - we can always find an α less than θ.

The maximum redshift is $v_r = v_w(1 - \cos\theta)/2$ for $\theta > \pi - 2\alpha_o$ and is $v_r = -v_w \cos\alpha_o \cos(\theta+\alpha_o)$ for $\theta < \pi - 2\alpha_o$.

For $v_w < \sqrt{2}v_d$ ($\alpha_o > 45^\circ$) the total width v_t is simply v_w for viewing

angles $\theta > \pi - 2\alpha_o$. For $\theta < \pi - 2\alpha_o$, $v_t = v_w[1 - \cos(2\alpha+\theta)]/2$. Hence the line width is always narrower than $\sqrt{2}v_d$. The 40 km s^{-1} example corresponds to $\alpha_o = 90^\circ$ and so we expect the full width to be 40 km s^{-1} independent of the observer's orientation. We found the width at 10% peak intensity to actually vary with θ (Table 2) since the true width of a planar line profile is ~ 90% and 50% of v_w for H_2 and H_2O cooling respectively. Thus the width of H_2O cases never exceeds the equivalent H_2 case. Furthermore, the distribution of line luminosity with α depends on the bow shape, the paraboloids producing a narrower, peakier profile due to the long low α tail.

All the other examples taken in Section 3 are of high velocity bows i.e. $\alpha_o < 45^\circ$. Then there are three categories: Case I $\alpha_o \leqslant \theta/2$ and $\alpha_o \leqslant 90^\circ - \theta/2$; Case II $\theta/2 \leqslant \alpha_o \leqslant 90^\circ - \theta/2$ or CaseIII $90^\circ - \theta/2 \leqslant \alpha_o \leqslant \theta/2$. Case III corresponds to $\theta \geqslant 90^\circ$ - the same lines of sight as in case II with just the direction reversed. Using the values of v_+ and v_-, it is straightforward to calculate the line width as a function of the viewing angle:

Case I: $v_t = 2v_d$ (cos α_o sin θ) for $\theta > 2\alpha_o$ with a maximum width $2v_d$ cos α_o at $\theta = 90^\circ$.

Case II: $v_t = v_w$ [1 - cos(θ + $2\alpha_o$)]/2 for $\theta < 2\alpha_o$. The maximum value is $4v_d(v_d/v_w)$ $\cos^2\alpha_o$ at $\theta = 2\alpha_o$ and corresponds to the minimum value of case I.

So the maximum width always occurs at $\theta = 90^\circ$ and has the value $2v_d$ [1 - (v_d/v_w)]$^{1/2}$. This is consistent with the numerical results in all cases with $v_w \geqslant 80$ km s^{-1} (e.g. Table 3). The line widths at 10%, however, do not even reach v_d. The widths at zero intensity reach 61 and 65 km s^{-1} for H_2 cooling with $v_w = 80$ and 160 km s^{-1}, respectively, which are 82% and 75% of the theoretical maxima. These values are consistent since the theory

presumes that a planar C-shock emits as $v_\perp \to 0$ whereas the true width, corresponding to the warm H_2 gas region, is at most 90% of $v_\perp$ (Paper 2). For H_2O cooling the planar lines are narrow due to the rapid n^2 cooling in the downstream flow and the warm H_2 region corresponds to less than 50% of $v_\perp$ - hence the much smaller widths consistently present in Table 1.

4.2 Line shapes

The widest lines found (at 10% of the peak intensity) correspond to examples of caps rather than paraboloids. This is interpreted as follows. The emission from a paraboloid is (relatively) strongly weighted to the tail (i.e. to low α). The width from an annulus is proportional to sin 2α (equation 4). Hence the tail adds a narrow peak which should reduce the line width as measured at some fixed level of the peak intensity.

The results are in contrast to the profiles for atomic lines in gas at 10^4K where the cooled deflected gas determines the profile shape (HRH). The profiles for HH-objects are very broad and wing-less, often with two peaks of varying intensity according to the orientation. The difference with our results is that the emission is here dominated by the warm gas which is still moving close to the initial velocity v_w at all locations of the surface. This leads to a single dominating peak and the same qualitative appearance of the profiles for wide ranges of all the variables.

5 COMPARISON WITH OBSERVED PROFILES

Many Herbig-Haro objects have been successfully observed in the 1-0 S(1) line at 2.12μm (upper energy level 6956K) and can be compared to the model results for T_j = 7000K. A summary of previous observations, along

with much new high resolution data, has been recently published (ZMGZ). There is good spatial coincidence between the optical and H_2 line emission with the optical emission often distributed in a bow-shaped structure (e.g. HH43, Schwartz et al. 1988). The displacement in several examples is only a few arcseconds and suggests that the optical lines are produced at or near to the apex of a bow shock and the H_2 from the tail/wings. The spatial distributions will be examined in detail in a following study. Here we note that the H_2 emission arises from tail angles relative to the direction of motion of $\alpha < \alpha_o = \sin^{-1}(v_d/v_w)$ where v_d is the breakdown velocity and v_w is the speed of the configuration.

The observed FWHMs of H_2 S(1) for 10 out of 11 HH objects are in the range 35 - 50 km s^{-1}. ZMGZ note that the true widths are in the range 10 - 25 km s^{-1}, depending on the instrumental and intrinsic profile shapes. The instrumental profile is 30 - 35 km s^{-1} full resolution. For comparison we have convolved model profiles with a Gaussian profile $\exp\{-(\Delta v)^2/(2\sigma^2)\}$. Fig. 11 displays convolved profiles with σ = 15 km s^{-1}. Intrinsic FWHMs of 7 and 12 km s^{-1} after convolution are 36 and 42 km s^{-1} respectively. Width measurements of the S(1) line in Cepheus A and NGC 2071 are in the range 27 - 61 km s^{-1} (FWHM) (Doyon & Nadeau 1988; Burton, Geballe & Brand 1989). Hence line widths are generally consistent with the present model.

The bow model does not predict the line widths from locations within three bipolar outflows. HH 11 has a FWHM of ~ 110 km s^{-1} (ZMGZ, but note the poor signal to noise). Various positions in OMC-1 have full widths exceeding 100 km s^{-1} (Burton et al 1988) including peak 1 with a width of 140 km s^{-1} (12 km s^{-1} resolution, Brand et al. 1989). The HH11 line can be split into at least four discrete components but the smooth OMC-1 profile is not so easily decomposed. Brand et al have shown that at least 30 narrow components would be required. Locations in the west lobe of DR21 possess convolved line widths of up to ~ 80 km s^{-1} (FWHM) and ~ 150 km s^{-1} at 10% peak intensity (Garden 1987, Nadeau & Beland 1988).

The bow C-shocks predict the H_2 peak, V_2, to lie very close to the pre-shock radial velocity. The bow models of HRH predict the Hα peak, V_1, (or the main peak if there are two) to also occur at this velocity for v_w = 200 km s^{-1} but to be shifted by roughly $v_w \cos \theta$ for v_w = 100 km s^{-1}. This is due to the main emitting region shifting from the cooled apex gas to the cooled wing gas as v_w is increased between these two velocities. We first assume that the optical lines and the H_2 lines arise from the same shock. Then v_w is also known - simply the Hα velocity full width at zero intensity (HRH). Given so much information we can restrict the possibilities for interpreting Herbig-Haro object emission.

HH 32A can be modelled by two components which appear both in the optical and H_2 lines. A one component model is ruled out by the high velocity of the H_2 emission (ZMGZ). For the low velocity component $V_1 - V_2$ ~ +50 km s^{-1} with an Hα width of ~ 140 - 160 km s^{-1} (HRH). This is consistent with the present model for a bow shock being driven into the ambient cloud at an angle to the line of sight of 20° - 50° (the peak position being intermediate between the values expected for v_w of 100 and 200 km s^{-1}). The proper motion of HH 32A is ~ 50 km s^{-1}. Given v_w = 140 -160 km s^{-1} this yields θ ~ 18° - 21°. The two peaks of the high velocity component roughly coincide (ZMGZ) at +270 km s^{-1} which is consistent with a high velocity bow or conical configuration. Indeed the full width of the optical lines of this component is at least 220 km s^{-1} (HRH). Nevertheless it is not clear how the two components could be pieced together.

A more consistent model for HH32A begins by considering all the optical emission to arise from a single bow shock. This has been successfully modelled by a bullet moving at 360 km s^{-1} at 30° to the line of sight through a medium which has a radial velocity component of 25 km s^{-1} (HRH). The two Hα peaks can then be reasonably fitted. The low velocity peak occurs from the oblique wings for which the pre-shock velocity is altered (but not greatly) in the cooled gas. The high velocity peak arises

from the apex and moves with a speed close to that of the bullet. The H_2 emission then finds a simple two component interpretation: the low velocity H_2 is also from the wings while the high velocity H_2 is from shocks within the bullet itself. The two high velocity peaks coincide, since the compressed apex gas and the pre-shock bullet gas are moving with roughly the same velocity. The nature of the bullet shock will depend on its origin and the bullet geometry. It could, for example, be caused by the collision with small dense clumps of filaments in the ambient medium which are overtaken by the leading bow shock and impact directly onto the bullet surface. These shocks would be of low velocity if the bullet is assumed to be much denser or much larger than the clumps. Therefore dissociation of the bullet's H_2 is not a problem. An alternative to the clumps is that the bullet has recently entered a denser region and the internal shock then acts to slow down the cloud. This model can thus explain a great deal but if the tangential motion of HH 32A is close to 50 km s^{-1} (Herbig & Jones 1983), which needs confirmation (HRH), then the bullet apparently must be moving at about 20° to the line of sight. The optical emission is, however, dominated by the tail which may appear slow moving if the bow shock is overtaking clumps i.e. the tangential motion is variable. In this picture the low velocity H_2 emission also comes from these clumps. Since the shock is weaker in the clumps than in the ambient medium, this advances the H_2 region to a position closer to the apex even though the bullet's velocity is so high. Also note that the speed of 360 km s^{-1} (HRH) predicts the Hα peak to occur at velocities higher than observed. A lower velocity, near to 300 km s^{-1}, would be more suitable as the possibility that the gas is not completely cooling locally might explain the wing to higher velocities in the Hα profile (HRH).

Asymmetries are evident in a number of H_2 profiles. For example HH 40 has an AI20 ~ 0.15 for the 1-0 S(1) line, which implies a much higher asymmetry for the intrinsic profile. The peak is at +16 km s^{-1} and a broad

red wing extends to 80 km s^{-1} (ZMGZ). The intrinsic profile must be 50 - 60 km s^{-1} wide. Out of the possibilities considered here, this suggests a conical configuration. HH 40 is indeed a knot within a jet and is not associated with the end of a flow. A bow shock interpretation is not valid for another reason: although the Hα profile is 200 km s^{-1} wide the H_2 and Hα peaks are separated by ~ 90 km s^{-1}, inconsistent with the statement above. A two component model involving shocks which propagate into the ambient medium or entrained slow moving H_2 is necessary.

A shift of the line peak by $V_2 \gtrsim 10$ km s^{-1} relative to the ambient cloud (as taken from CO measurements) was noted by ZMGZ. The shift is in the same sense as the shift of the optical peak in all cases where $|V_1| > 40$ km s^{-1}. These shifts are not accounted for by our bow shock model but are found for low-velocity 'open' cones (i.e. α large) which resemble planar shocks. Although planar shocks have V_2 within 5 km s^{-1} of the pre-shock velocity there is a plateau of emission out to ~ 32 km s^{-1} (for lines from an upper energy level of 7000K) if the cooling is dominated by H_2 (Paper 2). When combined with a Gaussian instrumental profile the peak is shifted by ~ 16 cos θ km s^{-1} for a 40 km s^{-1} shock. Such low velocity planar shocks may be present in the obstacle in the course of the flow. To be found as often as in the ZMGZ survey, however, puts unreasonable constraints on the shock velocity and angle to the line of sight. It would appear more likely that the ambient gas retains a residual velocity from previous encounters with bullets or blast waves, the effects of which have not been completely dissipated by turbulence.

6 CONCLUDING SUMMARY

Curved C-shock surfaces with a wide range of velocities have been investigated with the following conclusions.

1. The H_2 profiles from paraboloidal or hemispherical bow shocks are single-peaked and narrow. The widths at 10% of the peak intensity are $\leqslant$ 45 km s^{-1}. These line widths are generally much less than theoretical limits and only small variations with the upper energy level were found. These results are the consequences of two facts. First, H_2 can survive in the bow wings only and so is only weakly deflected by the shock. Second, the emission from a planar shock element is strongest at velocities close to the pre-shock velocity. Hence the emission is dominated by warm undeflected gas. Strong asymmetry occurs when the motion is within 30° of the line of sight. The line peak is generally within 1 km s^{-1} of the radial component of the pre-shock velocity.

2. Conical shocks result in wider lines but these depend critically on the cone angle which must be near to the maximum allowed by the breakdown condition. Exceptional cases of asymmetry and wider lines which involve particular field directions have been noted.

The gas cools well before being swept downstream past the obstacle (cloud, bullet or jet working surface) or into a turbulent wake provided the obstacle size is much greater than L_c, as given by equation (2). This condition is most difficult to satisfy in low density gas.

3. The theoretical profiles are consistent with observed profiles and fit the optical data with the exceptions of the smooth, wide profiles of OMC-1 and DR21, and the lumpy wide HH 11 profile. They are also inconsistent with the observed shift of the H_2 peaks being often $>$ 10 km s^{-1} relative to the large scale ambient cloud. While there are reasonable explanations for the latter two problems (Section 5) the extremely wide profiles remain enigmatic. Wider profiles may be obtainable if the breakdown conditions are somehow relaxed. One possibility is that the magnetic field in this source is exceptionally large so that the Alfvén Mach number is close to unity. Streaming velocities are then reduced, easing the self-ionisation breakdown limit, and dissociation is also reduced since this breakdown velocity is $\propto$

$B^{1/4}$ (Paper 1). This case will be explored elsewhere.

4. A bow shock forms around an obstacle when there is supersonic motion. A second shock propagates through the obstacle, very slowly if the obstacle is dense. This shock may appear close to planar in its emission line properties: very narrow and asymmetric. A profile may then have two distinct components separated in velocity and space.

Stringent tests of the C-type bow shock model will be applied in the near future by the latest generation of infrared detectors and spectrometers. Much improved optical data, which should refine the modelling, are now perhaps not far away. Further tests will be provided by determining and comparing line ratios and column densities with the theoretical values and this will be our next task.

ACKNOWLEDGEMENTS

This work was much improved through discussions with A. Moorhouse and the scrutiny of P. Williams and B. Guthrie. MDS acknowledges the financial support provided by SERC.

REFERENCES

Brand, P.W.J.L., Toner, M.P., Geballe, T.R. & Webster, A.S., 1989. Mon. Not. R. astr. Soc., 237, 1009.

Burton, M.G., Hough, J.H., Axon, D.J., Hasegawa, T., Tamura, M., McCaughrean, M.J. & McLean, I.S., 1988. Mon. Not. R. astr. Soc., 235, 161.

Burton, M.G., Geballe, T.R. & Brand, P.W.J.L., 1989. Mon. Not. R. astr. Soc., 238, 1513.

Choe, S.-U., Böhm, K.-H. & Solf, J., 1985. Astrophys. J., 288, 338.

Doyon, R. & Nadeau, D., 1988. Astrophys. J., 334, 883.

Draine, B.T., 1980. Astrophys. J., 241, 1021.

Draine, B.T., Roberge, W.G. & Dalgarno, A., 1983. Astrophys. J., 264, 485.

Garden, R.P., 1987. In *Star Forming Regions*, IAU Symp. 115, eds. Peimbert, M. & Jugaku, J., 325.

Hartigan, P., Raymond, J. & Hartmann, L., 1987. Astrophys. J., 316, 323.

Hartigan, P., 1988. In *Formation & Evolution of Low Mass Stars*, eds. Dupree, A.K. & Lago, M.T., (Kluwer Academic), 285.

Heckman, T.M., Miley, G.K., Breugel, W.J.M. & Butcher, H.R., 1981. Astrophys. J., 247, 404.

Herbig, G.H. & Jones, B.F., 1983. Astron. J., 88, 1040.

Hollenbach, D. & McKee, C.F., 1979. Astrophys. J. Suppl., 41, 555.

Hollenbach, D. & McKee, C.F., 1989. Astrophys. J., 342, 306.

Langer, W.D., 1985. In Protostars and Planets, eds. Black, D.C. & Matthews, M.S., p650.

Nadeau, D. & Béland, S., 1988. Astron. J., 95, 136.

Raymond, J.C., Hartigan, P. & Hartmann, L., 1988. Astrophys. J., 326, 323.

Schwartz, R.D., Williams, P.M., Cohen, M. & Jennings, D.G., 1988. Astrophys. J., 334, L99.

Shull, J.M. & McKee, C.F., 1979. Astrophys. J., 227, 131.

Smith, M.D. & Brand, P.W.J.L., (Paper 1) 1990a. Mon. Not. R. astr. Soc.,

in press.

Smith, M.D. & Brand, P.W.J.L., (Paper 2) 1990b. Mon. Not. R. astr. Soc., in press.

Smith, M.D. & Brand, P.W.J.L., (Paper 4) 1990c. Mon. Not. R. astr. Soc., in preparation.

Wardle, M. & Draine, B.T., 1987. Astrophys. J., 321, 321.

Zinnecker, H., Mundt, R., Geballe, T.R. & Zealey, W.J., 1989. Astrophys. J., 342, 337 (ZMGZ).

FIGURE CAPTIONS

Fig. 1. A sketch of the bow shock geometry. Molecular hydrogen dissociates in the speckled area around the apex. The observer lies in the x-z plane at an angle θ to the z-axis. The magnetic field direction is described by the angles Ψ and Π. The angle of incidence of the flow onto the bow surface is $\alpha = \sin^{-1} (v_{\parallel}/v_w)$.

Fig. 2. The profiles obtained for parabolic 'bows' and hemispherical 'caps' and also a comparison of H_2 cooling with the total cooling law including dissociative cooling. The upper energy level is 7000K. The full lines are for H_2 radiative cooling ($n = 10^6$ cm^{-3}, $\chi = 10^{-7}$, $B = 10^{-3}$G, with $v_w = 160$ km s^{-1} and the dashed lines are for the full cooling law ($\epsilon = 4.25 \times 10^{-4}$) with $v_w = 120$ km s^{-1}. The velocity shift is in kilometres per second.

Fig. 3. H_2 line profiles and the orientation to the observer. All figures employ the parameter set stated under Fig. 2 unless otherwise specified.

Fig. 4. The profile dependence on the effective magnetic field factor τ and the angle of the field in the x-y plane. For curves (i) and (ii) $\Psi = 90^{\circ}$.

Fig. 5. The profile dependence on the breakdown velocity $v_d = 40$ or 45 km s^{-1} and on the field direction.

Fig. 6. Profiles produced from widely spaced upper energy levels for a paraboloid.

Fig. 7. Profiles from various upper energy levels for an hemispherical cap.

Fig. 8. The full width at 10% peak intensity for cones with field orientation $\psi = 90^\circ$ and viewing angle $\theta = 90^\circ$. The full lines and dashed lines are for H_2 and H_2O cooling respectively. The left-hand and right-hand lines are for $v_w = 160$ and 80 km s^{-1}, as indicated. The figure demonstrates the narrow range of cone angles which give wide lines.

Fig. 9. Line profiles from $T_j = 7000K$ produced by cones at various orientations and for an oblique field $(\psi, \eta) = (30^\circ, 60^\circ)$.

Fig 10. Asymmetric line profiles can be produced by cones even when the motion is in the plane of the sky ($\theta = 90^\circ$). Here $v_w = 80$ km s^{-1} and (1) $\psi = 70^\circ$, $\eta = 90^\circ$ (2) $\psi = 45^\circ$, $\eta = 90^\circ$ and (3) $\psi = 35^\circ$, $\eta = 30^\circ$ and $\theta = 150^\circ$.

Fig. 11. Lines convolved with a Gaussian profile with $\sigma = 15$ km s^{-1}. (1) An H_2 cap, $\theta = 90^\circ$ and $v_w = 120$ km s^{-1} (deconvolved profile is shown in Figs 5) and (2) an H_2O bow, $\theta = 30^\circ$ and $v_w = 160$ km s^{-1} (Fig. 3). Note the full widths exceed 80 km s^{-1} and the loss of features such as the asymmetry.

Table 1. H_2 line widths for T_j = 7000K at 10% peak intensity. Orientation: θ = 90°, field direction (Ψ,Π) = (90°,0°). In all tables we take n = 10^6 cm^{-3}, χ = 10^{-7}, B = 10^{-3}G and ϵ = 4.25 10^{-4}. Paraboloidal (Par) and hemispherical (Hem) bow shocks are taken.

v_w	H_2 Par	H_2 Hem	H_2O Par	H_2O Hem	Total cooling. Par
40	23	21	15	14	16
80	31	41	24	18	26
120	31	41	25	24	26
160	31	41	25	24	26
200	31	39	25	24	26

Table 2. H_2 line widths for different viewing angles. Bow velocity: v_w = 40 km s^{-1}. Other parameters are as in Table 1.

θ	H_2 Par	H_2 Hem	H_2O Par	H_2O Hem
0°	15	26	15	17
30°	17	24	15	17
60°	21	22	15	14
90°	23	21	15	14

Table 3. H_2 line widths for different viewing angles. Bow velocity: v_w = 160 km s^{-1}. Other parameters are as in Table 1.

θ	H_2 Par	H_2 Hem	H_2O Par	H_2O Hem
0°	7	10	5	7
30°	16	21	15	13
60°	25	32	22	22
90°	31	41	25	24

Table 4. Line widths for bows demonstrating the degree of sensitivity to the breakdown velocity (row 2) and magnetic field direction (row 3). Orientation: $\theta = 90^{\circ}$. Basic parameters as for Table 1.

v_w	v_b	Ψ,Π	H_2 Par	H_2 Hem	H_2O Par	H_2O Hem
120	45	90°,0°	31	41	27	29
120	40	90°,0°	31	36	25	24
120	45	0°,0°	29	45	24	31

Table 5. Cones. Orientation of observer: $\theta = 90^{\circ}$. The maximum full width (10% peak intensity) and the cone half-angle (α) at which the maximum is reached are given. The field direction $(\Psi,\Pi) = (0^{\circ},0^{\circ})$ produces the wider profiles for v_w above 40 km s^{-1}.

v_w	H_2 (90°,0°)	H_2 (0°,0°)	H_2O (90°,0°)	H_2O (0°,0°)	Total cooling (0°,0°)
40	29 (45°)	26 (45°)	17 (45°)	14 (45°)	16 (45°)
80	51 (31°)	65 (34°)	28 (27°)	41 (30°)	52 (39°)
120	52 (19°)	73 (22°)	28 (16°)	45 (19°)	61 (25°)
160	49 (13°)	75 (16°)	29 (11°)	45 (14°)	64 (18°)

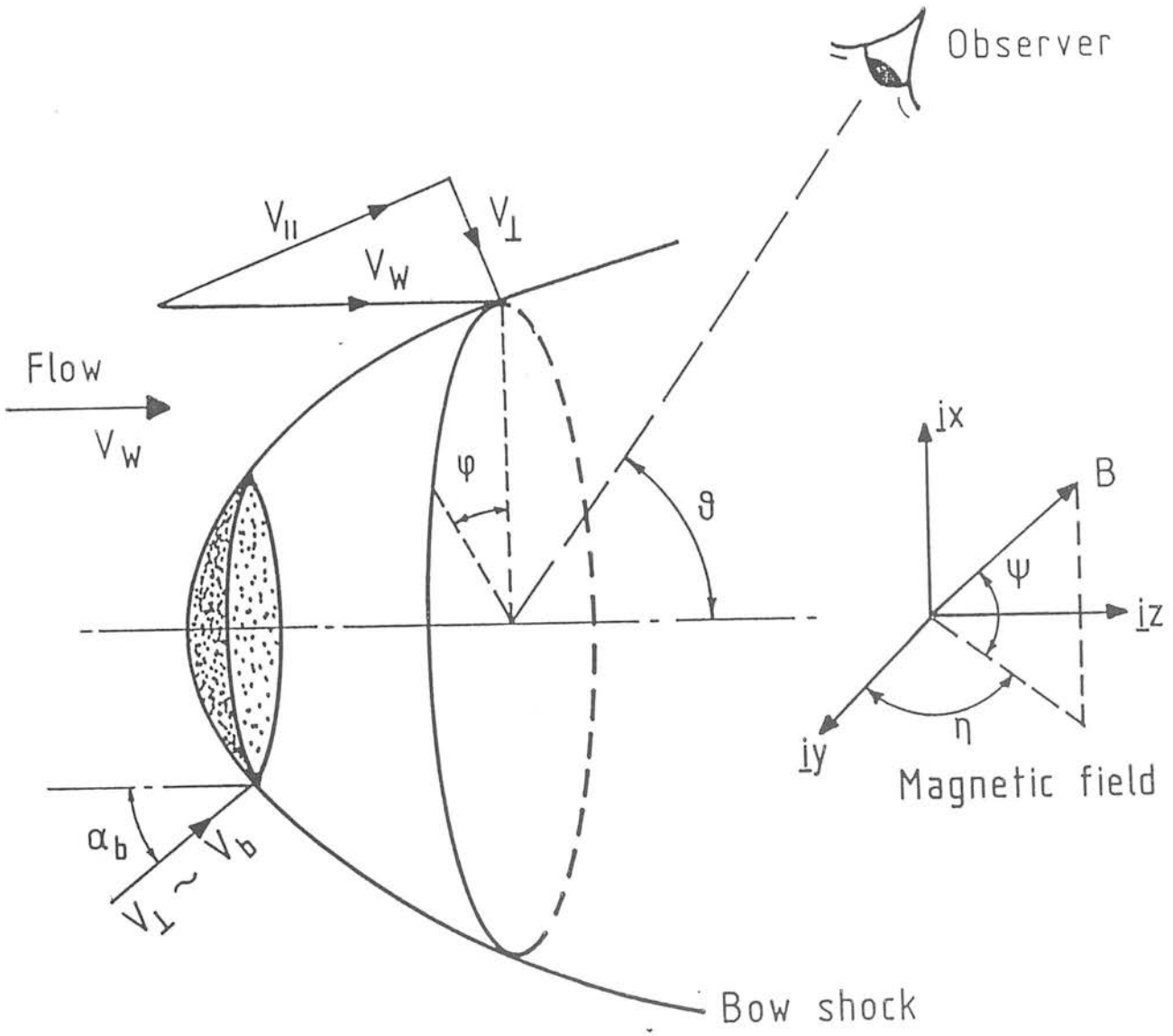

Observer
V_{II}
$V_\perp$
V_W
Flow
V_W
φ
ϑ
ix
B
Ψ
iz
η
iy
Magnetic field
α_b
V_b
$V_\perp$
Bow shock

Fig 1.

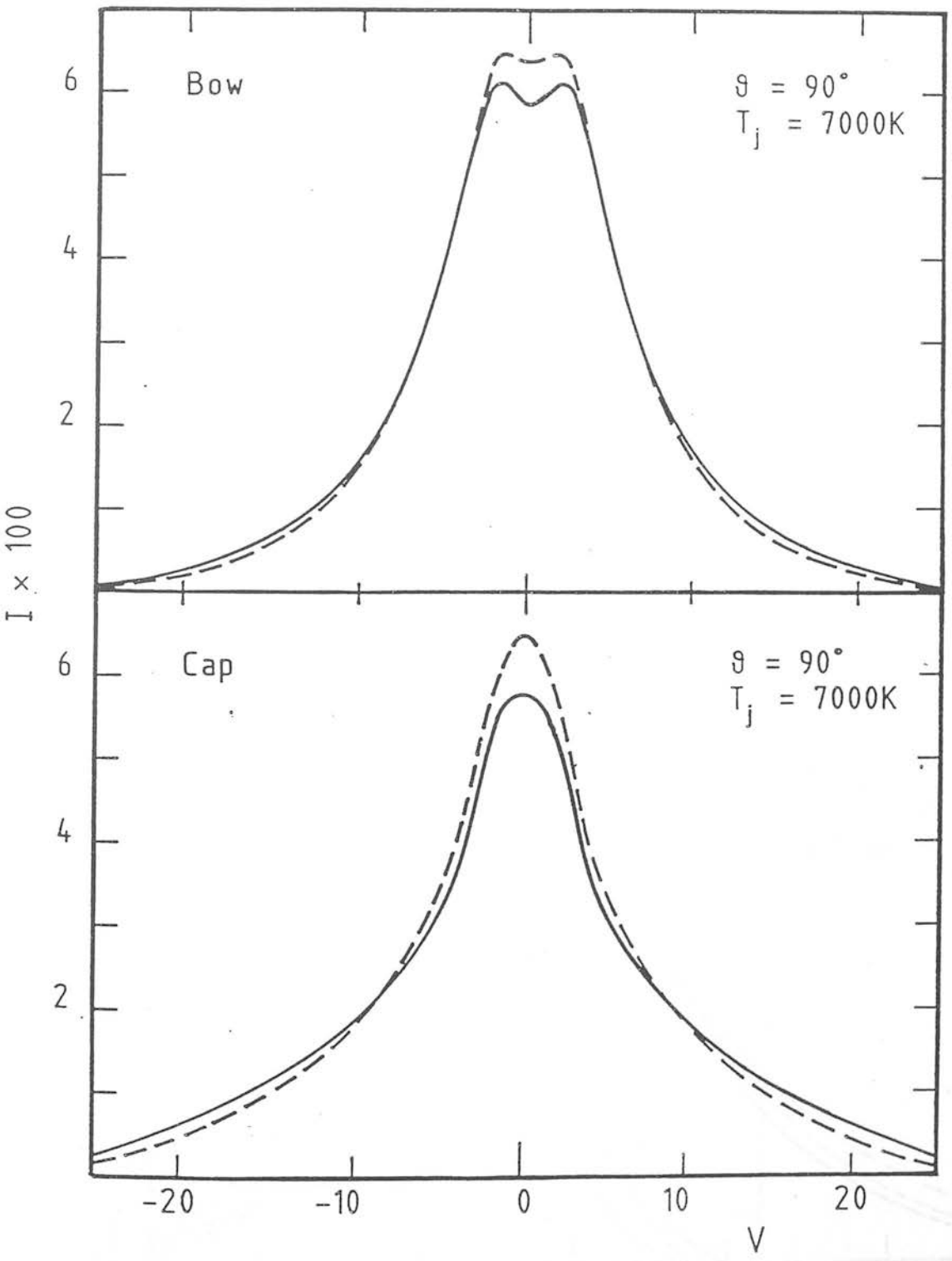


Fig.2

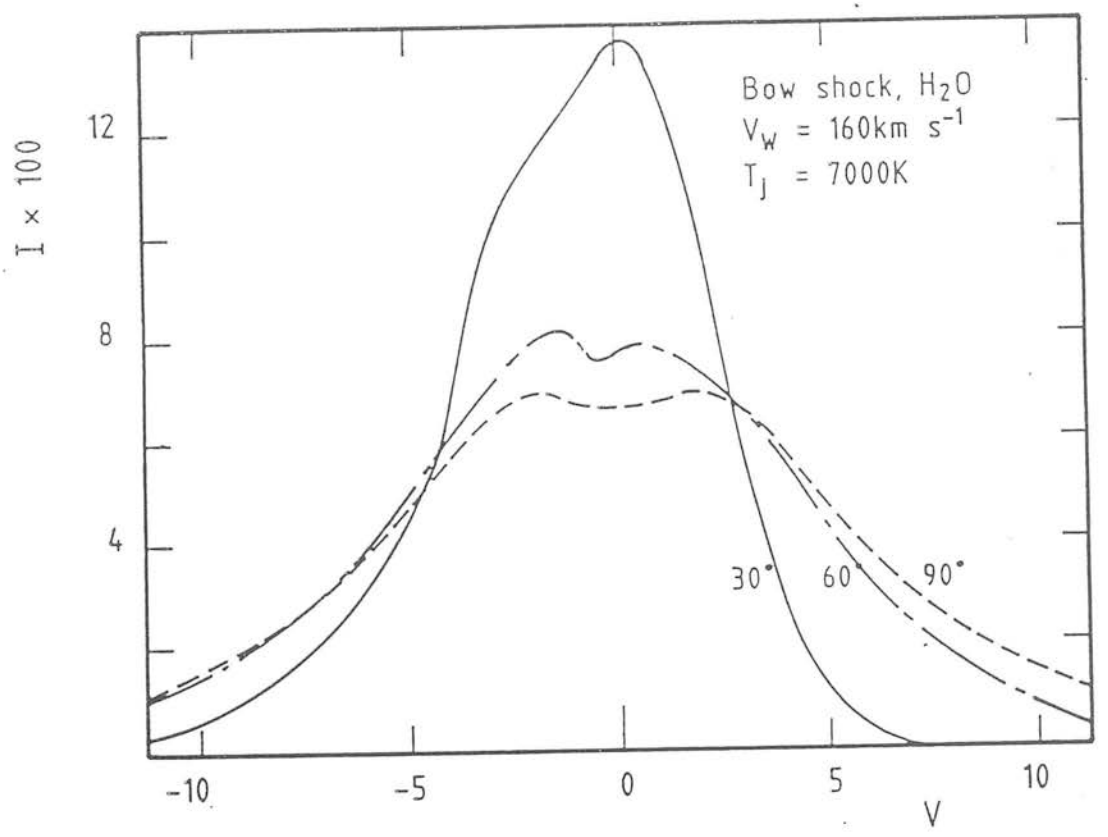

Bow shock, H_2O
V_W = 160km s^{-1}
T_J = 7000K
30°
60°
90°
I × 100
12
8
4
-10
-5
0
5
10
V

Fig. 3

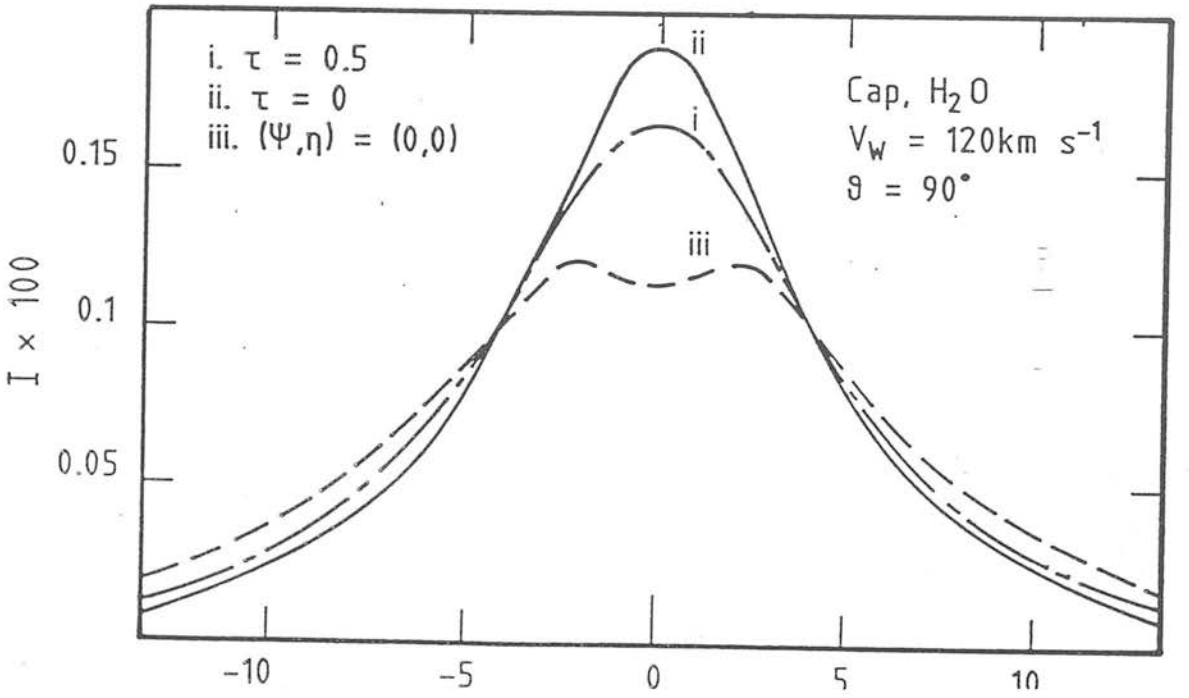

i. τ = 0.5
ii. τ = 0
iii. (Ψ,η) = (0,0)
Cap, H_2O
V_W = 120km s^{-1}
ϑ = 90°
i
ii
iii
I × 100
0.15
0.1
0.05
-10
-5
0
5
10

Fig. 4

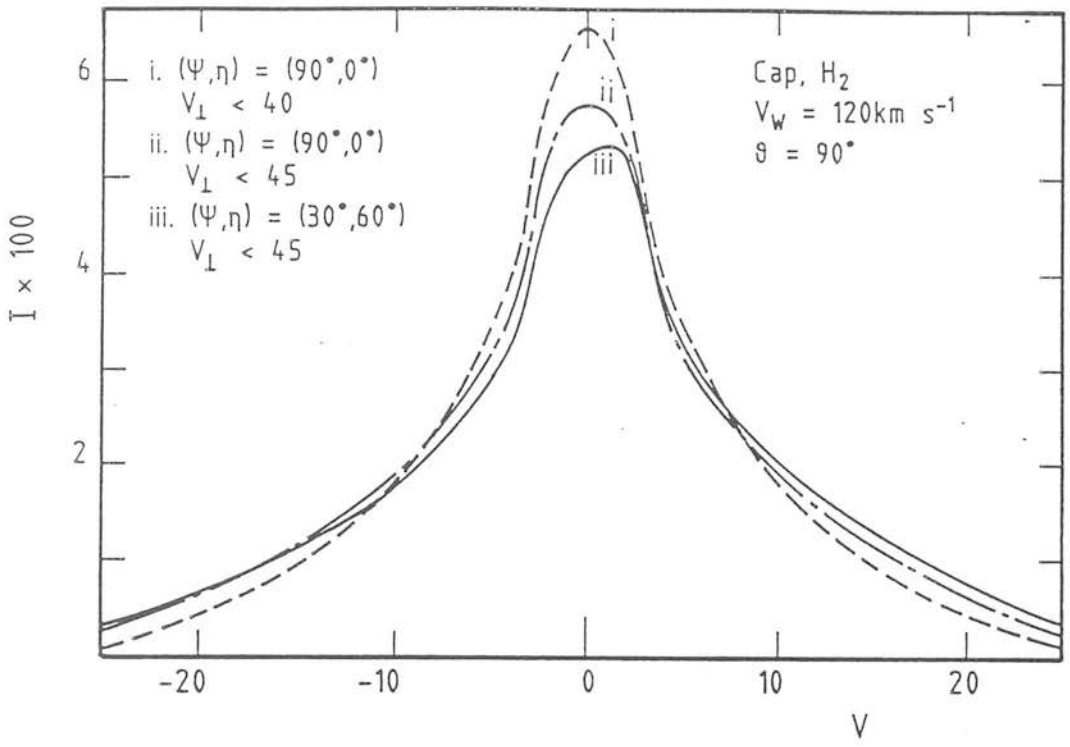

i. (Ψ,η) = (90°,0°)
V⊥ < 40
ii. (Ψ,η) = (90°,0°)
V⊥ < 45
iii. (Ψ,η) = (30°,60°)
V⊥ < 45
Cap, H2
Vw = 120km s-1
ϑ = 90°
I × 100
V

Fig. 5

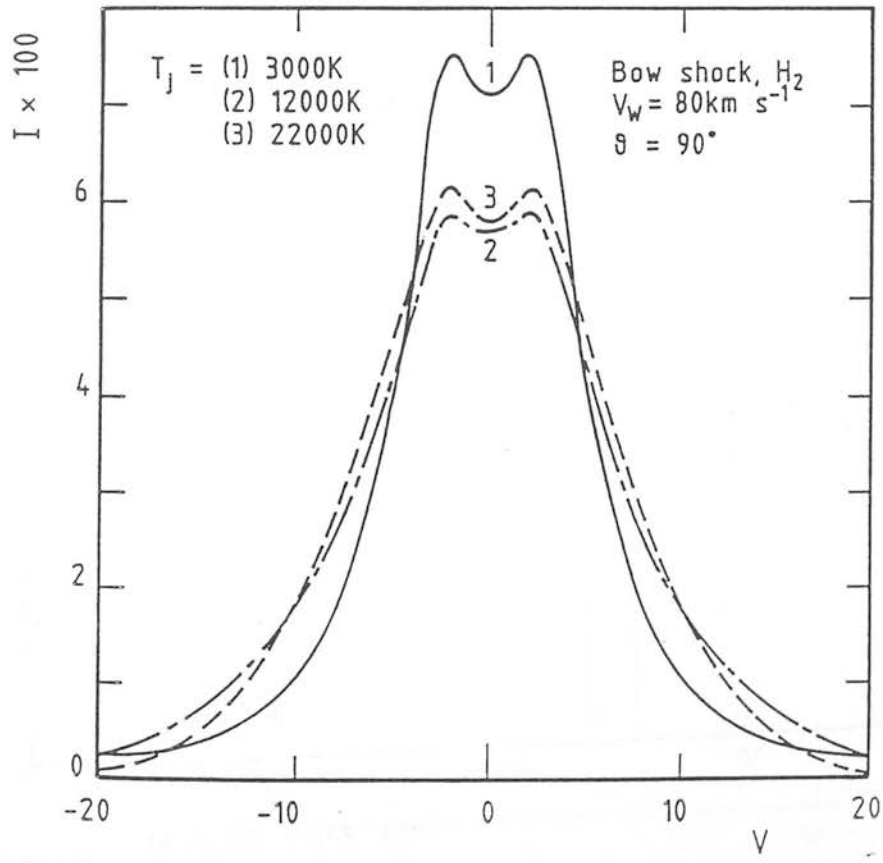

Tj = (1) 3000K
(2) 12000K
(3) 22000K
Bow shock, H2
Vw = 80km s-1
ϑ = 90°
I × 100
V

Fig. 6

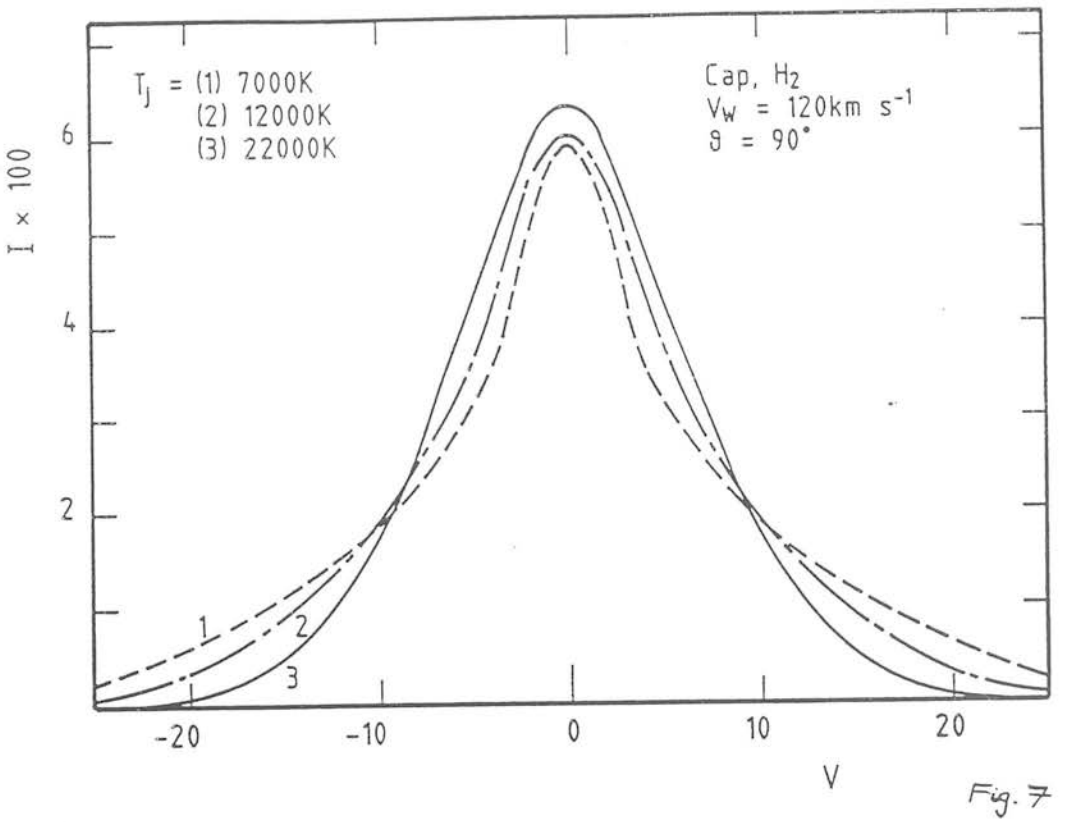

T_j = (1) 7000K
(2) 12000K
(3) 22000K
Cap, H_2
V_W = 120km s^{-1}
ϑ = 90°
I × 100
6
4
2
1
2
3
-20
-10
0
10
20
V

Fig. 7

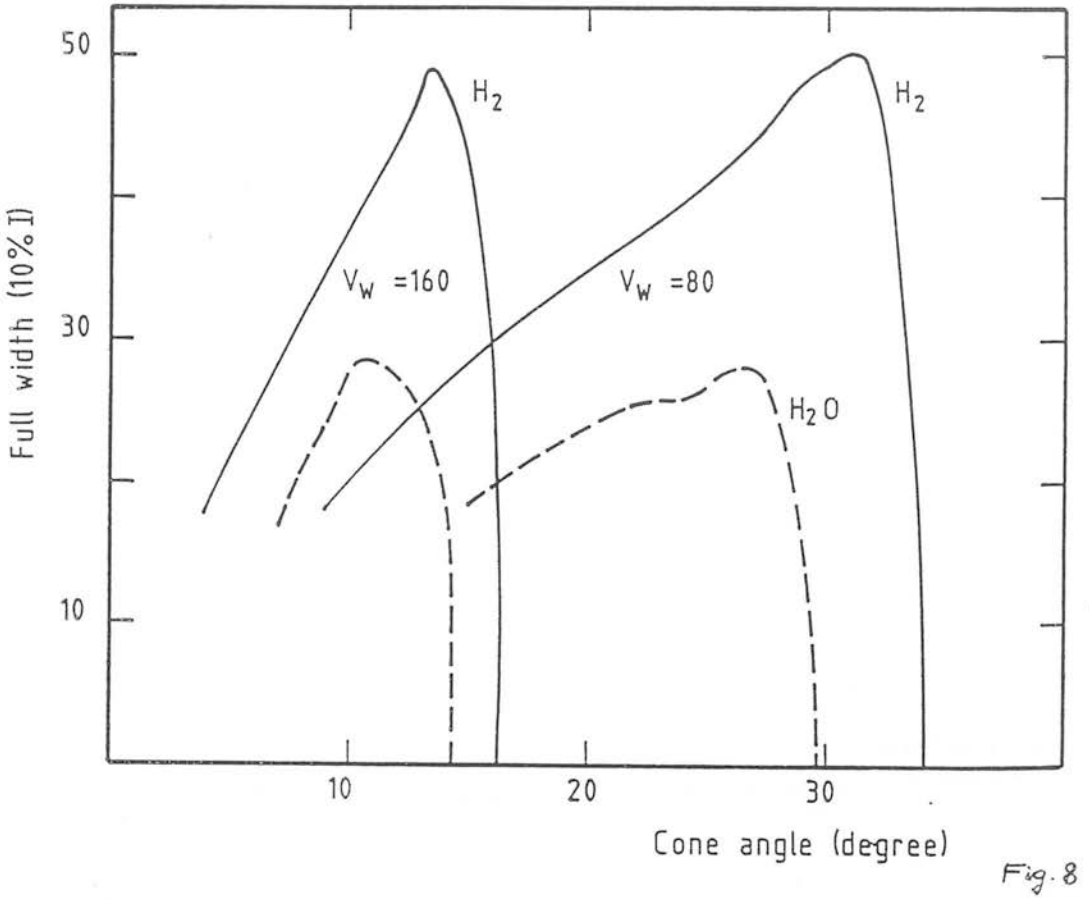

Full width (10% I)
50
30
10
H_2
H_2
V_W =160
V_W =80
H_2O
10
20
30
Cone angle (degree)

Fig. 8

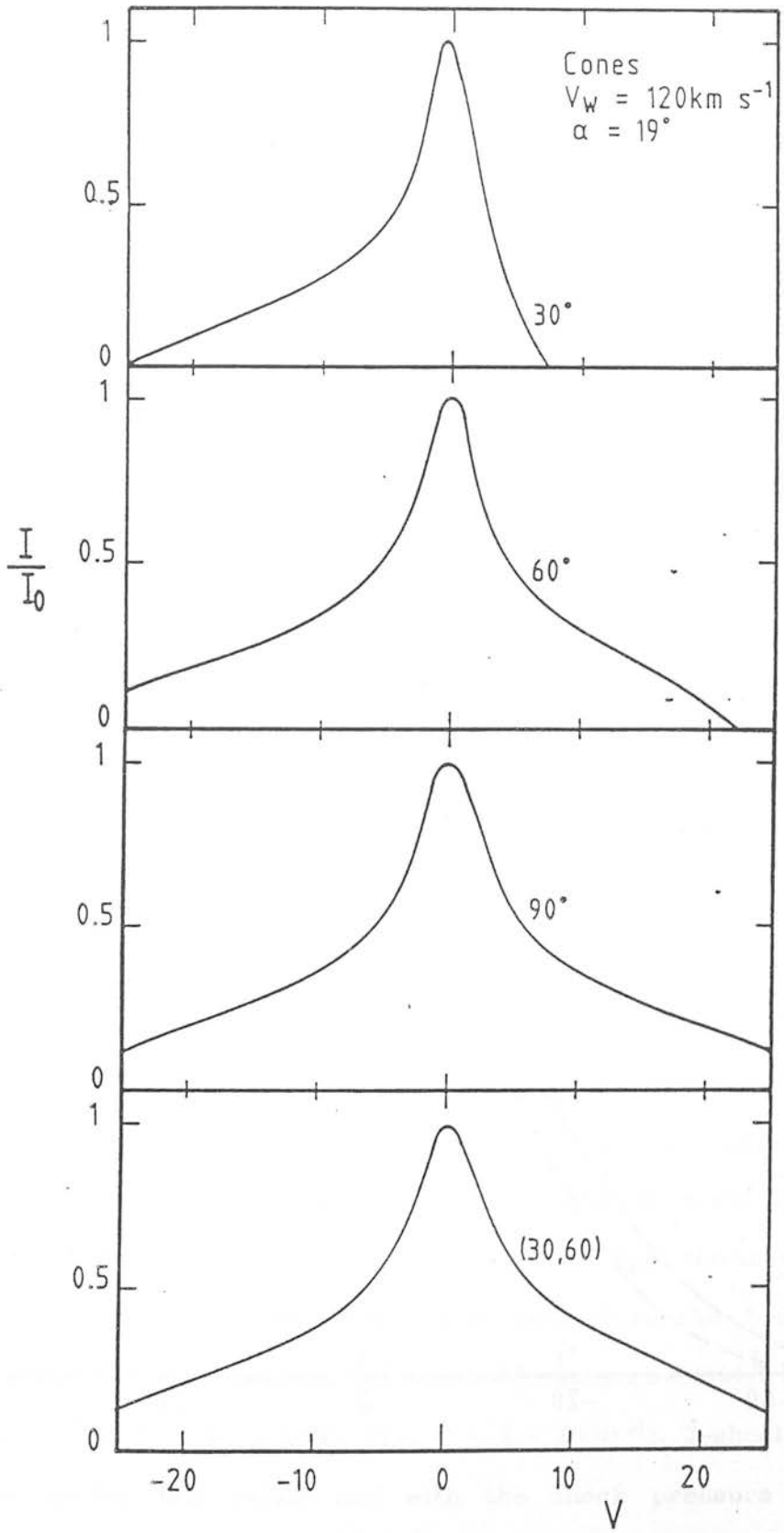


Fig. 9

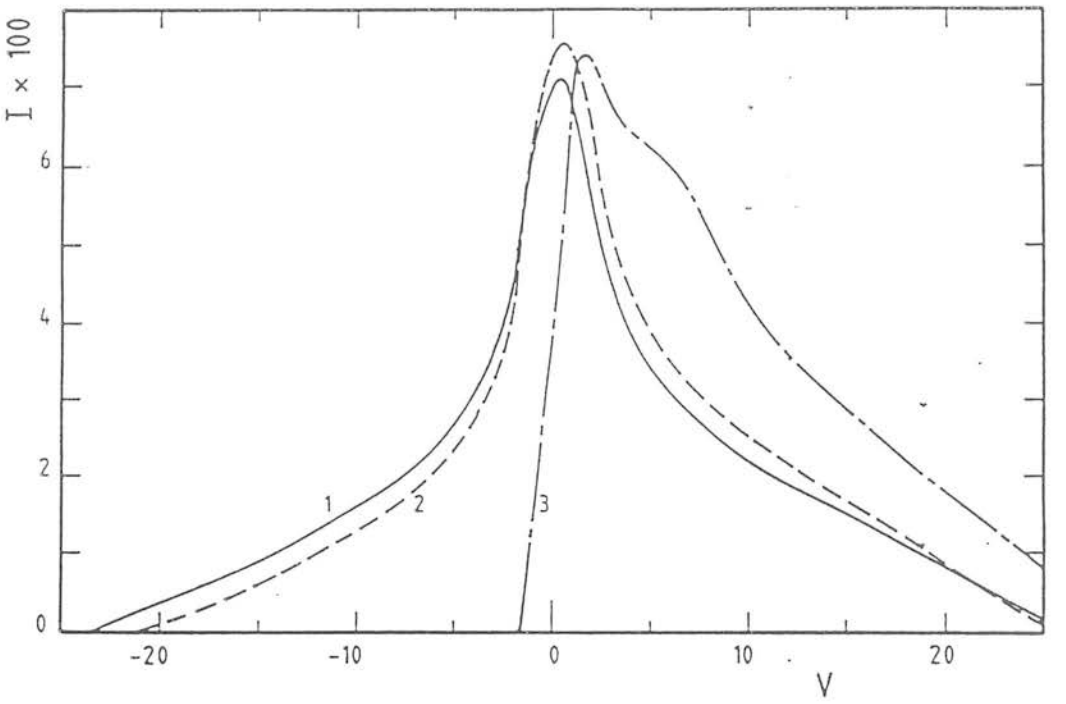


Fig. 10

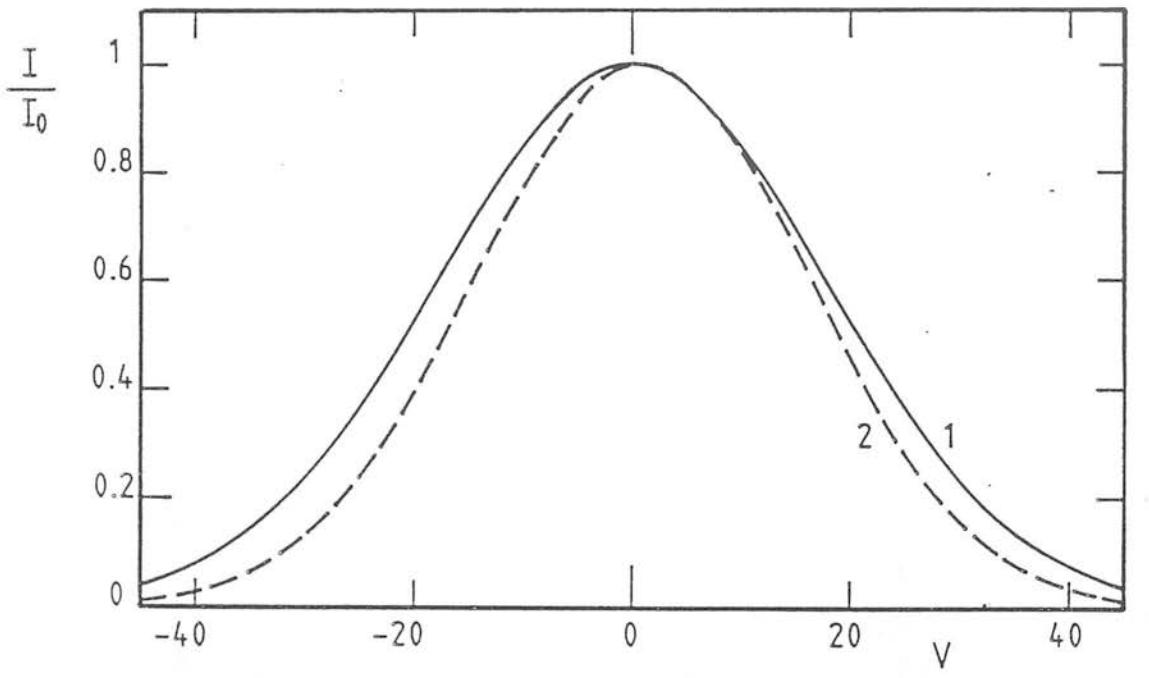


Fig. 11

BOW SHOCKS IN MOLECULAR CLOUDS: H_2 LINE STRENGTHS

M. D. Smith, P.W.J.L. Brand & A. Moorhouse

Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, U.K.

Submitted to *Monthly Notices R.A.S.*

SUMMARY

Molecular hydrogen line strengths from a model for bow-shaped C-shocks in high density molecular clouds are calculated. The line ratios are independent of the bow speed once it exceeds the breakdown value (typically 30 - 50 km s^{-1}). Then, the H_2 emitting section is simply located further towards the bow tail at higher speeds. The line ratios are sensitive only to the shape of the bow, the (shock-stimulated) H_2O abundance and the cloud density, provided that dissociation determines the breakdown limit. The presently available data are interpretable as paraboloidal bows of high velocity and high H_2O abundance (i.e. $\epsilon \sim 1 - 5\times10^{-4}$). J-shocks (with $\epsilon = 0$) produce similar line ratios but with the shock pressure as the critical parameter. The possibilities for distinguishing between C and J type shocks are discussed.

1 INTRODUCTION

Shocked molecular gas is present in the interstellar medium (Shull & Draine 1987). Much information about the shocks and molecular cloud is derivable from molecular hydrogen line strengths which, due to the many recent advances in infrared astronomy, can now be measured with some accuracy (Brand et al. 1988). To extract this information we must develop the theory to the equivalent stage. Here we calculate the H_2 line strengths for bow shock models. Each section of the bow is treated as a planar C-shock (see below). The results are then compared to those for bow J-shocks, to planar shocks of both C and J type as well as to the observed data.

A C-shock is a continuous MHD wave in which the magnetic field cushions a transition via ion-neutral collisions (Draine 1980; Draine, Roberge & Dalgarno 1983). Low ionization in the pre-shocked medium is necessary (Smith & Brand 1990a) but this is expected in high density molecular cloud cores (Langer 1985). A program to find the full signatures of C-shocks is underway. First analysed were planar C-shocks with high Alfvén Mach numbers. They produce H_2 emission line profiles and line ratios which are inconsistent with those of OMC-1 and DR21 (Smith & Brand 1990b).

Profiles from parabolic bow shocks, conical jet shocks and hemispherical caps have also been presented (Smith & Brand 1990c). These are quite narrow (less than 70 km s^{-1}) but generally consistent with the H_2 profiles associated with Herbig-Haro objects (Zinnecker et al 1989). A single bow shock is, however, not sufficient to produce a line as wide as observed at several locations in OMC-1 and DR21 ($\sim$ 150 km s^{-1})

J-shocks produce even narrower lines, since dissociation greatly limits the normal component of velocity ($<$ 25 km s^{-1}). Nevertheless a planar shock model can account for the line strengths in OMC-1 (Brand et al 1988) leading to the possibility that the emission arises from a large number of

components, accelerated to a range of velocities by some mechanism other than the shock (Brand et al 1989b). However progress from this hypothesis is difficult.

Bow shocks are expected when an obstacle is in supersonic motion relative to an ambient fluid. Bow shock structures have been recognised through mapping of knots of atomic emission lines (i.e. Herbig-Haro objects) associated with molecular outflows (e.g. Raga et al. 1988). Bow shock signatures have been found by comparing their emission line strengths and line profile shapes with detailed models (Hartigan, Raymond & Hartmann 1987). It is appropriate to ask if such models can account for the molecular emission lines.

The results are limited to the high density regions where local thermodynamic equilibrium is achieved ($n > 10^6$ cm^{-3} but approximate for n down to 10^4 cm^{-3}). Then, for planar shocks, the line strengths are simply related to the column densities of the upper energy levels. We calculate the column density ratio in the upper energy level T_j: CDR = N_j/N_{jB} – the ratio of the column density to that of a 2000K slab. This is normalised so that CDR = 1 at T_j = 7000K. We employ this method of analysis here except that we determine the total number of molecules in the energy level for the complete bow shock. This can be converted into a column density once the beam size is specified.

Power-law cooling functions are employed as approximations to the cooling mechanisms and ion-neutral drag and dust-neutral drag are approximated by simple, but realistic, expressions (Smith & Brand 1990a). The results are insensitive to these approximations. High Alfvén Mach numbers are assumed for the bows. This is consistent with high velocity bows (e.g. $v > 30$ km s^{-1}) moving through quiescent clouds (e.g. Alfvén velocity < 5 km s^{-1}).

2 THE LINE-EMITTING SECTION

The H_2 line emission is not produced from the whole bow shock. The normal component of shock velocity, $v_\perp$, is limited by the breakdown condition. This gives a maximum velocity v_b above which either direct H_2 dissociation occurs ($v_b = v_d$) or neutral-ion collisions lead to runaway self-ionisation and subsequent dissociation ($v_b = v_i$). Therefore the emission is confined to the section of the bow for which $v_\perp < v_b$. If the bow shock moves through the ambient cloud with a velocity v much greater than v_b, then the emission is confined to the far tail where the shock is sufficiently oblique.

A fairly precise determination of v_b is necessary for this study. We have not used the approximation developed by Smith & Brand (1990a) but calculated v_b numerically for a planar shock for each set of parameters chosen for the bow and undisturbed medium. This is calculated assuming the field to be perpendicular to the shock velocity. The result is then extended by utilising the approximation $v_b \propto B_\perp^{1/4}$ for the surface integration (Smith & Brand 1990c, equation 1) where $B_\perp$ is the effective perpendicular field component.

A sharp cut-off is applied at the velocity v_b. This is an approximation to the true physical situation in which the transition from partial to complete dissociation occurs over a velocity interval of a few km s^{-1}. We also ignore the hydrogen atoms and new ions created in the flow. The question then is: what level of fractional dissociation should be taken to correspond to v_b? As can be seen from Fig. 1a for $v_d < v_i$, this choice is only crucial to line ratios from widely-spaced upper energy levels. The general result is that planar shocks in which the total fractional dissociation is between 20% and 45% correspond to a range in v_b of ~ 4 km s^{-1} and a range in maximum temperature of ~ 300K. We shall presume v_b to correspond to a dissociation fraction close to 30%. At this level about one half of the neutrals would be atoms but dissociative cooling decreases and

collision rates are greatly altered.

If $v_d > v_i$, then we simply take $v_b = v_i$. This corresponds to a wide range in (relatively low) maximum temperatures according to the chosen parameters. This case corresponds to low ionisation fractions ($\chi < 10^{-7}$, Smith & Brand 1990a) or moderately high fields.

3 RESULTS

3.1 The bow velocity

We begin by fixing all parameters except the bow velocity. We take a paraboloid with $n_o = 3\text{x}10^6$ cm^{-3} (above the critical density for LTE cooling), the magnetic field $B_o/10^{-3}\text{G} = (n_o/10^6\text{cm}^{-3})^{½}$ directed perpendicular to the direction of motion, $\chi_o = 10^{-7}$ and an H_2O abundance $\epsilon_o = 4.25\text{x}10^{-4}$. This will be referred to as the standard set. A maximum dissociated fraction of 36% corresponds to $v_b = 48$ km s^{-1}. We consider bow velocities in the range $v \sim 30 - 200$ km s^{-1}, consistent with the bow shock models for the atomic component.

A crucial result is that the line ratios are constant whatever the bow velocity once the bow shock is moving faster than the breakdown velocity (Fig. 1b). The explanation is inherent in the fact that the maximum temperature of 3968K for this example is always reached at some location on the bow. This location is farther from the leading edge for higher bow velocities. The smooth paraboloid ensures that the cooler parts of the flow are still present and in the same proportion to the hot section for all velocities above v_b (see Section 4).

The line ratios are a strong function of velocity only for slow bow shocks, $v < v_b$, as can be seen for v in the narrow range 36 - 44 km s^{-1} in Fig. 1b. In this case the maximum temperature occurs at the leading

edge of the bow. It is determined by v since v_b is not reached.

We deduce, from these results and others not displayed, that we need examine only a single velocity above v_b. We choose v = 120 km s^{-1}.

3.2 Magnetic field and ionization

We first alter the magnetic field and the ionization level in a manner so that dissociation still provides the breakdown condition. As shown in Fig. 1c the ionization level and field strength then have little influence on the line ratios. The field direction is found to have a similar small effect. Again this indicates that the maximum temperature controls the line ratios and is practically a constant since dissociation is an exponential function of temperature (Smith & Brand 1990a).

Breakdown is via streaming ionisations (i.e. $v_d > v_i$) for high magnetic fields or low ionization fractions. Then v_b is fixed at $v_i \sim 49$ km s^{-1} (Smith & Brand 1990a). The maximum temperature, however, is not fixed and the CDRs are sensitive to the parameters (Fig. 1c). The values for which $v_d > v_i$ are $\chi < 7\times10^{-8}$ for $B = 10^{-3}(n/10^6 cm^{-3})^{1/2}$ G and $B > 1.3\times10^{-3}(n/10^6 cm^{-3})^{1/2}$ G for $\chi = 10^{-7}$.

3.3 Hemispherical caps

Hemispherical caps were considered as an alternative to the paraboloidal bows for the H_2 line profile analysis (Smith & Brand 1990c). Caps produce relatively weaker line emission from the lower energy levels (Fig. 1c). This is not unexpected since, although the maximum temperature is fixed, there is now much less surface area highly oblique to the flow direction. In contrast, for paraboloids this oblique section in which the gas remains cool

dominates the low excitation lines. We conclude that the shape of the shock wave is one important factor in interpreting line ratios.

3.4 H_2O abundance and density

The CDR dependence on the H_2O abundance ϵ relative to an assumed standard of $\epsilon_o = 4.25\ 10^{-4}$ per H atom is shown in Fig. 1d. Note that for a low abundance ($\epsilon \rightarrow 0$) or for a high abundance ($\epsilon > 10^{-4}$) the CDRs are approximately constant.

Densities above $3\ 10^6$ cm^{-3} have been considered. The shape of the curves are very close to those of Fig. 1d. The heating increases as well as the cooling and the breakdown velocity actually decreases as the density increases. The $\epsilon = \epsilon_o$ curve, for which H_2O cooling is dominant, is still approximately the upper envelope. This is due to the H_2O cooling and collisional heating both being $\propto n^2$. Thus the line ratios depend on the shape of the cooling function which is nearly fixed once either H_2 or H_2O dominates the cooling. This is explained in Section 4.

4 INTERPRETATION AND LINE INTENSITIES

Why are the line ratios independent of all parameters but the bow shape and the cooling function shape? Previously we have demonstrated the existence of an approximation for the H_2 line intensities from a planar shock (Smith & Brand 1990b). We can use this approximation to help understand the present results. Line ratios were estimated from the maximum temperature T_{max} reached in a C-shock. Thus the CDR curves are a one parameter family. A line strength however also depends on the Alfvén Mach number, M_A. Here we shall use the approximation for each surface

element of the bow wave. The total number of hydrogen molecules in the energy level kT_j is then given by integrating over the surface:

$$N_{totj}/g_j = \int N_j(T_{max},M_A)\ dA \qquad (1)$$

where $N_j/g_j = 44 n_{no} L_o\ (T_{max}/T_j + 0.1)\ \exp(-T_j/T_{max})/(T_{max}M_A)$ is the planar approximation (Smith & Brand 1990b). Here n_{no} is the pre-shock H_2 density and the scale length is $L_o = 1/(\chi n_{no}\sigma_o)$ with $\sigma_o = 10^{-15}cm^2$.

We combine the results for the area weighting presented by Smith & Brand (1990c). The area is $dA = L_p{}^2 \cos\alpha/(\sin\alpha)^m\ d\alpha\ d\Phi$ where m = 4 for a paraboloid, m = 0 for a hemispherical cap, L_p is the appropriate length scale and α and Φ are the angles of incidence and azimuth, respectively (cf Fig. 1 of Smith & Brand (1990c)). Also $M_A = 6.5\ (n_{no}/10^6cm^{-3})\ (v\ \sin\alpha/10\ km\ s^{-1})\ /\ (B/10^{-3}G)$.

We employ the approximations for T_{max} derived by Smith & Brand (1990a). These ignore dissociative cooling. We also replace the normal component of velocity by v sin α. Thus we rewrite equations (42) and (43) of Smith & Brand (1990a): $T_{max} \sim T_b\ (\sin\alpha/\sin\alpha_b)^{4/p}$ where p, the power-law cooling index, is taken as 3.3 for H_2 cooling and 1.5 for H_2O cooling. Here T_b and α_b are the limiting temperature and shock angle for a non-dissociative shock, corresponding to v_b i.e. $\sin\alpha_b = v_b/v$. Equation (1) can now be written as

$$N_{totj}/g_j = 4.3\ 10^{22} L_p^2 \left[\frac{B/10^{-3}G}{n_6{}^{1/2}\chi_{-7}T_{j4}}\right]\left[\frac{100\ km\ s^{-1}}{v}\right]\left[\frac{v}{v_b}\right]^m \int_o^1 \frac{(0.1+y)}{x^m y e^{1/y}}\,dx \qquad (2)$$

where $T_{j4} = T_j/10^4K$, $\chi_{-7} = \chi/10^{-7}$, $n_6 = n/(10^6cm^{-3})$ and $y = (T_b/T_j)x^{4/p}$.

It is clear that line intensities are functions of all parameters. In

contrast a line ratio depends only on p, m and T_b, the variables within the integral. When H_2 survival is limited by dissociation, the dependence on T_b is negligible since T_b remains in the narrow range 3700K - 4000K. Thus the bow shape (m) and cooling function (p) determine the CDR.

Equation (2) can be used as a rough but extremely quick method to determine CDRs and line strengths from bow C-shocks. The accuracy is demonstrated in Fig. 2.

The line intensities are found directly from the total number of molecules in the associated upper energy level provided that the density exceeds the critical value. Numerically, for T_j = 7000K we find N_{totj}/g_j to increase from $1.9\times10^{18}L_p^2$ to $2.8\times10^{20}L_p^2$ as a paraboloidal bow velocity v is increased from 40 to 200 km s^{-1} (with the standard values and L_p = $R^2/(2z)$ is the scale length of the paraboloid).

A size relevant to the observations is desirable. We take L_{ob} to be the width of the bow (i.e. 2R) at $\alpha = \alpha_b$, the location where the H_2 emission rapidly turns on. Then $L_{ob} \sim (v/20 \text{ km s}^{-1})L_p$ and $N_{totj} = (5\times10^{17} - 3\times10^{18})L_{ob}^2$. This is a remarkably narrow range consistent with the present data (see the discussion by Smith & Brand 1990b) since higher values have not been observed for the 1-0 S(1) line at 6956K and clumpiness or an unfilled beam can account for lower apparent values.

The line intensities are estimated from equation (2). We find

$$N_{totj}/g_j \sim 1.3\ 10^{18}\ L_{ob}^2 \left[\frac{v}{100 \text{ km s}^{-1}}\right] \left[\frac{B/10^{-3}G}{n_6^{1/2}\ X_{-7}}\right], \qquad (3)$$

for T_j = 7000K in agreement with the numerical results. This is useful as an indication of the dependence on the individual parameters.

5 DISCUSSION

We have uncovered a major difference between the properties of planar and bow C-shocks. In the latter the excitation temperature increases more rapidly with T_j. Bow C-shocks can explain the line ratio data of OMC-1 (Fig. 3a), IC 443 (Fig. 3b) and CRL 618 (Burton et al 1989). The planar shock version was unable to reproduce the observationally derived CDRs for T_j above 12,000K.

J-shocks are also able to explain the CDRs provided H_2O is not an important coolant. The ions and neutrals are co-moving and the field is taken to cushion the shock only after the H_2 has cooled. The CDR curve then simply depends on the pressure which determines the relative importance of dissociative cooling. Fig. 4 demonstrates the pressure dependence for J-shocks derived from the analysis of Brand et al. (1988). So how can we differentiate between the two models on the basis of their H_2 line predictions?

J-shock and bow C-shock models both incorporate a means of explaining the properties of the OMC-1 line data. It has been found that the intensity ratio of 1-0 O(7) to 0-0 S(13) is independent of position in the outflow (Brand et al 1989a). For planar J-shocks this implies a constant driving pressure which could be brought about by an expanding high pressure bubble. For C-shocks this is interpreted as a high H_2O formation rate and the widespread presence of bow-forming obstacles. A convincing scenario to explain the width of the lines in either model is lacking.

The line ratios are generally not constant from source to source. This is interpreted by the pressure dependence in the J-shock model. The low CDRs in IC443 (844"W, 635"N) imply a pressure ~ 6×10^{11} K cm^{-3}, two orders of magnitude greater than the driving SNR pressure (Moorhouse et al 1990). Such a high pressure could be attained in a collision of a clump with a shocked and cooled layer (Moorhouse et al 1990). By contrast, in the

C-shock model the pressure is not constrained and is consistent with the supernova blast. The low CDRs are interpreted as either a low H_2O abundance (e.g. photodissociated) or a magnetic field of value (2 - 3) x $(n/10^6 cm^{-3})$.

The H_2 line emission from a C-shock occurs in the warm gas with density and speed close to the pre-shock values. For a J-shock the emission is from the dense cooled layer moving close to the shock velocity. The lines from planar shocks are thus narrow. Bow J-shocks are subject to the same theoretical limits on the line width as bow C-shocks (Smith & Brand 1989c) except with a breakdown velocity of 25 km s^{-1}. Hence the upper bound to a line width is ~ $2v_b$ ~ 50 km s^{-1}.

The CDR curves from J-shocks are insensitive to the shape of the shock front. The line ratios from bow J-shocks will not greatly differ from those for planar J-shocks because of the limited range in pressure across the shock surface. Whereas the CDR curves shown in Fig. 4 cover orders of magnitude in pressure, a pressure range across a bow of only 6.25 results from a unifom medium with normal component of velocity in the range 10 - 25 km s^{-1}. C-shocks are considerably more sensitive to the shock configuration as demonstrated by the hemispherical cap (Fig. 1b) and the planar shock (Smith & Brand 1989b).

Molecular reformation may lead to high velocity H_2 far downstream even for a completely dissociating shock. This H_2 will be cold and one must invoke further non-dissociative shocks to follow. Thus the reformation mechanism does not help to interpret line strengths. The reformation behind the initial shock will occur after a column density ~ $5x10^{21} cm^{-2}$ (Hollenbach & McKee 1989) which yields a length scale ~ $5x10^{15}/(n_6 v_{100})$ cm^{-2} where v_{100} is the shock velocity in units of 100 km s^{-1}. This scale is not excessive in comparison to bow shock sizes provided $n > 10^6 cm^{-3}$. This mechanism may possibly explain the OMC-1 H_2 data including the ~ 140 km s^{-1} line widths but requires the secondary shocks to propagate into the

dense, cold magnetically-supported H_2 layer. Twin shocks could work in other contexts with two consecutive C-shocks able to accelerate H_2 to $2v_b$, producing widths of up to 200 km s^{-1} in a non-planar configuration.

To summarise the bow C-shock properties:

(1) The ratios of the column densities of the upper energy levels (CDRs) are plotted against the upper energy level to produce CDR curves. The curves define a particular family described by the maximum temperature, bow shape and cooling function. For velocities above ~ 50 km s^{-1} the CDRs are restricted to a narrow range determined by just the bow shape and cooling function. The latter is determined by the H_2O abundance and density. This holds provided that dissociation determines the breakdown condition which typically requires $B_{-3} \leqslant n_6^{1/2}$ and $X \geqslant 10^{-7}$.

(2) Parabolic bows with H_2O abundances of 10^{-4} - 5×10^{-4} yield CDRs consistent with observed line data. Hemispherical caps and planar C-shocks do not.

(3) Total line strengths are roughly proportional to the square of the paraboloid scale length, are not highly dependent on the individual parameters and are sufficient to cover the observed range.

We have assumed the density to be above $10^6 cm^{-3}$. For lower densities the cooling function and level populations rapidly diverge from the values calculated here. The above conclusions apply to all parameter ranges for which the ion-neutral streaming velocity is a large fraction of the shock velocity which fixes the ceiling to the breakdown velocity at ~ 49 km s^{-1} (Smith & Brand 1989a). This converts to an M_A greater than about 5 or $B_{-3} < 11.6\ n_6^{1/2}$ (v/40 km s^{-1}). Thus these conclusions apply to shocks propagating through what may prove to be quiescent cloud conditions. They are not necessarily applicable to regions of strong magnetic field.

REFERENCES

Brand, P.W.J.L., Moorhouse, A., Burton, M.G., Geballe, T.R., Bird, M. & Wade, R.,1988. Astrophys. J., 334, L103.

Brand, P.W.J.L., Toner, M.P., Geballe, T.R., Webster, A.S., Williams, P.M. & Burton, M.G., 1989a. Mon. Not. R. astr. Soc., 236, 929.

Brand, P.W.J.L., Toner, M.P., Geballe, T.R. & Webster, A.S., 1989b. Mon. Not. R. astr. Soc., 237, 1009.

Burton, M.G., Brand, P.W.J.L., Geballe, T.R. & Webster, A.S., 1989. Mon. Not. R. astr Soc., 236, 409.

Draine, B.T., 1980. Astrophys. J., 241, 1021.

Draine, B.T., Roberge, W.G. & Dalgarno, A., 1983. Astrophys. J., 264, 485.

Hartigan, P., Raymond, J. & Hartmann L., 1987. Astrophys. J., 316, 323.

Hollenbach, D. & McKee, C.F., 1989. Astrophys. J., 342, 306.

Langer, W.D., 1985. In Protostars and Planets, eds. Black, D.C. & Matthews, M.S., p650.

Moorhouse, A., Brand, P.W.J.L., Geballe, T.R. & Burton, M.G., 1990. Mon. Not. R. astr. Soc., submitted.

Raga, A.C., Mateo, M., Böhm, K.H. & Solf, J., 1988. Astr. J., 95, 1783.

Shull, J.B. & Draine, B.T., 1987. In *Interstellar Processes*, eds D. Hollenbach & H. Thronson, D. Reidel, p283

Smith, M.D. & Brand, P.W.J.L., 1990a. Mon. Not. R. astr. Soc., in press.

Smith, M.D. & Brand, P.W.J.L., 1990b. Mon. Not. R. astr. Soc., in press.

Smith, M.D. & Brand, P.W.J.L., 1990c. Mon. Not. R. astr. Soc., submitted.

Zinnecker, H., Mundt, R., Geballe, T.R. & Zealey, W.J., 1989. Astrophys. J., 342, 337.

FIGURE CAPTIONS

Fig. 1. Column density ratio (CDR) curves demonstrating the dependence on various parameters. The standard set of parameters given in Sub-section 3.1 are taken unless specifically changed. The upper energy level, expressed in units of 1000K, is shown on the x-axis.
(a) The imposed *breakdown velocity* (indicated in km s^{-1}) has a small influence on the results. Here $\epsilon = 0$ and $\epsilon_o = 4.25\ 10^{-4}$ are the H_2O abundances. The lower and upper velocity limits correspond to dissociated fractions of ~ 20% and ~ 45% respectively. Elsewhere we take the median values for v_b which correspond to ~ 33 ± 3% dissociated or the 49 km s^{-1} corresponding to streaming ionisation breakdown, whichever is lower.
(b) The *bow velocity* (in km s^{-1}) has no influence on the CDRs once above the breakdown velocity (48 km s^{-1} – corresponding to a maximum temperature of 3968K and a dissociated fraction of 36%).
(c) The CDRs are sensitive to the *ionization fraction and magnetic field* only if streaming ionization provides the breakdown limit. Here, the maximum temperatures (and total dissociated) in the shocks are 2908K (0.5%) and 2394K (0.01%) for $\chi = \chi_o/3$ and $B = 3B_o$ respectively. Also shown is the result for a hemispherical cap with the standard parameters.
(d) The *H_2O abundance* is a vital factor. Total dissociation of 33% yields breakdown speeds of 39 ($\epsilon = 0$), 40 ($\epsilon_o/16$), 42 ($\epsilon_o/4$) and 48 (ϵ_o) in km s^{-1}.

Fig. 2. A comparison of the CDR results (a) to the approximate model (b). (1) $\epsilon = \epsilon_o$, $v_b = 48$ km s^{-1}, $T_{max} = 3968$K and (2) $\epsilon = 0$, $v_b = 39$ km s^{-1}, $T_{max} = 3737$K.

Fig. 3. (a) The data for Peak 1 of OMC-1 from Brand et al (1988) superimposed on the standard C-type bow shock model with the two indicated H_2O abundances and (b) the data for IC 443 from Moorhouse et al. (1990), interpreted by a higher H_2O abundance model.

Fig. 4. CDRs for planar J-shocks of low velocity. The driving pressure determines the line ratios with nT going from (1) 10^{10} K cm^{-3} to (4) 10^{13} K cm^{-3} in jumps of 10.

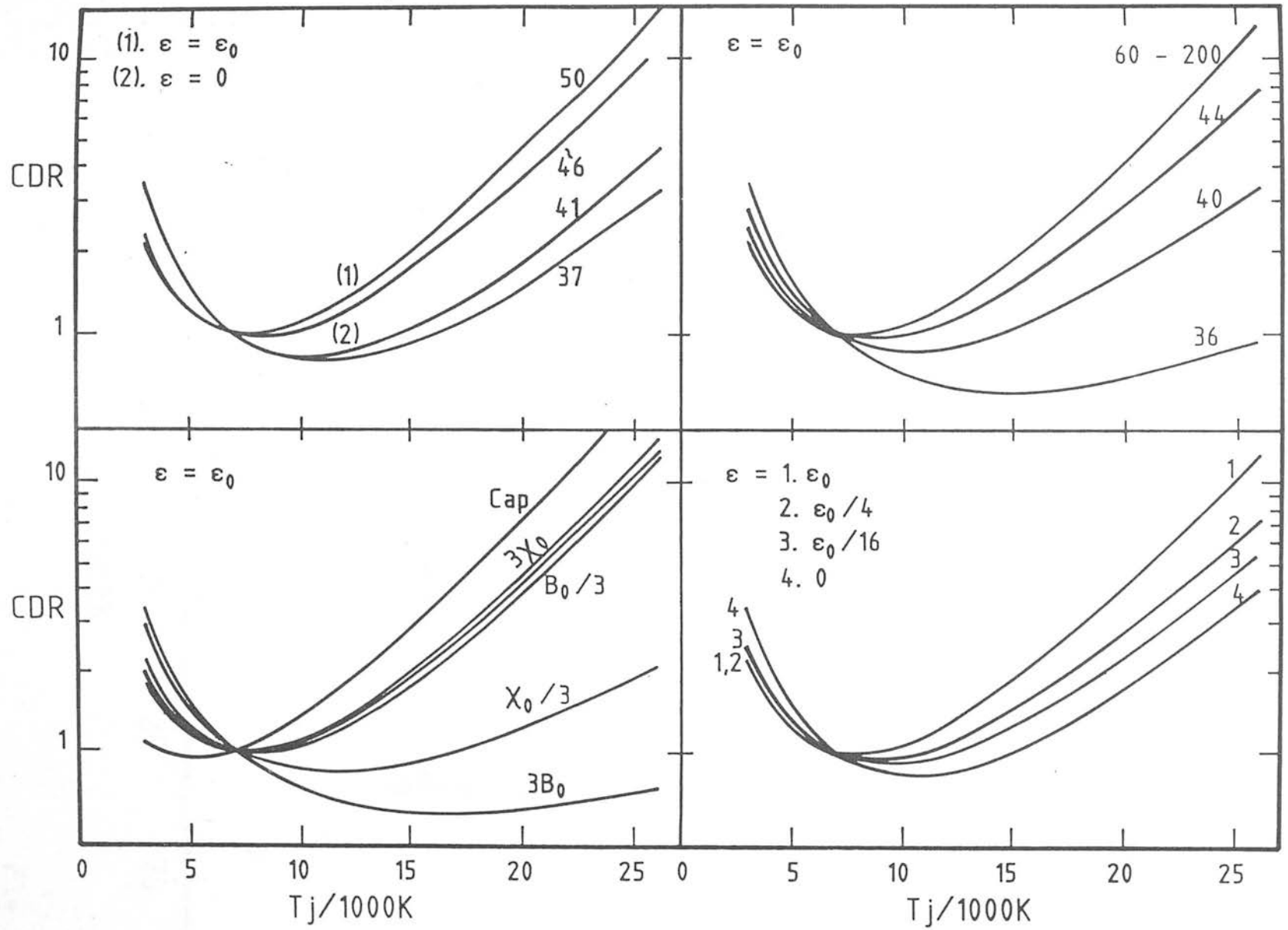


Fig. 1

Figure 1

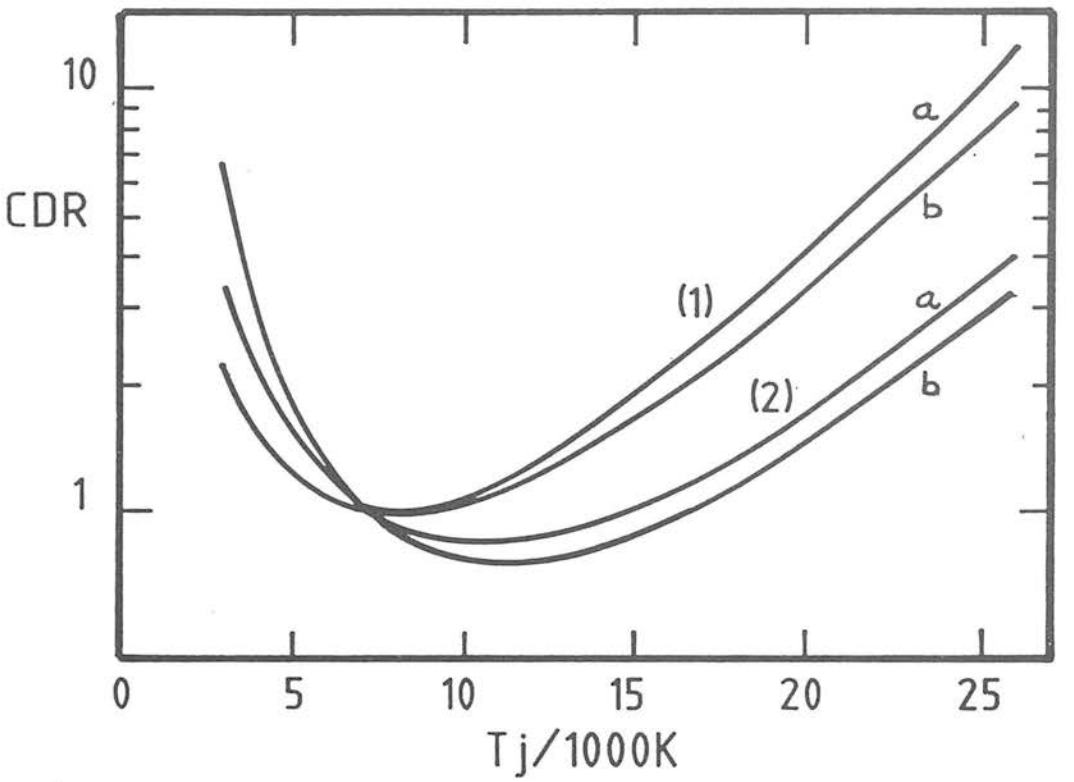


Fig. 2.

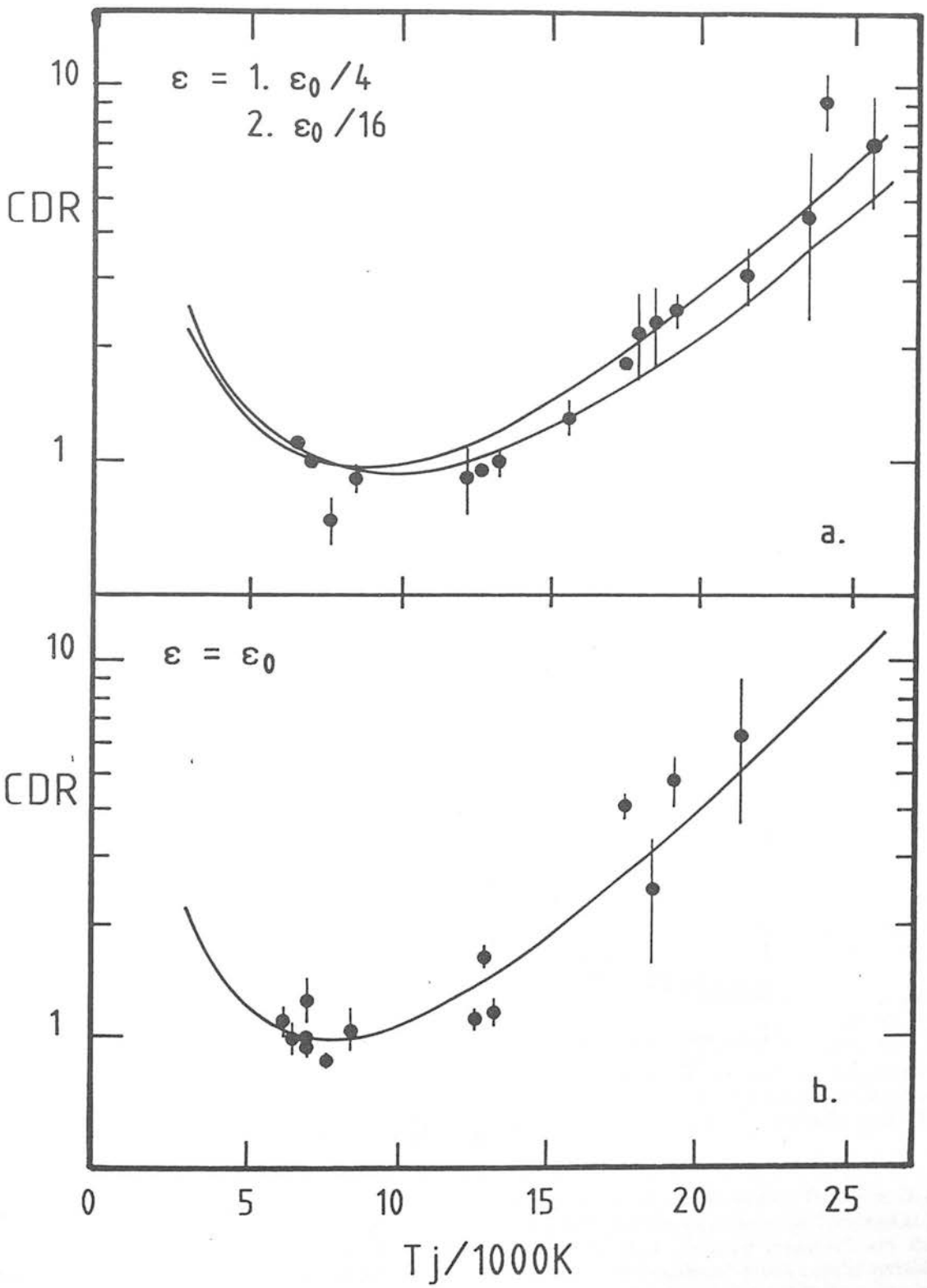


Fig. 3

Figure 2

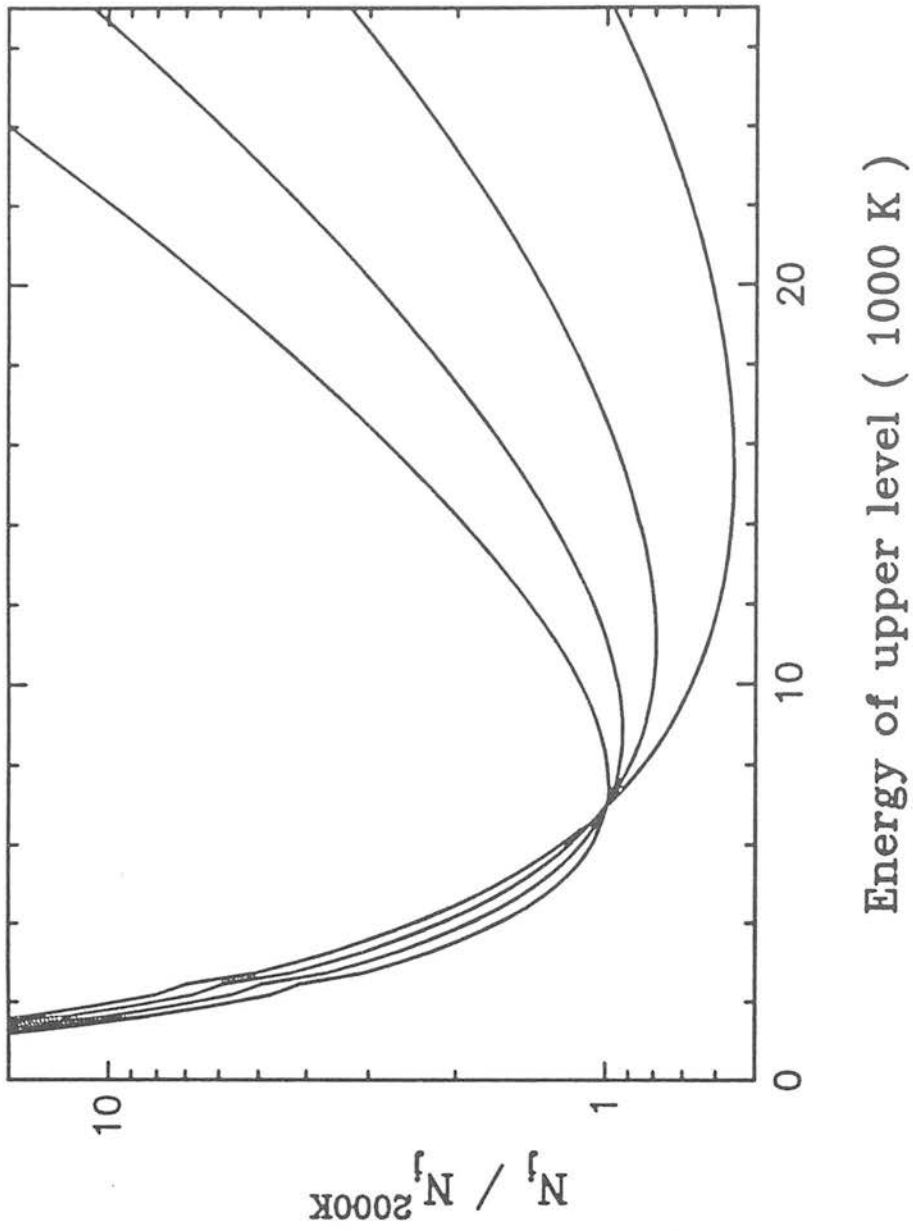


Fig. 4

THE ASTROPHYSICAL JOURNAL, 318:L77–L80, 1987 July 15

LEVEL POPULATION AND PARA/ORTHO RATIO OF FLUORESCENT H_2 IN NGC 2023

TETSUO HASEGAWA
Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo
IAN GATLEY
United Kingdom Infrared Telescope Unit of the Royal Observatory, Edinburgh
RON P. GARDEN AND PETER W. J. L. BRAND
Astronomy Department, University of Edinburgh
AND
MASATOSHI OHISHI, MASAHIKO HAYASHI, AND NORIO KAIFU
Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo
Received 1986 September 29; accepted 1987 April 17

ABSTRACT

Observations of vibrationally excited H_2 in 10 lines at 2.03–2.39 μm toward the reflection nebula NGC 2023 are presented and compared with the corresponding data in Orion-KL. Level populations of H_2 in NGC 2023 are characterized by a high vibrational temperature ($T_v > 3600$ K) and a low rotational temperature ($T_r \approx$ 900–1500 K). This confirms the identification of the H_2 emission as due to UV excitation followed by fluorescence. The population of the levels of H_2 in para form are found to be systematically larger than one-third of that of the ortho form. The estimated para/ortho ratio is 1/(1.4–2.0). The population distribution of H_2 in Orion-KL is expressed by a single curve representing an excitation temperature $T_v = T_r \approx 1600$–3300 K with no sign of departure of the para/ortho ratio from one to three. This is in accordance with earlier observations and theoretical predictions for shocked gas. The implications of the observed para/ortho ratio are discussed.

Subject headings: interstellar: molecules — molecular processes — nebulae: reflection

I. INTRODUCTION

A new aspect of the observational study of interstellar molecular hydrogen was opened by the discovery of the fluorescent infrared H_2 line emission in the reflection nebula NGC 2023 (Gatley and Kaifu 1987; Gatley *et al.* 1987) and successive detection of other fluorescent sources and the diffuse emission in the Orion Nebula (Hayashi *et al.* 1985) and some other reflection nebulae (Sellgren 1986; Tanaka *et al.* 1987). The fluorescent emission is expected to carry information on the physical processes occurring at the photodissociation regions, i.e., the surfaces of molecular clouds exposed to intense radiation field (e.g., Black and Dalgarno 1976; Tielens and Hollenbach 1985).

In this *Letter*, we present a comparative study of the excitation of molecular hydrogen in the fluorescent source NGC 2023 and Orion-KL, a prototypical source of shocked emission. To get enough information on the rotational and vibrational excitation of molecular hydrogen in para and ortho forms, we have observed 10 spectral lines of $v = 1$–0, $v = 2$–1, and $v = 3$–2 vibration-rotation transitions of H_2 which arise from energy levels 6471–19086 K above the ground state. A detailed level population of the fluorescent H_2 and its para/ortho ratio are revealed for the first time. To interpret the level population and the para/ortho ratio, a theoretical model has been developed by Takayanagi, Sakimoto, and Onda (1987), also appearing in this issue.

The reflection nebula NGC 2023 is excited by a B1.5 V star HD 37903. The visible nebula is dominated by the scattered light from HD 37903, and any H II region around this star is expected to be too small to be detected in the radio and infrared recombination lines (e.g., Knapp, Brown, and Kuiper 1975). Behind the nebula lies a relatively dense molecular cloud (e.g., Milman *et al.* 1975). Molecules at the front surface of the cloud are exposed to the light from the illuminating star and are excited and dissociated by it.

Orion-KL, on the other hand, is powered by the activity of a protostar embedded deeply in dense molecular gas. The high-velocity molecular outflow interacts with the ambient molecular gas and drives shocks. The level population of the H_2 there has been found to be consistent with the collisional excitation in the postshock gas (e.g., Beckwith *et al.* 1983).

II. OBSERVATIONS AND RESULTS

The observations were made on 1985 December 28 and 29 using the 3.8 m United Kingdom Infrared Telescope at Mauna Kea, Hawaii. The instrumentation employed was the photometer system UKT-9 equipped with a circular variable filter (CVF) in tandem with a Fabry-Pérot (FP) filter. The spectral resolution of the FP was 100 km s^{-1}, and its orders were sorted by the CVF. The FP and the CVF were controlled simultaneously to scan sequentially across a number of spectral lines skipping the wavelengths between the lines, so that

TABLE 1

OBSERVED H_2 LINE INTENSITIES AND COLUMN DENSITIES

			NGC 2023 80″S			ORION-KL PEAK 1			
	WAVELENGTH	E_u/k	Flux per Beam[a]		N_u	Flux per Beam[a]		N_u Unreddened[b]	
LINE	(μm)	(K)	($\times 10^{-20}$ W cm^{-2})		($\times 10^{15}$ cm^{-1})	($\times 10^{-18}$ W cm^{-2})		($\times 10^{17}$ cm^{-2})	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
$v = 1$–0 $S(0)$	2.2233	6471	2.6	(8%)	2.0	0.435	(2%)	2.71	14.2
$v = 1$–0 $S(1)$	2.1218	6956	5.5	(4%)	3.0	1.67	(2%)	7.97	45.3
$v = 1$–0 $S(2)$	2.0338	7584	2.3	(5%)	1.05	0.49	(20%)	2.2	13.1
$v = 2$–1 $S(0)$	2.3556	12095	0.8	(30%)	0.46	0.031	(40%)	0.125	0.60
$v = 2$–1 $S(1)$	2.2477	12550	1.5	(15%)	0.60	0.163	(4%)	0.51	2.62
$v = 2$–1 $S(2)$	2.1542	13150	1.1	(10%)	0.38	0.060	(10%)	0.175	0.97
$v = 2$–1 $S(3)$	2.0736	13890	1.2	(12%)	0.38	0.140	(4%)	0.414	2.45
$v = 3$–2 $S(1)$[c]	2.3866	17818	0.6	(50%)	0.25	0.042	(15%)	0.12	0.56
			0.4	(100%)	0.17	0.018	(40%)	0.05	0.24
$v = 3$–2 $S(2)$	2.2874	18386	0.3	(70%)	0.11	0.022	(20%)	0.060	0.30
$v = 3$–2 $S(3)$	2.2021	19086	0.6	(17%)	0.21	0.031	(15%)	0.087	0.47

[a] Numbers in parentheses are the uncertainties estimated from the statistical noise and the repeatability of the observations. In Orion-KL, the intensities and the column densities are systematically underestimated by about 30% (see § II).

[b] Correction is applied for the estimated extinction of 5×10^{-4} mag cm^{-1} (col. [9]; Beckwith *et al.* 1983) and 9×10^{-4} mag cm^{-1} (col. [10]; Davis, Larson, and Hofmann 1986) assuming the λ^{-1} reddening law.

[c] Identification is ambiguous (see § II).

all the lines were observed in an entire scan through 2.03–2.39 μm. The entire scans were repeated until a good enough signal-to-noise ratio was achieved for the weaker lines. This method enabled us to measure the relative intensities of the lines accurately. For each of the lines, 21 spectral points were observed to cover ± 0.006 μm centered on the line. For the weaker three lines at the longest wavelengths, the spectral coverages of the scans were doubled to get better estimates of the baselines. The flux density scale has been calibrated by observations of a standard star HD 26846 (K3 III, $m_K = 2.28$ mag).

Ten spectral lines listed in Table 1 were selected for observations. The measurements were made toward 80″ south of HD 37903, the exciting star of NGC 2023, and the Peak 1 of the H_2 emission in Orion-KL (Beckwith *et al.* 1978). The aperture was 19″.6 in diameter, and the sky chopping was made with a throw of 3′ east-west.

The measured line intensities are listed in Table 1. Eight and nine out of the 10 lines have been firmly detected in NGC 2023 and in Orion-KL, respectively. The identification of the $v = 3$–2 $S(1)$ line is ambiguous because two lines separated by 0.001 μm have been detected very close to its nominal wavelength, so that we list in Table 1 the intensities of both the stronger line at the longer wavelength and the weaker line at the shorter wavelength. Although the spacing between the $v = 1$–0 $S(0)$ and the $v = 2$–1 $S(1)$ lines is close to the order spacing of the FP, a careful test proved that there is no aliasing problem in our results. The broad lines in Orion-KL are partially resolved by the 100 km s^{-1} FP, so that the flux observed through the FP centered on the line is about 70% of the total flux. As our main interest is on their relative intensities, we will not correct for this.

The column densities of H_2 in the upper energy levels of the transitions, N_u, are derived as listed in Table 1 from the measured intensities, I (= flux per beam) through the equation,

$$N_u = (4\pi I)/(Ah\nu\Omega), \quad (1)$$

where A and ν are the Einstein coefficient and the frequency of the transition, respectively, and Ω is the beam solid angle. Correction for the reddening is applied for the Orion-KL data based on the two recent estimates of the extinction; 5×10^{-4} mag cm^{-1} (Beckwith *et al.* 1983) and 9×10^{-4} mag cm^{-1} (Davis, Larson, and Hofmann 1986). In both cases the λ^{-1} reddening law is assumed. In NGC 2023, the reddening may be far smaller and is not corrected for.

III. DISCUSSION

Figure 1 shows the level population of H_2 deduced from the present observations. The column densities, N_u, have been divided by the statistical weights, g_u, which include the rotational ($g_J = 2J + 1$) and the spin ($g_s = 1$ for even J and $g_s = 3$ for odd J) degeneracies. Then they are scaled to the population of the $v = 1$, $J = 3$ level. The slopes of the lines in the figure are direct measures of the excitation temperatures.

Comparison of the population in NGC 2023 with that in Orion-KL readily shows three remarkable points as follows:

1. NGC 2023 shows two separate sequences of energy levels corresponding to the para and ortho forms of H_2 (Fig. 1*a*). As the statistical weight applied in Figure 1 includes the spin degeneracies, the para and ortho levels should align on a single sequence if the para/ortho abundance ratio is 1/3, as is indeed the case for Orion-KL (Fig. 1*b*). The separation between the two sequences in the $v = 1$ and $v = 2$ states corresponds to the para/ortho ratio of 1/2.0 and 1/1.4, respectively.

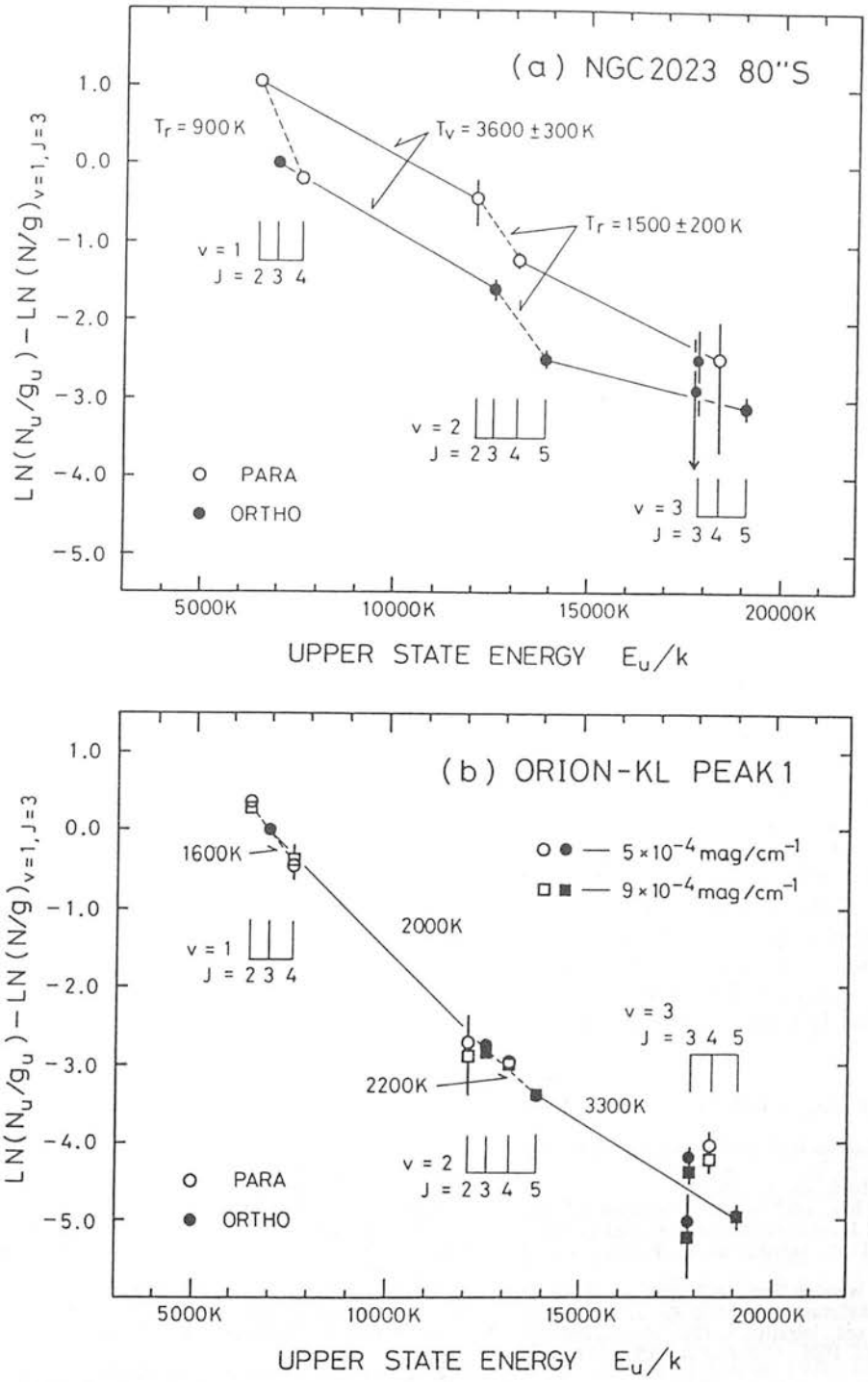


FIG. 1.—The level population of H_2 normalized to the $v = 1$ $J = 3$ state in NGC 2023 (*a*) and Orion-KL (*b*). For Orion-KL, reddening correction has been applied based on the two estimates indicated in the upper right-hand corner. Excitation temperatures measured from the slopes of the lines are shown.

2. Both of the two sequences of the level population in NGC 2023 are characterized by high vibrational excitation temperatures, T_v, and low rotational excitation temperatures, T_r. This is in contrast to the case of Orion-KL, in which the levels at higher energy are characterized by higher excitation temperatures without any systematic difference between the vibrational and rotational temperatures; this trend is in good agreement with the result of Beckwith *et al.* (1983).

3. The two sequences of population in NGC 2023 show similar excitation characteristics. Both the populations of para-H_2 and ortho-H_2 are consistent with the same $v = 2$–1 vibrational temperatures and the same rotational tempera-

tures in the $v = 2$ state. They may also be consistent with the same $v = 3$–2 vibrational temperatures. This suggests that the para/ortho ratio measured from the separation of the two sequences on Figure 1*a* (point [1] above) may represent the ratio of the total abundance of para-H_2 and ortho-H_2.

The high vibrational temperature and the low rotational temperature found in NGC 2023 are expected from theoretical calculations by Black and Dalgarno (1976), Black and van Dishoeck (1987), and Takayanagi, Sakimoto, and Onda (1987). Especially, the last model can reproduce the level population in Figure 1*a* fairly well. This strongly supports the identification by Gatley *et al.* (1987) of the H_2 emission in NGC 2023 as due to the fluorescent cascade after absorption of a UV photon. No effect to selectively excite para-H_2 or ortho-H_2 is found in this scheme. This is consistent with the similar excitation temperatures for the two forms (point [3] above). The observed column densities of H_2 in vibrationally excited states (Table 1) are in rough agreement with theoretical expectations by Shull (1978) and Tielens and Hollenbach (1985), lending another support to the UV excitation of the observed emission.

What is reflected in the observed para/ortho ratio? One of the possible ways to interpret the observational result is outlined as follows: The para/ortho ratio is first established at the formation of H_2, reflecting probably the physical conditions of the H_2 formation (e.g., the grain temperature). Although radiative transitions cannot cause interchanges between the para and ortho forms, they are possible in exchange reactions of H_2 with H^+, H, and H_3^+. The reaction with protons produced through cosmic-ray ionization can slowly alter the para/ortho ratio of H_2 in dense molecular clouds (Dalgarno, Black, and Weisheit 1973; Flower and Watt 1984). Finally, the ratio may change in the regions of H_2 emission due to various processes. In the case of Orion-KL, the equilibrium para/ortho ratio of 1/3 may be attained through rapid reactions in the hot ($\sim$ 2000 K), dense postshock gas (Chernoff, Hollenbach, and McKee 1982; Draine, Roberge, and Dalgarno 1983), and the memory of the para/ortho ratio in the past may be lost. In the photodissociation region of NGC 2023, on the other hand, the para/ortho ratio may change through the exchange reactions of H_2 with abundant H and possibly with H^+ and H_3^+ along with H_2 formation in the warm H_2/H I transition layer where the H_2 fluorescence arises (Tielens and Hollenbach 1985; Takayanagi, Sakimoto, and Onda 1987). It is not clear at present whether the ratio observed in NGC 2023 retains any trace of the para/ortho ratio of H_2 in the molecular cloud behind the nebula or if it is a pure reflection of the local processes as in Orion-KL.

The model by Takayanagi, Sakimoto, and Onda (1987) can reproduce the observed para/ortho ratio in NGC 2023 by assuming the H_2 being formed in this region with the para/ortho ratio expressed by a "formation temperature," $T_f \approx 60$–70 K. It is noteworthy that the para/ortho ratio of H_2 in diffuse clouds observed in UV absorption is expressed by $T_{01} = 77 \pm 17$(rms) K (Savage *et al.* 1977), a value close to T_f. Measurements of the para/ortho ratio in sources of various physical conditions and developments of more detailed models would contribute to the understandings of the physical and chemical processes during the formation and evolution of interstellar clouds.

We are grateful to the staff of UKIRT for their development of the excellent instrumentation and efficient support during the observations. We have benefited from helpful discussions with Drs. K. Takayanagi, K. Sakimoto, K. Onda, J. Bally, J. Black, E. van Dishoeck, and A. Tielens. This work was done under a UK-Japanese collaboration supported by SERC of the United Kingdom and the Japanese Ministry of Education.

REFERENCES

Beckwith, S., Evans, N. J., II, Gatley, I., Gull, G., and Russel, R. W. 1983, *Ap. J.*, **264**, 152.
Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978, *Ap. J.*, **223**, 464.
Black, J. H., and Dalgarno, A. 1976, *Ap. J.*, **203**, 132.
Black, J. H., and van Dishoeck, E. F. 1987, in *IAU Symposium 115, Star Forming*, ed. M. Peimbert and J. Jugaku (Dordrecht: Reidel), p. 139.
Chernoff, D. F., Hollenbach, D. J., and McKee, C. F. 1982, *Ap. J. (Letters)*, **259**, L97.
Dalgarno, A., Black, J. H., and Weisheit, J. C. 1973, *Ap. Letters*, **14**, 77.
Davis, D. S., Larson, H. P., and Hofmann, R. 1986, *Ap. J.*, **304**, 481.
Draine, B. T., Roberge, W. G., and Dalgarno, A. 1983, *Ap. J.*, **264**, 485.
Flower, D. R., and Watt, G. D. 1984, *M.N.R.A.S.*, **209**, 25; erratum 1985, *M.N.R.A.S.*, **213**, 991.
Gatley, I., *et al.* 1987, *Ap. J. (Letters)*, **318**, L73.
Gatley, I., and Kaifu, N. 1987, in *IAU Symposium 120, Astrochemistry*, ed. M. S. Vardya and S. T. Tarafdar (Dordrecht: Reidel), p. 153.
Hayashi, M., Hasegawa, T., Gatley, I., Garden, R., and Kaifu, N. 1985, *M.N.R.A.S.*, **215**, 31P.
Knapp, G. R., Brown, R. L., and Kuiper, T. B. H. 1975, *Ap. J.*, **196**, 167.
Milman, A. S., Knapp, G. R., Kerr, F. J., Knapp, S. L., and Wilson, A. 1975, *A.J.*, **80**, 93.
Savage, B. D., Bohlin, R. C., Drake, J. F., and Budich, W. 1977, *Ap. J.*, **216**, 291.
Sellgren, K. 1986, *Ap. J.*, **305**, 399.
Shull, J. M. 1978, *Ap. J.*, **219**, 877.
Takayanagi, K., Sakimoto, K., and Onda, K. 1987, *Ap. J. (Letters)*, **318**, L81.
Tanaka, M., *et al.* 1987, in preparation.
Tielens, A. G. G. M., and Hollenbach, D. 1985, *Ap. J.*, **291**, 722.

PETER W. J. L. BRAND: Astronomy Department, Edinburgh University, Blackford Hill, Edinburgh EH9 EHJ, Scotland

ROGNVALD P. GARDEN: Space Science Laboratory, University of California, Berkeley, CA 94720

IAN GATLEY: National Optical Astronomy Observatories, 950 North Cherry Avenue, P. O. Box 26732, Tucson, AZ 85726-6732

TETSUO HASEGAWA, MASAHIKO HAYASHI, NORIO KAIFU, and MASATOSHI OHISHI: Nobeyama Radio Observatory, Tokyo Astronomical Observatory, Nobeyama, Minamisaku, Nagano 384-13, Japan

THE ASTROPHYSICAL JOURNAL, 318:L73–L76, 1987 July 15

FLUORESCENT MOLECULAR HYDROGEN EMISSION FROM THE REFLECTION NEBULA NGC 2023

IAN GATLEY
United Kingdom Infrared Telescope Unit of the Royal Observatory Edinburgh
TETSUO HASEGAWA AND HIROKO SUZUKI
Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo
RON GARDEN AND PETER BRAND
Astronomy Department, University of Edinburgh
JOHN LIGHTFOOT AND WILLIAM GLENCROSS
Department of Physics and Astronomy, University College London
HARUYUKI OKUDA
Astrophysics Division, Institute of Space and Astronautical Sciences, Tokyo
AND
TETSUYA NAGATA
Institute for Astronomy, University of Hawaii
Received 1986 September 29; accepted 1987 April 17

ABSTRACT

The near-infrared spectrum of NGC 2023 is shown to be rich with emission lines of vibrationally excited molecular hydrogen. The relative intensities of these lines are in excellent agreement with theoretical predictions for ultraviolet pumped fluorescence.

A map of NGC 2023 in the $v = 1–0$ $S(1)$ line of H_2 shows that the fluorescent emission originates from a thin shell, of radius ~ 0.3 pc, which surrounds the central star of the reflection nebula. Within this shell the hydrogen is radiatively dissociated. Density variations within the nebula cause the shell to appear nonspherical, broken, and clumpy. A map of NGC 2023 in ^{12}CO ($J = 1–0$) reveals a bright rim of emission lying immediately outside the brightest portion of the fluorescent shell, and so demonstrates the density variations directly.

Maps of NGC 2023 at wavelengths of 1.65 μm (H) and 2.2 μm (K) confirm the existence of bright continuum emission, which Sellgren has attributed to thermal emission from very small grains (of radius ~ 10 Å), heated briefly to temperatures $\sim$ 1000 K by absorption of individual ultraviolet photons. The shell of fluorescent H_2 lies immediately outside Sellgren's nebula, at those radii where attenuation of the stellar radiation field (by dust and by self-shielding of H_2) becomes significant, in good agreement with theory.

Subject headings: interstellar: molecules — molecular processes — nebulae: reflection

I. INTRODUCTION

The ultraviolet irradiation of an interstellar cloud of molecular hydrogen has been considered in considerable theoretical detail. Radiation in the wavelength interval 912–1108 Å is absorbed in the Lyman and Werner band systems of the molecule. Then, in a mechanism first suggested by Solomon (see Field, Somerville, and Dressler 1966), fluorescence leads either to dissociation (Stecher and Williams 1967; Nishimura and Takayanagi 1969; Dalgarno and Stephens 1970) or to the population of excited vibration-rotation levels (Gould and Harwit 1963; Black and Dalgarno 1976; Shull 1978; Black and van Dishoeck 1987). Approximately one in 10 absorptions leads to dissociation, and, in steady state, dissociation is balanced by formation of molecules on grain surfaces (Tielens and Hollenbach 1985).

The remainder of the fluorescent gas radiates. The homonuclear hydrogen molecule has no permanent dipole moment, and so the excited states decay by electric quadrupole transitions, producing emission lines in the infrared. Early searches for emission lines of H_2 were based on the predictions for fluorescence (Werner and Harwit 1968; Gull and Harwit 1971).

Infrared emission from vibrationally excited H_2 was first observed in Orion OMC1 (Gautier *et al.* 1976) and in NGC 7027 (Treffers *et al.* 1976). Measurements of the relative intensities of the lines in Orion showed that the gas was thermally excited (Gautier *et al.* 1976; Grasdalen and Joyce 1976; Beckwith *et al.* 1978), and successful models based on shock excitation were developed (Hollenbach and Shull 1977; Kwan 1977; London, McCray, and Shull 1977). Many further examples of shock-excited sources of vibrationally excited H_2 have been discovered (Shull and Beckwith 1982), and the detection of an infrared emission line of H_2 has come to be regarded as a routine diagnostic for the presence of shocks.

Yet theory predicts that ultraviolet pumped fluorescence should be a common phenomenon. In particular, it should

occur on the surfaces of molecular clouds illuminated by young massive stars, producing large sheets of diffuse emission. The techniques of infrared astronomy are typically not well suited to the detection of diffuse emission, because they involve sky-chopping. We have used a frequency switching method and a large beam size to optimize our experiment for the detection of diffuse emission.

In this *Letter* we report observations of the reflection nebula NGC 2023 which demonstrate the detection of ultraviolet pumped fluorescent molecular hydrogen. The agreement between theory and experiment is excellent.

II. OBSERVATIONS

NGC 2023 is a conspicuous reflection nebula in the molecular cloud L1630 created by the B1.5 star HD 37903. Far-infrared observations show directly that much of the radiation from the star is absorbed by the molecular cloud within a radius of only a few arcminutes (Harvey, Thronson, and Gatley 1980). Diffuse near-infrared continuum radiation from NGC 2023 has been attributed to thermal emission from very small grains (of radius ~ 10 Å) heated briefly to temperatures $\sim$ 1000 K by absorption of individual ultraviolet photons (Sellgren, Werner, and Dinerstein 1983; Sellgren 1984). A strong dust emission feature is observed at 3.3 μm; this feature may be of fluorescent origin and be produced when polycyclic aromatic hydrocarbons (PAHs) are irradiated by ultraviolet light (Leger and Puget 1984; Sellgren *et al.* 1985). Therefore NGC 2023 is an obvious target in which to search for fluorescent H_2.

The infrared observations were made at the United Kingdom Infrared Telescope (U.K.I.R.T.), Mauna Kea, Hawaii, in 1984 August and 1985 December. The standard observatory photometers UKT6 and UKT9 were used; these consist of a single solid-nitrogen–cooled InSb photovoltaic detector with cooled filters, including a circular variable filter wheel (CVF) of $\sim 1\%$ spectral resolution. The beam size was chosen to be 19″.6, the largest available.

In order to search for diffuse H_2 emission, an ambient temperature tunable Fabry-Perot interferometer was placed directly in the f/35 beam of the telescope ahead of the photometer. The CVF was set to the wavelength of the $v = 1$–0 $S(1)$ line of H_2 at 2.122 μm, both to act as an order sorter for the Fabry-Perot and to restrict the background radiation falling on the detector.

Wavelength and flux calibration were established by observations of NGC 7027; the Fabry-Perot was scanned across the H_2 line while measuring the DC signal from the detector. The resolution of the Fabry-Perot was measured to be 130 km s^{-1}. In this configuration the system proved to be detector noise–limited, as expected, with a noise equivalent line strength of $\sim 6 \times 10^{-21}$ W cm^{-2}.

For the observations of NGC 2023 the Fabry-Perot was switched at 1 Hz between the wavelength of the H_2 line and a reference wavelength 300 km s^{-1} shorter; the signal difference was measured. Sky spectra, taken by the method used for NGC 7027, demonstrated that this method is satisfactory; the reference wavelength was chosen to avoid strong atmospheric emission features.

A map of the $v = 1$–0 $S(1)$ line of H_2 in NGC 2023 was constructed by rastering the telescope under computer control to successive offset positions on a 20″ spatial grid. The total integration time at each position was 8 s. Near real-time monitoring of the result allowed the area of the map to be extended until the edges of the source were reached.

The map of the H_2 line emission establishes suitable reference positions for sky-chopping measurements. Low-resolution spectra were measured with the CVF at positions 80″ south and 160″ north of the central star HD 37903 using a sky separation of 180″ east-west between the beams produced by the oscillating secondary mirror of the telescope. These spectra cover the wavelength interval 1.4–2.5 μm, that is, the whole of the so-called H and K windows. At Mauna Kea the atmospheric transmission is nonzero at wavelengths between these windows, and so measurements were included across the whole CVF range for the position 80″ south. The results obtained in the interval 1.8–2.0 μm should be viewed with caution.

In order to compare the illumination of the dust in NGC 2023 with that of the molecular gas three more infrared maps were constructed, at wavelengths of 1.65 μm (H), 2.2 μm (K), and 3.3 μm; the H and K maps were measured through the standard photometric filters, and the 3.3 μm map, through the CVF. The CVF was tuned on source to the peak wavelength of the bright 3.3 μm dust emission feature. In each case the telescope was rastered as before, but the size of region to be mapped was determined from inspection of the H_2 result. Each map was made in two pieces, one to the east of the central star and the other to the west, again with a beam separation of 180″ east-west. The maps were made without beamswitching; the reference positions were to the east of the eastern map and to west of the western map, in order to minimize the possibility of chopping onto the source.

A map of the ^{12}CO ($J = 1$–0) emission from NGC 2023 was made using the 45 m telescope of the Nobeyama Radio Observatory,[1] Japan. Spectra were measured using a high-resolution acousto-optical radiospectrometer (AOS) with a velocity resolution of 0.1 km s^{-1} on a 20″ spatial grid with a beam size of 16″ throughout the region of brightest H_2 emission.

III. RESULTS

Figure 1 (Plate L1) shows the map of NGC 2023 in the $v = 1$–0 $S(1)$ line of H_2. There is diffuse emission over a region $\sim 5'$ in diameter, corresponding to a physical size of ~ 0.6 pc at the distance of NGC 2023, which is ~ 450 pc (Racine 1968; de Boer 1983). The brightest emission forms a broken, clumpy, and irregular shell.

Figure 2 shows 1.4–2.5 μm spectra measured at bright positions on this shell, located 80″ south and 160″ north of the central star. Despite their low resolution these spectra show a considerable number of emission lines of H_2, identifiably arising from states as high as $v = 6$. The significant

[1] Nobeyama Radio Observatory, a branch of Tokyo Astronomical Observatory, University of Tokyo, is a facility open for general use by researchers in the field of astronomy and astrophysics.

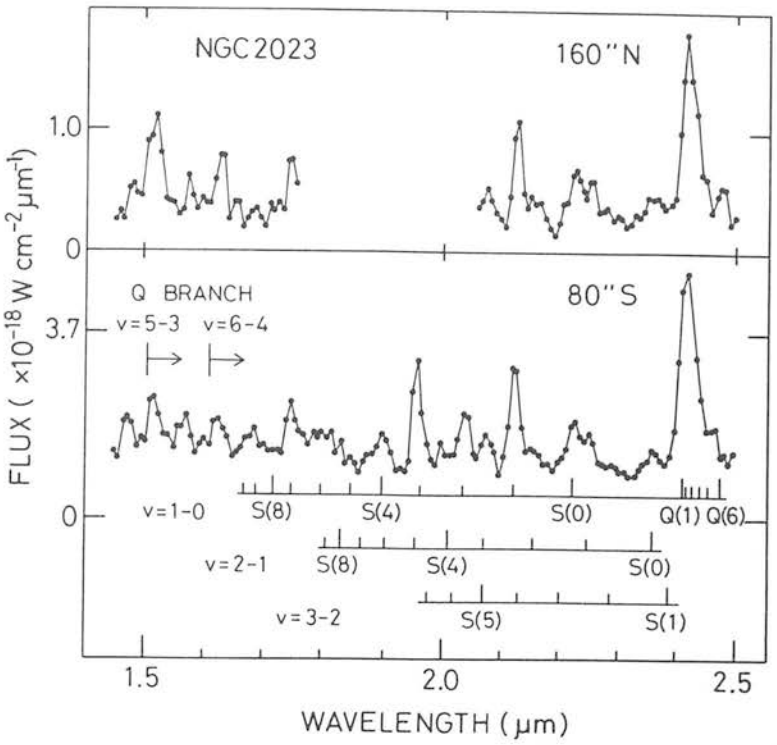


FIG. 2.—CVF spectra measured at positions 160" north (*upper*) and 80" south (*lower*) of the star HD 37903. The wavelengths of some emission lines of H_2 are indicated.

continuum emission (see Fig. 4, below) and large number of lines present in these spectra makes it difficult to determine accurately the strength of a particular line at this resolution. Measurements made with the Fabry-Perot interferometer are presented in the paper which follows (Hasegawa *et al.* 1987). Black and van Dishoeck (1987) have synthesized model spectra matched to the resolution of the present experiment under the assumption that the H_2 undergoes ultraviolet pumped fluorescence; the agreement between experiment and theory is excellent.

As Black and van Dishoeck (1987) have demonstrated, fluorescent H_2 is detected most straightforwardly by measurement of the lines at wavelengths less than 2.0 μm. Most infrared spectroscopy of H_2 to date has been performed at wavelengths longer than 2.0 μm; a signature of fluorescence in this wavelength region is provided by the strength of the emission in $v = 1$–0 $S(0)$ and $v = 2$–1 $S(1)$. These lines are more conspicuous in the fluorescent case than in the shocked, because the rotational temperature is lower and the vibrational temperature higher when the gas is radiatively excited than when it is shocked; an example is shown by Gatley and Kaifu (1987).

Figure 3 (Plate L2) shows a map of NGC 2023 in ^{12}CO ($J = 1$–0) in the velocity interval 10–11 km s^{-1}. This gas is redshifted by ~ 1 km s^{-1} from the velocity of the quiescent cloud (T. Hasegawa *et al.*, private communication); the emission forms a bright rim immediately outside the brightest portion of the fluorescent H_2 shell.

Figure 4 (Plate L3) shows maps of NGC 2023 at 1.65 μm (H), 2. μm (K), and at the wavelength of the dust emission feature at 3.3 μm. The emission due to the central star HD 37903 has been subtracted from the maps at H and K in order to display the diffuse emission clearly; the 3.3 μm map is uncorrected for the emission from the star. The diffuse near-infrared continuum is more conspicuous at K than at H, in agreement with the results of Sellgren (1984). Comparison of Figures 1 and 4 shows that the shell of fluorescent H_2 lies precisely at the edge of Sellgren's nebula. The striking difference in appearance between the line and continuum maps shows clearly that they arise from different physical origins, and that, although the contribution from H_2 lines contaminates the broad-band fluxes, there is distinct and significant continuum emission of the type described by Sellgren (1984) and by Witt, Schild, and Kraiman (1984).

The 3.3 μm map is also very different from the H and K maps. This is interesting because all measure dust emission. The emitting grains which form Sellgren's nebula are not responsible for a large fraction of the emission in the 3.3 μm feature. Leger and Puget (1984) have discussed the difficulties inherent in the attempt to assign a common origin to both the dust feature and the continuum emission based on the theory of PAHs. The similarity of the 3.3 μm map to that of H_2 suggests either that the carrier of the 3.3 μm feature is absent or depleted within the volume where the H_2 is dissociated, or that conditions favorable to the production of the 3.3 μm feature occur where the H_2 is excited.

IV. DISCUSSION

The observations presented in this *Letter* give a simple picture of the ultraviolet irradiation of an interstellar cloud. Near the star the intense radiation field both destroys the molecular hydrogen and heats the dust. The very small grains produce a near-infrared emission nebula, while larger grains radiate away the bulk of the power from the star in the far-infrared (Harvey, Thronson, and Gatley 1980). At the interface between the dissociated region and the quiescent molecular cloud, where the radiation field is severely attenuated by the dust, H_2 continues to be excited and destroyed by ultraviolet radiation.

Radiative dissociation is very rapid following the turn-on of the star; for an assumed density of 10^4 cm^{-3} (Harvey, Thronson, and Gatley 1980), in the absence of attenuation by dust, a dissociation sphere of the observed size develops in only 10^4 yr. The observed strength of the H_2 emission suggests, however, that only $\sim 1\%$ of the stellar ultraviolet flux is presently absorbed in the molecular gas; the predicted rate of growth of the dissociation sphere is therefore slowed by the dust to a modest speed, less than 1 km s^{-1}. The fresh material may be supplied either from the quiescent cloud or by formation of H_2 on grain surfaces (Tielens and Hollenbach 1985). It is interesting to note that the CO structure shown in Figure 3 is redshifted from the velocity of the quiescent cloud by 1 km s^{-1}, and that the bulk of the molecular material is believed to lie behind the exciting star (Harvey, Thronson, and Gatley 1980).

Measurements of the relative strengths of the H_2 emission lines in NGC 2023 made with higher spectral resolution (Hasegawa *et al.* 1987) and theoretical interpretation of those results (Takayanagi, Sakimoto, and Onda 1987) are presented elsewhere in this issue.

We thank Professor K. Takayanagi, Norio Kaifu, John Black, Ewine van Dishoeck, and Kris Sellgren for their enthusiastic help with this work.

REFERENCES

Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978, *Ap. J.*, **223**, 464.
Black, J. H., and Dalgarno, A. 1976, *Ap. J.*, **203**, 132.
Black, J. H., and van Dishoeck, E. F. 1987, in *IAU Symposium 115, Star Forming Regions*, ed. M. Peimbert and J. Jugaku (Dordrecht: Reidel), p. 139.
Dalgarno, A., and Stephens, T. L. 1970, *Ap. J. (Letters)*, **160**, L107.
de Boer, K. S. 1983, *Astr. Ap.*, **125**, 258.
Field, G. B., Somerville, W. B., and Dressler, K. 1966, *Ann. Rev. Astr. Ap.*, **4**, 207.
Gatley, I., and Kaifu, N. 1987, in *IAU Symposium 120, Astrochemistry*, ed. M. S. Vardya and S. P. Tarafdar (Dordrecht: Reidel), p. 153.
Gautier, T. N., Fink, U., Treffers, R. R., and Larson, H. P. 1976, *Ap. J. (Letters)*, **207**, L129.
Gould, R. J., and Harwit, M. 1963, *Ap. J.*, **137**, 694.
Grasdalen, G. L., and Joyce, R. R. 1976, *Bull. AAS*, **8**, 349.
Gull, T. R., and Harwit, M. O. 1971, *Ap. J.*, **168**, 15.
Harvey, P. M., Thronson, H. A., Jr., and Gatley, I. 1980, *Ap. J.*, **235**, 894.
Hasegawa, T., Gatley, I., Garden, R. P., Brand, P. W. J. L., Ohishi, M., Lightfoot, J. H., Hayashi, M., and Kaifu, N. 1987, *Ap. J. (Letters)*, **318**, L77.
Hollenbach, D. J., and Shull, J. M. 1977, *Ap. J.*, **216**, 419.
Kwan, J. 1977, *Ap. J.*, **216**, 713.
London, R., McCray, R., and Chu, S.-I. 1977, *Ap. J.*, **217**, 442.
Leger, A., and Puget, J. L. 1984, *Astr. Ap.*, **137**, L5.
Nishimura, S., and Takayanagi, K. 1969, *Pub. Astr. Soc. Japan*, **21**, 111.
Racine, R. 1968, *A.J.*, **73**, 233.
Sellgren, K. 1984, *Ap. J.*, **277**, 623.
Sellgren, K., Allamandola, L., Bregman, J., Werner, M., and Wooden, D. 1985, *Ap. J.*, **299**, 416.
Sellgren, K., Werner, M. W., and Dinerstein, H. L. 1983, *Ap. J. (Letters)*, **271**, L13.
Shull, J. M. 1978, *Ap. J.*, **219**, 877.
Shull, J. M., and Beckwith, S. 1982, *Ann. Rev. Astr. Ap.*, **20**, 163.
Stecher, T. P., and Williams, D. A. 1967, *Ap. J. (Letters)*, **149**, L29.
Takayanagi, K., Sakimoto, K., and Onda, K. 1987, *Ap. J. (Letters)*, **318**, L81.
Tielens, A. G. G. M., and Hollenbach, D. 1985, *Ap. J.*, **291**, 747.
Treffers, R. R., Fink, U., Larson, H. P., and Gautier, T. N. 1976, *Ap. J.*, **209**, 793.
Werner, M. W., and Harwit, M. 1968, *Ap. J.*, **154**, 881.
Witt, A. N., Schild, R. E., and Kraiman, J. B. 1984, *Ap. J.*, **281**, 708.

PETER W. J. L. BRAND: Astronomy Department, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

ROGNVALD P. GARDEN: Space Science Laboratory, University of California, Berkeley, CA 94720

IAN GATLEY: National Optical Astronomy Observatories, 950 North Cherry Avenue, P. O. Box 26732, Tucson, AZ 85726-6732

WILLIAM GLENCROSS and JOHN F. LIGHTFOOT: Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, England

TETSUO HASEGAWA and HIROKO SUZUKI: Nobeyama Radio Observatory, Tokyo Astronomical Observatory, Nobeyama, Minamisaku, Nagano 384-13, Japan

TETSUYA NAGATA: Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822

HARUYUKI OKUDA: Astrophysics Division, Institute of Space and Astronautical Sciences, 4-6-1 Komaba, Meguro-ku, Tokyo 153, Japan

PLATE L1

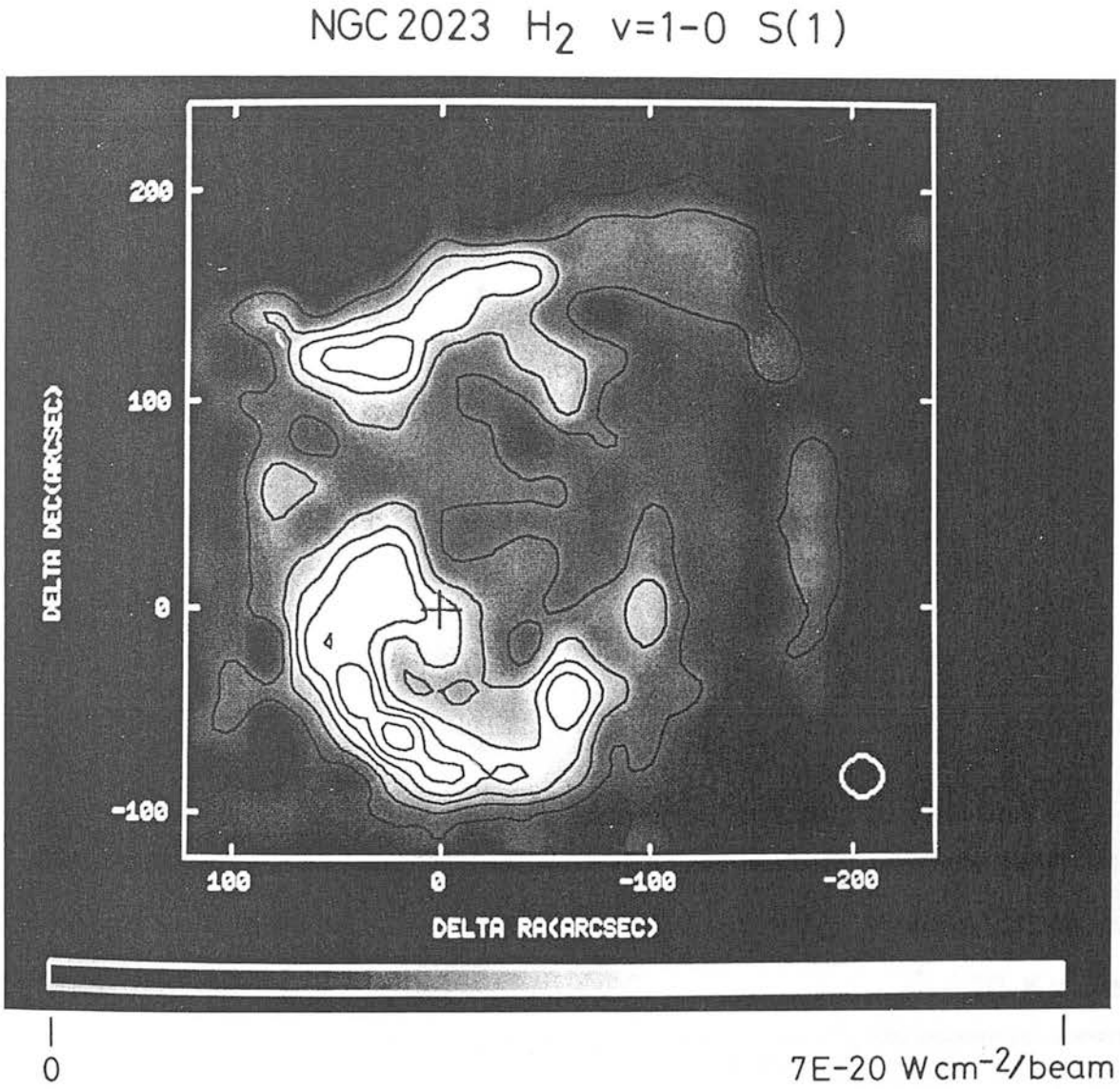


FIG. 1.—A map of the emission in the $v = 1–0$ $S(1)$ line of H_2 from NGC 2023. The contour interval and the level of the lowest contour correspond to an emission line strength of 9×10^{-21} W cm^{-2} in a 19″.6 beam. The position of the star HD 37903 is indicated by a cross.

GATLEY *et al.* (*see* **318**, L74)

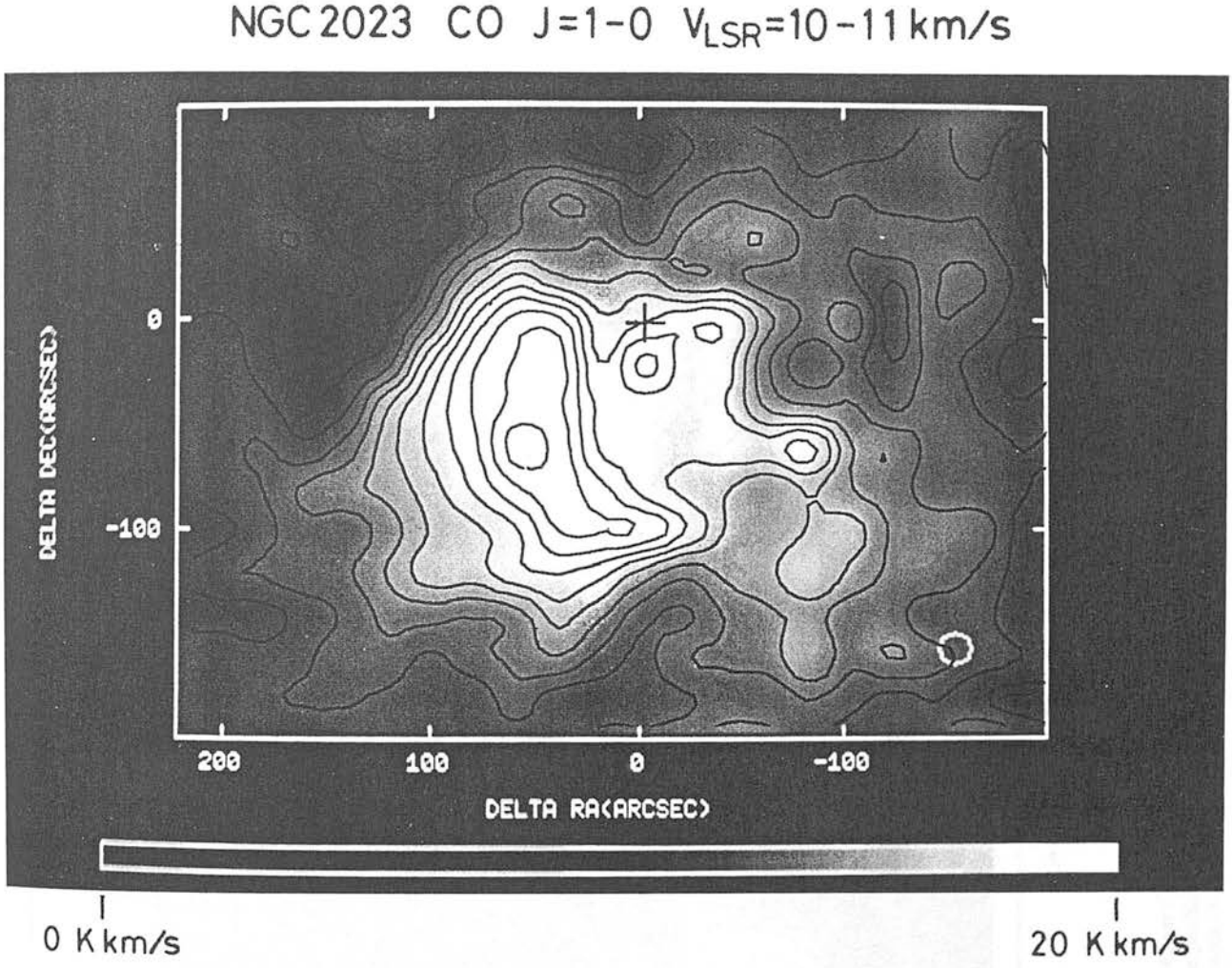


FIG. 3.—A map of the emission in ^{12}CO ($J = 1–0$) in the velocity interval $V_{LSR} = 10–11$ km s^{-1} from NGC 2023, measured with a beam size of 16″. The contours interval is 2 K km s^{-1} of $\int T_A^* dv$. The position of the star HD 37903 is indicated by a cross.

GATLEY *et al.* (*see* **318**, L75)

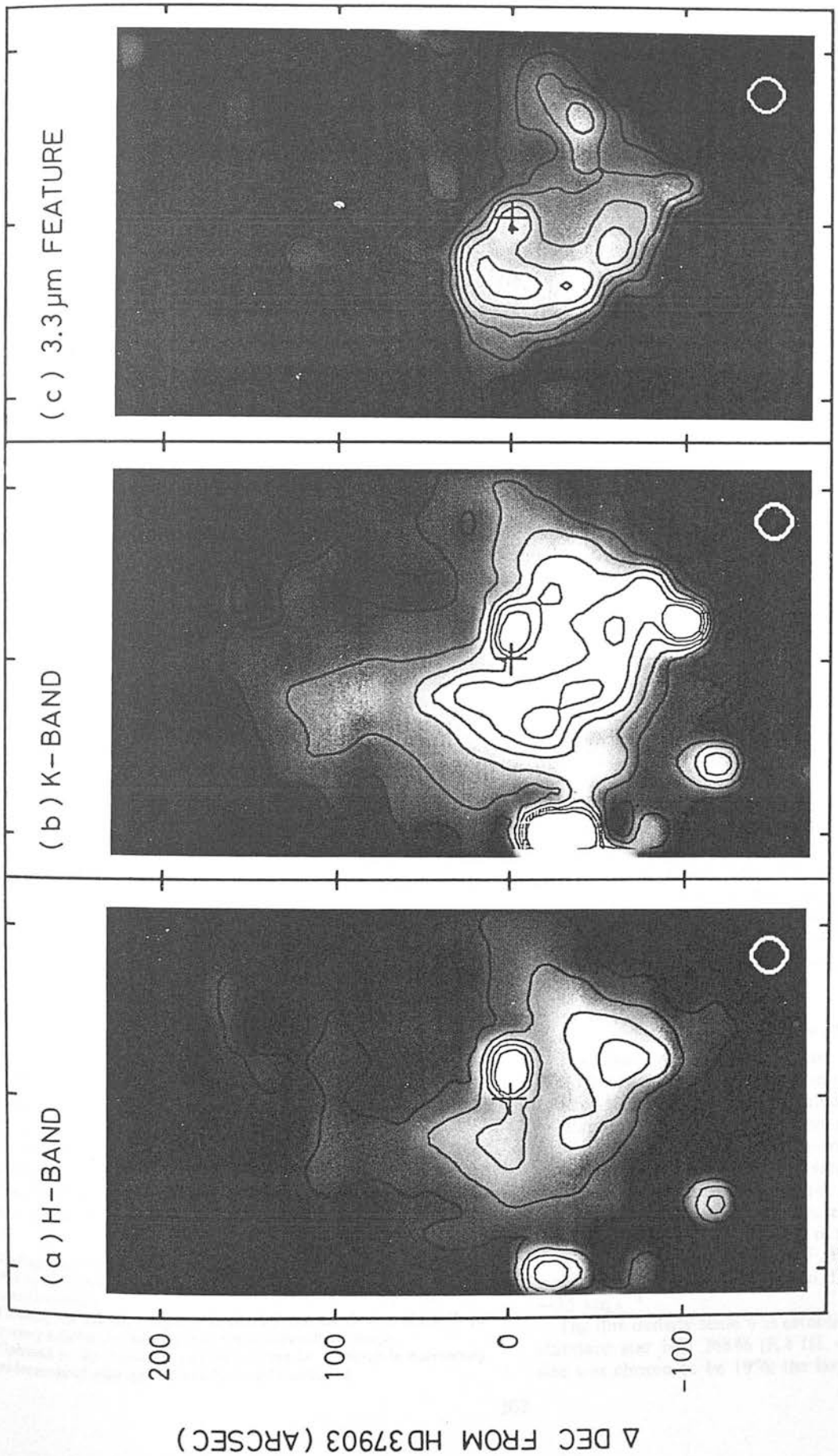


FIG. 4.—Maps of NGC 2023 at wavelengths of 1.65 μm (H), 2.2 μm (K), and 3.3 μm, measured with a beam size of 19″.6. The contour interval is 11.3 mJy per beam at H, 8.3 mJy per beam at K, and 7×10^{-20} W cm^{-2} per beam at 3.3 μm. The position of the star HD 37903 is indicated by a cross. The emission from the star itself has been subtracted from the maps at H and K and would correspond to approximately 100 contours at each of these wavelengths.

GATLEY *et al.* (*see* **318**, L75)

THE ASTROPHYSICAL JOURNAL, 336:207–211, 1989 January 1

INFRARED SPECTROSCOPY OF INTERSTELLAR MOLECULAR HYDROGEN: DECOMPOSITION OF THERMAL AND FLUORESCENT COMPONENTS

MASUO TANAKA[1] AND TETSUO HASEGAWA[1]
Nobeyama Radio Observatory, Tokyo Astronomical Observatory, University of Tokyo
SAEKO S. HAYASHI
Joint Astronomy Center[2]
PETER W. J. L. BRAND[1]
Astronomy Department, University of Edinburgh
AND
IAN GATLEY[1]
National Optical Astronomy Observatories[3]
Received 1988 February 3; accepted 1988 June 7

ABSTRACT

High-resolution spectroscopy of vibrationally excited molecular hydrogen emission in a variety of Galactic sources allows us to assess the relative contributions of thermal and fluorescent excitation. Several examples are dominated by one process or the other; collisional excitation in the Orion-KL outflow and the Galactic center, fluorescence in M17 and vdB 130. Both excitation processes contribute in NGC 7027, M1-78, S106, and NGC 2023.

When fluorescent emission is present, we can estimate the ortho/para ratio for the radiatively excited gas; it appears to lie consistently in the range 1.0–1.8 in the present examples, significantly different from 3, the statistical equilibrium value at high temperature. The near-constancy of the ortho/para ratio may imply efficient ortho-para interchange reactions in photodissociation regions.

Subject headings: interstellar: molecules — molecular processes — radiation mechanisms

I. INTRODUCTION

In two earlier papers we have demonstrated that the excitation of interstellar H_2 by ultraviolet fluorescence can result in emission-line spectra whose line ratios differ dramatically from those produced by thermal excitation in shocks (Gatley *et al.* 1987, hereafter Paper I; Hasegawa *et al.* 1987, hereafter Paper II). The agreement between theory and observation is excellent (Takayanagi, Sakimoto, and Onda 1987, hereafter TSO; Black and van Dishoeck 1987, hereafter BvD). The fluorescent H_2 emission gives information about physical processes occurring in "photodissociation regions"—the surfaces of molecular clouds exposed to intense UV radiation fields (Tielens and Hollenbach 1985).

In recent years H_2 emission has been detected from a wide variety of astronomical sources: outflows from young stellar objects, planetary nebulae, a supernova remnant, the Galactic center, H II regions, reflection nebulae, and external galaxies (see, for example, Shull and Beckwith 1982; Gatley and Kaifu 1987; Gatley 1988). The relative importance of thermal and fluorescent excitation in these various types of sources has not yet been systematically investigated.

We have shown in Paper II that the ortho/para ratio in the fluorescent source NGC 2023 is significantly smaller than the value of 3 measured in shocks. From that single result it was unclear whether the observed ortho/para ratio was a measure of physical conditions prevailing throughout the molecular cloud or was, rather, locally established in the photodissociation region by rapid chemical reactions. If the former were the case, then information on the formation and evolution of H_2 in molecular clouds could be obtained by measurements of the ortho/para ratio in various sources.

In this paper we present the results of a first survey to assess the relative importance of thermal and fluorescent excitation, and we illustrate simple graphical and analytical methods by which to distinguish their relative contributions to the measured spectrum. After separation of the thermal contribution from the observed spectra, we estimate the ortho/para ratio of the fluorescent H_2 in five sources. The ratio is found to lie within the limited range of 1.0–1.8, and an implication of this result is discussed.

II. OBSERVATIONS AND RESULTS

The observations were made on 1987 August 8 and 9 using the 3.8 m United Kingdom Infrared Telescope (UKIRT) at Mauna Kea, Hawaii. The instrumentation was that employed in Papers I and II, that is, the photometer system UKT9 equipped with a liquid nitrogen–cooled circular variable filter wheel (CVF) of ~1% spectral resolution in tandem with an ambient temperature Fabry-Perot (FP) interferometer placed directly in the f/35 beam of the telescope ahead of the photometer. The spectral resolution of the FP was 130 km s^{-1}; its orders were sorted by the CVF. Each line was observed by scanning the FP over 11 spectral points each separated by ~35 km s^{-1}.

The flux density scale was established by observations of the standard star HD 26846 (K3 III, $m_K = 2.28$ mag). The beam size was chosen to be 19″.6, the largest available. The separa-

[1] Visiting Astronomer at the United Kingdom Infrared Telescope, which is operated by the Royal Observatory Edinburgh for the Science and Engineering Research Council.
[2] Formerly the United Kingdom Infrared Telescope Facility of the Royal Observatory Edinburgh and the James Clerk Maxwell Telescope.
[3] Operated by the Association of Universities for Research in Astronomy, Inc., under contract with the National Science Foundation.

tion on the sky between the beams produced by the oscillating secondary mirror of the telescope was 180″ east-west. All observations were taken with the reference beam position located to the west of the source, except in the case of M17. Six spectral lines listed in Table 1 were selected for observations. The Doppler-broadened lines of H_2 in Orion and Sgr A are partially resolved by the FP. Since our interest is in line ratios, we will not correct for this effect (cf. Paper II).

Our targets were (1) the Doppler-shifted northeast and southwest H_2 peaks in the molecular ring surrounding the Galactic center (Gatley *et al.* 1986); (2) the bright peak in the northern bar of the H II region M17, which is offset from SAO 161357 by 210″ east and 220″ north (Gatley and Kaifu 1987); (3) the position of the CO maximum in the reflection nebula vdB 130 (van den Bergh 1966), which is offset by 100″ west and 40″ north from the illuminating B star (Hasegawa *et al.*, in preparation); (4) a position 40″ south of the exciting star in the reflection nebula NGC 1333 (cf. Sellgren 1986); (5) the H_2 peak in S106 offset 5″ east and 15″ north from IRS 4 (Hayashi *et al.*, in preparation); (6) the planetary nebula NGC 7027; (7) the planetary nebula M1-78 (= PK 93.0 − 1.1).

The measured line strengths are listed in Table 1 together with our earlier results (from Paper II) obtained at the position of peak 1 of the H_2 emission in Orion-KL (Beckwith *et al.* 1978) and at a position 80″ south of HD 37903, the exciting star of the reflection nebula NGC 2023. Also shown in Table 1 are the adopted values of the visual extinction used in making corrections for reddening before further analysis; the citations for the origins of these values are listed in footnotes to Table 1.

There are many telluric absorption lines in the 2 μm region of the spectrum. The spectra of Orion and Vega given by Scoville *et al.* (1983) show that, of the lines measured here, only $v = 1$–0 $S(2)$ is strongly affected by a telluric line. This telluric line lies 60 km s^{-1} to the blue of $v = 1$–0 $S(2)$ and has a width of 60 km s^{-1}. The size of the correction applied in our analysis depends on the continuum strength and the Doppler shift of the H_2 emission; it ranges from 25% to 55% for sources other than the Galactic center, and in the Galactic center it is 70% for the blueshifted southwest peak and is negligible for the redshifted northeast peak.

After these corrections for reddening and telluric absorption, we calculate the rotation and vibration temperatures, the column densities of H_2 in the upper energy levels of the transitions, N_u, and the ortho/para ratio in $v = 1$ and $v = 2$ (cf. Paper II).

III. DISCUSSION

It has been widely recognized that comparison of the rotational and vibrational temperatures of H_2 emission is a straightforward and intuitive diagnostic of the excitation mechanism. In thermal sources these two temperatures are similar ($\sim$2000 K), while fluorescent sources are characterized by a much higher vibrational temperature and a much lower rotational temperature (e.g., Gatley and Kaifu 1987).

Therefore, in Figure 1 we plot these two temperatures for our sample of sources, and show for comparison the typical location of thermal and fluorescent emission in this T_{vib}-T_{rot} plane. For thermal emission the relationship $T_{rot}(v = 1) = 0.8T_{vib}(v = 2\text{–}1)$ is used here to approximate an empirical fact: Orion-KL shows higher excitation temperatures at higher energy without any systematic difference between vibrational and rotational temperatures (Beckwith *et al.* 1983; Paper II). The region designated "fluorescent" is from the theoretical works of TSO and BvD; these independent calculations agree well.

Figure 1 shows very clearly that the H_2 emission from Orion-KL and the Galactic center is purely thermal, and that the H_2 emission from M17 and vdB 130 is purely fluorescent. Next we proceed to demonstrate that a unique decomposition into thermal and fluorescent contributions can be achieved for the remaining sources in Figure 1.

Some intuitive understanding of this decomposition is gained from Figure 2, which shows diagrams of relative population for the upper states of the observed transitions of H_2. Comparison of Figures 2*a* and 2*b* shows the very large differences between thermal and fluorescent emission, which were

TABLE 1

H_2 LINE INTENSITIES[a]

	FLUX/BEAM ($\times 1 \times 10^{-20}$ W cm^{-2})[b]						
	$v = 1$–0			$v = 2$–1			
SOURCE	$S(0)$ (2.2227 μm; 6471 K)	$S(1)$ (2.1213 μm; 6956 K)	$S(2)$ (2.0332 μm; 7584 K)	$S(1)$ (2.2471 μm; 12,550 K)	$S(2)$ (2.1536 μm; 13,150 K)	$S(3)$ (2.0729 μm; 13,890 K)	A_V
Orion-KL[c]	43.5 (2%)	167 (2%)	49 (20%)	16.3 (4%)	6.0 (10%)	14.0 (4%)	~40[d]
Sgr A NE[c]	0.93 (21%)	3.35 (9%)	1.03 (25%)	0.70 (31%)	...	...	~30[e]
Sgr A SW[c]	1.01 (17%)	4.00 (6%)	0.67 (30%)	0.92 (18%)	...	...	~30[e]
M17	1.38 (14%)	3.00 (9%)	1.00 (20%)	1.54 (16%)	0.81 (27%)	1.06 (24%)	...
vdB 130	0.41 (30%)	0.94 (20%)	0.46 (42%)	0.46 (23%)	...	...	...
NGC 1333	...	1.0 (24%)	...	0.3 (30%)	...	...	...
NGC 7027	4.2 (8%)	16.4 (2%)	3.1 (12%)	1.1 (25%)	1.0 (29%)	1.0 (33%)	...
M1-78	2.4 (13%)	9.0 (6%)	2.1 (16%)	1.2 (23%)	...	...	...
S106	2.1 (11%)	5.5 (5%)	1.5 (17%)	1.17 (13%)	0.71 (18%)	0.73 (33%)	...
NGC 2023	2.6 (8%)	5.5 (4%)	2.3 (5%)	1.5 (15%)	1.1 (10%)	1.2 (12%)	...

[a] No correction for extinction is made (see § II). The column heading for each line includes the wavelength (from Black and van Dishoeck (1987) and the upper state energy E_u/k.

[b] Numbers in parentheses are the uncertainties estimated from the statistical noise and the repeatability of the observations.

[c] The Doppler-broadened lines of H_2 in Orion and Sgr A are partially resolved by the FP, and the intensities are systematically underestimated (see § II).

[d] Davis, Larson, and Hofmann 1986.

[e] Geballe *et al.* 1984.

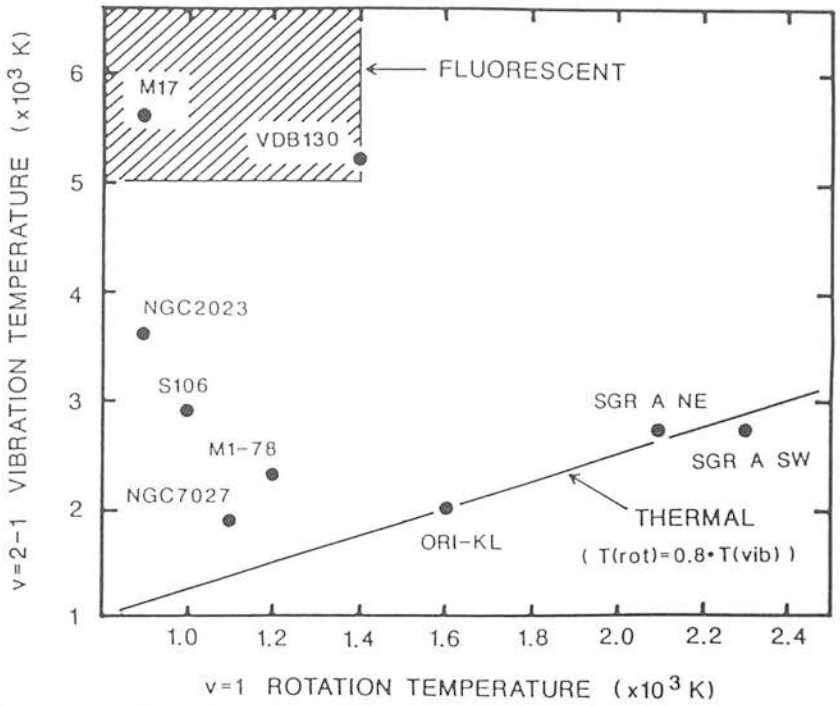


FIG. 1.—The $v = 2$–1 vibration temperature [determined from the $S(1)$ lines] is compared with the $v = 1$ rotation temperature [determined from the $S(0)$ and $S(2)$ lines]. The locus $T_{rot} = 0.8T_{vib}$ illustrates the typical location of purely thermal sources, and the hatched area that of purely fluorescent sources.

discussed in detail in Paper II. Simplistically, deviations from a straight-line fit to *all* the data for a single source in Figure 2 is a general symptom of fluorescence. Our "mixed cases" shown in Figure 2*c* immediately satisfy this modest condition, but more careful examination of these four mixed cases shows an interesting range of behavior within the subgroup.

Consider first the $v = 1$ populations alone. The planetary nebulae NGC 7027 and M1-78 share the property that the $v = 1$ states fit a straight line. Consequently, before $v = 2$ measurements were available, these sources had been interpreted as examples of pure thermal excitation from relatively cool gas (e.g., Smith, Larson, and Fink 1981). Nevertheless, planetary nebulae do contain appreciable amounts of ultraviolet radiation and are expected to exhibit fluorescence in some cases; BvD have pointed out that the $v = 1$ measurements alone are incapable of determining the presence of a small admixture of fluorescent gas.

Obviously the fit to the $v = 1$ data for the planetaries ($T_{rot} \approx$ 1100 K) fails dramatically to extrapolate through the $v = 2$ data points in Figure 2. That is, the $v = 2$–1 vibration temperature is much larger than the $v = 1$ rotation temperature. Furthermore, the $v = 2$ data for NGC 7027 cannot be fitted by a single straight line. This implies that the ortho/para ratio in $v = 2$ is noticeably different from 3 in this object.[4] The contribution of thermal excitation, which dominates $v = 1$ levels, gets smaller rapidly for higher energy levels, and the fluorescence becomes visible in $v = 2$ states. Consequently the case of Sgr A, for which we did not obtain adequate data in $v = 2$, warrants future attention.

The mixed cases S106 and NGC 2023 both show fluorescence even in $v = 1$. The component of fluorescence is larger here; BvD have calculated a mixture of thermal and fluorescent excitation in NGC 2023, and so have demonstrated a method to decompose mixtures of excitation.

The foregoing discussion suggests another useful graphical approach, which relies on the variation of the apparent ortho/para ratio in the $v = 1$ and $v = 2$ levels;[5] this variation is a diagnostic of mixed excitation. The apparent vibration temperature rises as the contribution of fluorescent gas increases. Therefore, we choose to plot in Figure 3 the "apparent ortho/para ratio" in both $v = 1$ and $v = 2$ versus the "apparent vibration temperature." If the excitation is pure (Fig. 3*a*), the apparent ortho/para ratios converge to a single value equal to the ratio of total column densities of ortho- and para-H_2. The thermal and fluorescent cases separate clearly in the ortho/para ratio versus vibration temperature plane. In mixed cases (Fig. 3*b*), on the other hand, a discord in apparent ortho/para ratios emerges. The $v = 1$ values are always larger than the $v = 2$ values; this reflects the more pronounced contribution from thermal emission in the $v = 1$ states. The thermal contribution also manifests itself as a reduction in the $v = 2$–1 vibration temperature. These trends emphasize the inability of $v = 1$ measurements alone to diagnose the presence of a minor fluorescent component. The decomposition of the mixed sources into their thermal and fluorescent components determines the true ortho/para ratio of the fluorescent component.

To quantify the heuristic arguments offered in support of

[4] It is important to remember that the normalization in Fig. 2 presupposes that the ortho/para ratio equals 3 in thermal sources, and is chosen to put ortho and para states of thermal emission onto a single sequence, as in Fig. 2*a*. The interested reader should review Paper II for more detail.

[5] The ortho/para ratio is defined as the ratio of the total column densities of ortho-H_2 (odd J) and para-H_2 (even J) as usual. We also use "apparent ortho/para ratios" in particular vibrational states, which are defined as the enhancement of a population of ortho-H_2, $N(v, J)$, relative to that interpolated or extrapolated from measured para-H_2 levels, disregarding the spin degeneracy. In Fig. 2 the apparent ortho/para ratios in $v = 1$ and $v = 2$ states are measured from the vertical displacement of the $v = 1$, $J = 3$ and $v = 2$, $J = 4$ populations from the dashed lines.

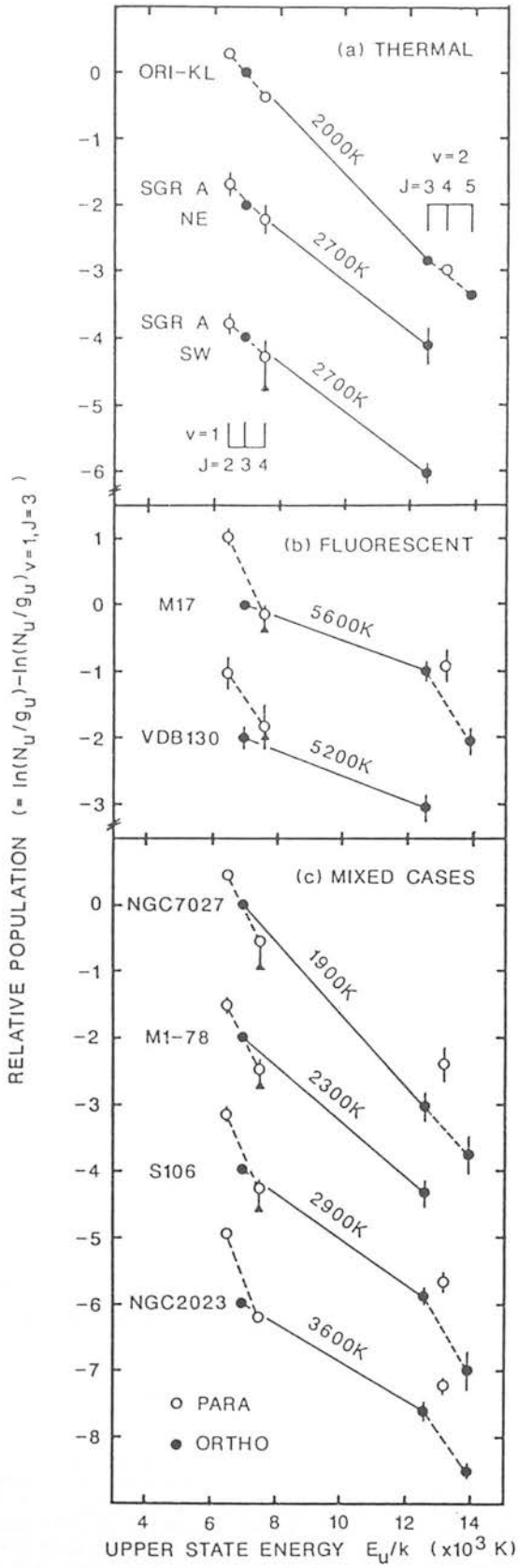


FIG. 2.—Relative populations of the upper levels of the observed transitions. The targets have been grouped according to their locations in Fig. 1. Triangles show the $v = 1$, $J = 4$ population before correcting for the telluric line absorption (see § II). The $v = 2$–1 vibrational temperatures shown correspond to the slopes of the lines connecting the $J = 3$ levels in the two vibrational states.

these graphical methods, we have made a numerical decomposition, shown in Table 2, as follows. We model the fluorescent H_2 emission on theoretical calculations by TSO and BvD; these are applicable to cases where the predissociation molecular gas temperature and density are $T < 300$ K and $n(H_2) \lesssim 10^4$ cm^{-3}, respectively. First, we assume that the fluorescent component has a vibrational temperature of 6000 K (TSO, BvD), and we calculate a grid of values for the fraction[6] of thermal emission [in the $v = 1$–0 $S(1)$ line] implied by the apparent $v = 2$–1 vibrational temperature throughout the expected range of the thermal excitation temperature (that is, 1000–3000 K). Within this grid we reject solutions having a fluorescent rotation temperature outside the theoretical range 1000 ± 200 K (TSO; BvD). As a final check we insist that the resultant ortho/para ratio of the fluorescent component in $v = 1$ and $v = 2$ be comparable. The results in Table 2 confirm and quantify the trends deduced by inspection of Figures 1–3.

After decomposition, the ortho/para ratio of fluorescent H_2 in the present sample is found to lie within the limited range of 1.0–1.8. In Paper II we suggested two possible ways to interpret this observed ratio, namely, (1) that the ratio is established in fluorescent dissociation zones or (2) that the ratio corresponds to that in the molecular cloud before radiative excitation. In the latter case we may anticipate a large range of observed ortho/para ratios (including values close to zero); variations would follow from differing histories and physical conditions within molecular clouds (age, density, temperature, etc; Dalgarno, Black, and Weisheit 1973; Flower and Watt 1984). The near-constancy of the ortho/para ratio found in our rather diverse sample of targets can be simply understood if ortho-para interchange reactions with H^+ and H_3^+ are efficient enough in the photodissociation zone to establish an ortho/para ratio close to the thermal ratio (corresponding to our case 1). Since the temperature in photodissociation regions is typically a few hundred degrees (Tielens and Hollenbach 1985), ortho/para ratios of 1–2 are expected. Observations of a larger sample of sources are needed to complete this discussion.

A limitation of the present decomposition could arise from our adoption of the models by TSO and BvD, which apply

[6] It is not necessary to assume in this decomposition that the thermal and fluorescent gas are distributed similarly within the beam. The two components may occupy different volumes of space, as indeed they do in the Orion bright bar (Hayashi *et al.* 1985).

TABLE 2

DECOMPOSITION OF THERMAL AND FLUORESCENT EMISSION

SOURCE	THERMAL FRACTION[a]	T_{ex}[b]	RADIATIVE ORTHO/PARA RATIO[c]	
	%	K	$v = 1$	$v = 2$
Orion-KL	100	2000	...	...
Sgr A NE	>90	2500	...	...
Sgr A SW	>90	2500	...	...
NGC 7027	90	1200	...	1.0 (0.3)
M1-78	80	1200	...	...
S106	65	<1400	1.4 (0.2)	1.4 (0.3)
NGC 2023	60	2000	1.1 (0.2)	1.2 (0.2)
M17	<10	...	1.8 (0.3)	1.7 (0.3)
vdB 130	<15	...	1.7 (0.4)	...

[a] Fraction of thermal component in $v = 1$–0 $S(1)$ line emission. Typical errors are 10%.

[b] The excitation temperature for the thermal component. Typical errors are 200 K.

[c] Numbers in parentheses are uncertainties which arise mainly from observational errors.

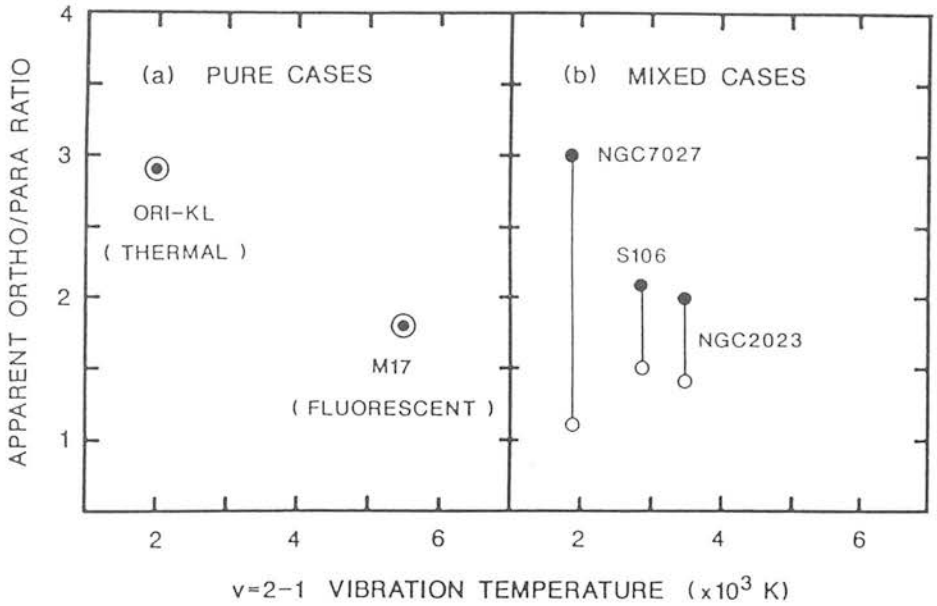


FIG. 3.—Values of the ortho/para ratio calculated for $v = 1$ (*filled circles*) and $v = 2$ (*open circles*) compared with the $v = 2$–1 vibration temperature. The discrepancy in the values for the two vibrational levels in the mixed cases is an indication of the fluorescent contribution (see text).

only to predissociation densities $\lesssim 10^4$ cm^{-3}. At higher density the collisional excitation and de-excitation in photodissociation regions thermalizes lower ($v \lesssim 2$) vibrational levels at 1000–3000 K (Hollenbach 1988). This situation could be included in the "thermal component" separated by our decomposition in some cases; the "thermal component" does not necessarily mean shocked gas. Observations of higher ($v = 3$–2) transitions would be helpful in assessing the importance of radiative excitation in "thermal" sources.

In this paper we have illustrated the methods by which analysis of H_2 spectra proceeds when excitation is by a mixture of thermal and radiative processes, and have demonstrated the experimental fact that these two processes commonly occur together within a single astrophysical system. This demonstration is a testament to the rapidly increasing sophistication of infrared instrumentation; we may confidently expect that detailed studies of the H_2 excitation structure within a single source will soon be routine.

The authors wish to thank Professor K. Takayanagi and Dr. D. L. DePoy for inspiring discussions, Professor N. Kaifu for promotion of the international collaboration which produced this paper, and the UKIRT staff for assistance during the observations. Comments by an anonymous referee improved the clarity of our discussion. Ms. Janet Grace and Ms. Linda Africano prepared the manuscript with great care and patience. M. T. and T. H. were supported by the Japanese Society for the Promotion of Science.

REFERENCES

Beckwith, S., Evans, N. J., II, Gatley, I., Gull, G., and Russell, R. W. 1983, *Ap. J.*, **264**, 152.
Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978, *Ap. J.*, **223**, 464.
Black, J. H., and van Dishoeck, E. F. 1987, *Ap. J.*, **332**, 412 (BvD).
Dalgarno, A., Black, J. H., and Weisheit, J. C. 1973, *Ap. Letters*, **14**, 77.
Davis, D. S., Larson, H. P., and Hofmann, R. 1986, *Ap. J.*, **304**, 481.
Flower, D. R. and Watt G. D. 1984, *M.N.R.A.S.*, **209**, 25; **213**, 991.
Gatley, I. 1988, *Galactic and Extragalactic Star Formation*, ed. R. Pudritz and M. Fich (Hingham, MA: Kluwer Academic), p. 159.
Gatley, I., *et al.* 1987, *Ap. J. (Letters)*, **318**, L73 (Paper I).
Gatley, I., Jones, T. J., Hyland, A. R., Wade, R., Geballe, T. R., and Krisciunas, K. 1986, *M.N.R.A.S.*, **222**, 299.
Gatley, I., and Kaifu, N. 1987, in *IAU Symposium 120, Astrochemistry*, ed. M. S. Vardya and S. T. Tarafdar (Dordrecht: Reidel), p. 153.
Geballe, T. R., Krisciunas, K., Lee, T. J., Gatley, I., Wade, R., Duncan, W. D., Garden, R., and Becklin, E. E. 1984, *Ap. J.*, **284**, 118.
Hasegawa, T., Gatley, I., Garden, R., Brand, P. W. J. L., Ohishi, M., Hayashi, M., and Kaifu, N. 1987, *Ap. J. (Letters)*, **318**, L77 (Paper II).
Hayashi, M., Hasegawa, T., Gatley, I., Garden, R., and Kaifu, N. 1985, *M.N.R.A.S.*, **215**, 31P.
Hollenbach, D. 1988, *Ap. Letters Comm.*, **26**, 191.
Scoville, N., Kleinmann, S. G., Hall, D. N. B., and Ridgway, S. T. 1983, *Ap. J.*, **275**, 201.
Sellgren, K. 1986, *Ap. J.*, **305**, 399.
Shull, J. M., and Beckwith, S. 1982, *Ann. Rev. Astr. Ap.*, **20**, 163.
Smith, H. A., Larson, H. P., and Fink, U. 1981, *Ap. J.*, **224**, 835.
Takayanagi, K., Sakimoto, K., and Onda, K. 1987, *Ap. J. (Letters)*, **318**, L81 (TSO).
Tielens, A. G. G. M., and Hollenbach, D. 1985, *Ap. J.*, **291**, 722.
van den Bergh, S. 1966, *A.J.*, **71**, 990.

P. W. J. L. BRAND: Astronomy Department, Edinburgh University, Blackford Hill, Edinburgh EH9 EHJ, Scotland, UK

IAN GATLEY: National Optical Astronomy Observatories, 950 North Cherry Avenue, P.O. Box 26732, Tucson, AZ 85726-6732

TETSUO HASEGAWA and MASUO TANAKA: Nobeyama Radio Observatory, Nobeyama, Minamisaku, Nagano 384-13, Japan

SAEKO S. HAYASHI: Joint Astronomy Center, 665 Komohana Street, Hilo, HI 96720

HIGH SPECTRAL RESOLUTION OBSERVATIONS OF FLUORESCENT MOLECULAR HYDROGEN IN MOLECULAR CLOUDS

MICHAEL G. BURTON, T. R. GEBALLE,
P. W. J. L. BRAND and A. MOORHOUSE

Reprinted for private circulation from
THE ASTROPHYSICAL JOURNAL, Vol. 352, No. 2, Part 1, 1990 April 1

THE ASTROPHYSICAL JOURNAL, 352:625–629, 1990 April 1

HIGH SPECTRAL RESOLUTION OBSERVATIONS OF FLUORESCENT MOLECULAR HYDROGEN IN MOLECULAR CLOUDS

MICHAEL G. BURTON
NASA Ames Research Center; and University of California at Irvine
T. R. GEBALLE
Joint Astronomy Centre, Hilo, Hawaii; and Foundation for Astronomical Research in The Netherlands—ASTRON
P. W. J. L. BRAND AND A. MOORHOUSE
Department of Astronomy, University of Edinburgh
Received 1989 May 19; accepted 1989 September 20

ABSTRACT

The 1–0 S(1) line of molecular hydrogen has been observed at high spectral resolution (16 km s^{-1} FWHM) in several sources where the emission was suspected of being fluorescent. In NGC 2023, the Orion Bar and Parsamyan 18 the *S*(1) line is unresolved, and the line center close to the rest velocity of the ambient molecular cloud. Such behavior is expected for UV-excited line emission. The H_2 line widths in molecular clouds thus can serve as a diagnostic for shocked and UV-excitation mechanisms; if the lines are broader than several km s^{-1} or velocity shifts are observed across a source it is likely that shocks are responsible for the excitation of the gas.

Subject headings: infrared: spectra — interstellar: molecules — molecular processes — nebulae: Orion nebula.

I. INTRODUCTION

The fluorescence of molecular hydrogen (H_2) gas in the interstellar medium is now a firm observational phenomenon (e.g., Gatley *et al.* 1987). In this process infrared line emission from H_2 follows the absorption of a UV photon by the molecule, raising it to an excited electronic state, and subsequent decay to an excited vibrational level of the ground electronic state (e.g., Black and Dalgarno 1976). The fluorescence is the radiative decay, via infrared vibration-rotation transitions, down the vibrational ladder. It has been a matter of common practice to distinguish between fluorescent and shock-excited H_2 emission by comparing intensities of lines from several vibrational levels. In low-density gas ($\lesssim 10^4$ cm^{-3}), fluorescence is characterized by strong emission lines from excited vibrational levels, with, in particular, 1–0 *S*(1)/2–1 *S*(1) line ratio of ~ 2 (e.g., Black and Dalgarno 1976; Hollenbach and Shull 1977; Black and van Dishoeck 1987). The observed spectra in the reflection nebulae NGC 2023 (Gatley *et al.* 1987) and Hubble 12 (Dinerstein *et al.* 1988) are characteristic of fluorescent emission. In contrast, shock-excited spectra, in which the H_2 levels are collisionally excited, are dominated by emission lines from the $v = 0$ and 1 levels, with, in particular a 1–0 *S*(1)/2–1 *S*(1) line ratio of ~ 10 (e.g., Kwan 1977; London, McCray, and Chu 1977; Shull and Hollenbach 1978; Draine and Roberge 1982; Chernoff, Hollenbach, and McKee 1982). Shock-excited emission has been observed in many sources, most notably the star-forming region OMC 1 (e.g., Gautier *et al.* 1976; Beckwith *et al.* 1978; Scoville *et al.* 1982; Brand *et al.* 1988).

Recent theoretical work has shown that discriminating between the two exciting mechanisms on the basis of line intensities (in particular, by using the ratio of the 1–0 *S*(1) and 2–1 *S*(1) lines) is not straightforward in all cases. In a photodissociation region (PDR) (e.g., a molecular cloud illuminated by UV radiation; see Tielens and Hollenbach 1985*a*) which is sufficiently dense ($\gtrsim 10^5$ cm^{-3}), collisional de-excitation of radiatively populated levels can move the level populations toward a thermal distribution and give the spectrum the appearance of shocked emission (Sternberg and Dalgarno 1989; Burton, Hollenbach and Tielens 1989*b*). In addition, it is possible that emission from certain kinds of shocks may appear "fluorescent." If the shock speed is sufficiently fast ($\gtrsim 40$ km s^{-1}) to dissociate all of the H_2, which then reforms on the surfaces of dust grains in cooler downstream regions, and is subsequently ejected, a "reformation" spectrum will result. The details of this process are somewhat uncertain, but Hollenbach and McKee (1989) suggest that highly excited vibrational lines may result, similar to those produced in the UV-excited radiative cascade.

Several recent observations of H_2 emission line spectra show characteristics of both emission processes. In the Orion Bar, an ionization front (IF) associated with the Trapezium stars, Hayashi *et al.* (1985) found that the ratio of the 1–0 and 2–1 *S*(1) lines varies from unity to ~ 10, the latter value occurring just behind the IF. The authors speculate that radiative excitation is occurring everywhere, but that just behind the IF a shocked layer exists, the shock wave being driven by the expansion of the H II region. However, it is hard to drive such a shock wave faster than ~ 3 km s^{-1} (Hill and Hollenbach 1978), which is too slow to significantly excite the vibrational levels of H_2. A plot of energy level against column density shows the lower level lines having apparent excitation temperatures of ~ 2000 K, typical of "shocks," while ratios of higher level lines are typical of "fluorescence" (Hippelein and Münch 1989). Tanaka *et al.* (1989) find that both radiative and thermal contributions are required to explain the excitation temperatures measured in some planetary nebulae, H II regions and reflection nebulae. In other planetary nebulae (e.g., the Dumbbell, the Ring, and NGC 6720; Zuckerman and Gatley 1988), the H_2 appears to be shock-excited on the evidence of line ratios. However, there appear to be no plausible forces available to drive a shock in order to match the intensity of the H_2 line emission (unless the collision rate coefficients have been severely underestimated and/or H-H_2 collisions

dominate over H_2-H_2); yet there is an abundance of UV photons available for pumping the molecules.

In molecular clouds there is another way to distinguish between shock-excited and fluorescent line emission, through the velocity dispersion of the excited gas. Shocks are associated with bulk motions and high velocities. Velocity widths are observed to be large in shocks, occasionally greater than 100 km s^{-1} (e.g., as in the bipolar outflows OMC 1 [Nadeau and Geballe 1979; Scoville *et al.* 1982] and DR 21 (Garden *et al.* 1986), and the supernova remnant IC 443 [Burton 1987]. Even in shocked sources where the lines are currently unresolved, such as the bipolar outflow NGC 2071, evidence for motions may be seen by comparing the emission velocities at different locations (Burton *et al.* 1989*a*). In contrast, fluorescent emission should occur at the ambient cloud velocity and the line width should be just the thermal or turbulent velocity dispersion of the cloud. This is less than 3 km s^{-1} and thus unresolvable with current observing techniques.

We have therefore conducted a program to measure the velocity profiles of H_2 lines in several fluorescent, or suspected fluorescent sources. We have observed the profiles at high spectral resolution, although the low surface brightness of the line emission necessitated the use of a large aperture, which in turn somewhat degraded the resolution. In this paper we show that the H_2 lines in most or all of the sources observed are quite narrow, a result consistent with their line emission being due to fluorescence.

II. OBSERVATIONS

The profiles of the H_2 1–0 $S(1)$ (2.1212544 μm) line in the sources listed in Table 1 were measured at the United Kingdom Infrared Telescope (UKIRT) on Mauna Kea, Hawaii, on 1988 January 20 and 21, utilizing a scanning Fabry-Perot interferometer (FP) in series with a circular variable filter (CVF) with resolving power 120, tuned to transmit at 2.12 μm. The aperture diameter was 8″ FWHM, which degraded the resolution of the FP from 12 km s^{-1} (its value in parallel light) to $\sim$16 km s^{-1}. Spatial chopping and nodding were performed, with a throw of 120″ E-W. Other details of the observations are listed in Table 1. The FP was scanned over ranges of 60–85 km s^{-1}, in steps of 5 km s^{-1}. Adjacent orders of the FP are separated by 450 km s^{-1}, thus five orders of the FP are included within the FWHM of the CVF. We are aware of four lines which may possibly contaminate the 1–0 $S(1)$ profiles. The 3–2 $S(4)$ line, two orders away at 2.1274 μm, lies -28 km s^{-1} from the line center of the 1–0 $S(1)$ line, and the 8–6 O(4) line at 2.1210 μm lies at -36 km s^{-1}. The 7–5 O(6) line, at 2.1084 μm, four orders away, and thus significantly attenuated, lies at -39 km s^{-1} (all wavelengths in standard air, from Black and van Dishoeck 1987 and Bragg, Brault, and Smith 1982). We estimate that in a shocked source the contamination by each line will be less than 1% of the 1–0 $S(1)$ line (Brand *et al.* 1988), but that in a (low-density) fluorescent source the contamination will be (not including attenuation) $\sim$6%, 7%, and 3%, respectively (Black and van Dishoeck 1987). The He I 3^3P–4^3S multiplet, with weighted mean wavelength 2.112022 μm (Litzén 1970), will occur at $+31$ km s^{-1} if it is emitted in the same source as the 1–0 $S(1)$ line and at the same local velocity. (N.B. The uncertainty in the last figure of these wavelengths results in a ± 3 km s^{-1} error in the velocity shifts above.)

TABLE 1

OBSERVING LOG

Object	(0, 0) Position	Integration Time per Point	Number of Spectral Points
Orion Bar ..... θ_2 Ori A	5^{h}32^{m}55^s.4 – 5°26′50″.7	40	13
OMC 2 ..... IRS 4 – S	5 32 59.8 – 5 11 30.1	35	18
NGC 2023 ..... HD 37903	5 39 07.3 – 2 16 58	50	15
Parsamyan 18 ..... Star A	6 57 16.6 – 7 42 16	60	17
OMC 1 ..... Peak 2	5 32 48.2 – 5 24 30	4	43

The $S(1)$ profiles are shown in Figure 1, together with the profile of an argon lamp line (19 km s^{-1} FWHM; the Ar profile is pressure broadened from the 16 km s^{-1} FWHM of the instrument) at 2.1338708 μm (in vacuo; Norlén 1973). The peak emission velocities have been shifted to 0 km s^{-1} and the fluxes normalized to unity. Details are given in Table 2. Velocities of line centers were measured relative to OMC 1 Peak 2, and are estimated to be correct to ± 3 km s^{-1}. Drifting of the FP plates was monitored by periodically measuring the frequency of maximum line emission at peak 2; the drift over the course of a night was steady and totaled 5 km s^{-1}, and thus was negligible for any single observation. Line fluxes are given relative to peak 2.

III. RESULTS

In Figure 1 it can be seen that the 1–0 $S(1)$ line at OMC 1 Peak 2 is resolved, with an observed FWHM 37 km s^{-1} (consistent with Nadeau and Geballe 1979). For the Orion Bar, NGC 2023, and Parsamyan 18, the $S(1)$ line is essentially unresolved with FWHMs ranging from 16 to 20 km s^{-1}. The $S(1)$ line in OMC 2 is somewhat broader. There are features at about the 10% level of the peak flux at ~ -30 km s^{-1} in several of the profiles, which may be due to higher excitation lines coming through other orders of the FP (see § II). For Parsamyan 18 8 W there is a feature at $\sim +35$ km s^{-1}, which may be due to the He I 3^3P–A^3S triplet; this feature is not seen in any other sources. For the the Orion Bar, OMC 2, NGC 2023 and Parsamyan 18 the $S(1)$ line emission velocity is, within the errors, at the velocity of the ambient molecular cloud determined by CO line observations (see Table 2).

IV. DISCUSSION

The observed narrow line profiles are centered at the rest velocity of the respective ambient molecular clouds, which is as expected for UV-excited line emission. However, such profiles could be produced by a single shock moving perpendicular to the line of sight. Thus these data are in themselves insufficient to demonstrate that the emission is fluorescent. Therefore, in the rest of this section we discuss additional evidence for fluorescence in the four sources that are observed.

NGC 2023.—The reflection nebula NGC 2023, excited by a B1.5 star HD 37903, is the source where the first clear identification of fluorescent H_2 line emission was made (Gatley *et al.* 1987; Black and van Dishoeck 1987). Many other observed phenomena are best interpreted as occurring on the surface of the molecular cloud as a result of UV irradiation (e.g., Pankonin and Walmsley 1976, Crawford *et al.* 1985; Witt and Schild 1988; Burton *et al.* 1989*c*) and have a close spatial coincidence with the H_2 line emission. Thus the evidence for H_2 fluorescence is particularly convincing in NGC 2023.

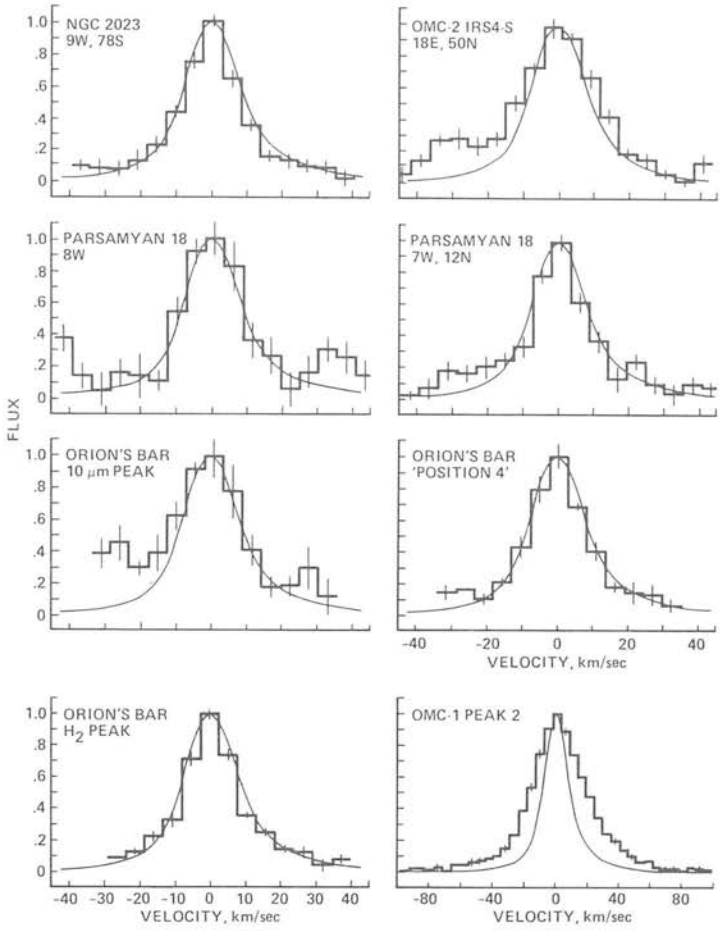


FIG. 1.—The observed profiles of the H_2 1–0 $S(1)$ line, at 2.1218 μm, in the sources NGC 2023, OMC 2, Parsamyan 18, the Orion Bar, and, for reference, in the shock-excited source OMC 1. An argon lamp line is shown for each profile. All profiles have been normalized to unit flux and have had their central velocity set to 0 km s^{-1}. Relative fluxes, FWHMs and V_{LSR} emission velocities are given in Table 2.

TABLE 2

SOURCE PARAMETERS

Source	Offset (″)	Observed FWHM of Line (km s^{-1})	Velocity of 1–0 S(1) Line (km s^{-1})	Velocity of Ambient Cloud (km s^{-1})	Relative Flux
Orion Bar	38.5 W 9.3 S (10 μm peak)[a]	20	+14.5	+11[b]	0.0095
	30 W 18 S (Position 4)[b]	17	+14		0.029
	45 W 25 S (H_2 peak)	17	+11.5		0.046
OMC 2	18 E 50 N (H_2 peak)	23	+13	+10.5[c]	0.030
NGC 2023	78 S 9 W (H_2 peak)	17	+11	+11[d]	0.031
Parsamyan 18	8 W	19	+15.5	+13.4[e]	0.011
	7 W 12 N	16	+15.5		0.015
OMC 1 Peak 2	0 E 0 N	37	+13	+9	1.0

[a] Becklin *et al.* 1976.
[b] Hayashi *et al.* 1985.
[c] Fischer *et al.* 1985.
[d] Gatley *et al.* 1987.
[e] D. Wooden and M. Cohen (private communication).
[f] An approximate line flux for OMC 1 Peak 2 is 4×10^{-19} W cm^{-2} through an 8″ aperture (Beck and Beckwith 1983). All velocities are with respect to the local standard of rest. H_2 velocities are estimated to be accurate to ± 3 km s^{-1}. They are measured relative to the line center at Peak 2, for which the values given are from Nadeau and Geballe 1979.

Parsamyan 18.—The reflection nebula P 18 has a 1–0/2–1 S(1) line ratio of 1.7, as expected for fluorescent line emission (Sellgren 1986). The UV-excited infrared emission bands at 3.3 μm (Gatley *et al.* 1987), 6.2, 7.7 μm, and 11.3 μm (Cohen *et al.* 1986), have been observed. The H_2 1–0 $S(1)$ line and the 3.3 μm emission feature are morphologically similar (Burton *et al.* 1989*c*). Thus, the observed narrow H_2 line profile is entirely consistent with the previous understanding of the excitation of the H_2.

The Orion Bar.—From the variation of the 1–0/2–1 $S(1)$ line ratio across the ionization front in the Orion Bar, Hayashi *et al.* (1985) concluded that both UV-excitation from the Trapezium stars, and shock-excitation driven by the expansion of the H II region were responsible for the H_2 line emission. As pointed out earlier, however, it is difficult to drive such a shock wave sufficiently fast to excite the vibrational H_2 lines. In addition, the H_2 line emission peaks 15″ away from the ionization front, whereas it would be expected to occur immediately adjacent to the front if the H_2 were shock-excited. A detailed model of the region by Tielens and Hollenbach (1985*b*) fits the observed [O I] 63 μm and 146 μm, [C I] 609 μm, [C II] 158 μm and low-lying CO rotational lines by a PDR model with density 2×10^5 cm^{-3} and a UV-field appropriate to that measured for the Trapezium stars. The high-excitation H_2 lines observed by Hippelein and Münch (1989) require a density of $\sim 10^6$ cm^{-3} in the same UV-field for this model (Burton *et al.* 1989*b*), but a possible underestimation of H_2 collisional de-excitation rates may lower this density to a value closer to that determined from the fine-structure lines. The 3.3 μm emission feature also originates in a neutral region behind the ionization front (Sellgren 1981; Aitken *et al.* 1979). However, images of the 3.3 μm feature and the $S(1)$ line are not identical, with the $S(1)$ line peaking a further 10″ away from the ionization front than the 3.3 μm feature (Burton *et al.* 1989*c*). The strongest $S(1)$ emission occurs in the region assigned as shock-excited by Hayashi *et al.* (1985). Emission from the 3.3 μm feature does arise here too, although its intensity is reduced relative to that closer to the ionization front.

On the basis of these data and the current H_2 spectra, we cannot rule out the possibility that some of the H_2 line emission from the Orion Bar is shock-excited, although it seems clear from the narrow line widths that the bulk of the emission is fluorescent. It appears likely that much of the H_2 line emitting gas is sufficiently dense ($> 10^5$–10^6 cm^{-3}) that collisions thermalize the vibrational levels of the H_2 molecules before they can radiate. The emission spectrum therefore appears similar to that expected from hot, shocked gas, and the lines are narrow and emitted at the rest velocity of the cloud, as observed.

OMC 2.—H_2 line emission in OMC 2 is centered on IRS4, (Fischer, Righini-Cohen, and Simon 1980; Thronson and Thompson 1982) and extends $\sim 90''$ NNE as a narrow "jet" (Burton, Garden, and Russell 1990). The present observations were centered on a bright clump situated about half way along the "jet." No infrared sources have been observed in the region of this "jet." On the basis of a 1–0/2–1 $S(1)$ line ratio of 7 for the emission on IRS 4, Thronson and Thompson (1982) argue for shock-excitation of the gas. The noise level on their 2–1 $S(1)$ line is high, however, so the interpretation is subject to some doubt. ^{12}CO line emission is seen with wings extending to ± 6 km s^{-1} away from line center around IRS 4 (Fischer *et al.* 1985). The emission is slightly bipolar, with axis extending NNE from IRS 4 and centered at about (10 E, 20 N) from the source; i.e., in a direction similar to the "jet" seen in the $S(1)$ line. IRS 4-N is the most luminous member of the star cluster making up OMC 2 and appears to be surrounded by a disk oriented E-W (Pendleton *et al.* 1986; Rayner *et al.* 1989). This disk may collimate an outflow from IRS 4-N. Thus, on this evidence, it seems plausible to assume that the $S(1)$ line emission from OMC 2 is shocked. However, the line profile observed is narrow (although somewhat broader than one would expect from a quiescent cloud), and the emission velocity is close to the rest velocity of the cloud. These are indicative of fluorescence, but could also be consistent with a low-velocity shock, or a faster one traveling in the plane of the sky. The observed CO velocities are also too small to significantly populate the vibrational H_2 levels if there are shocks moving with the same velocities. However, examples are known (e.g., DR 21; Garden *et al.* 1986) where observed H_2 velocities far exceed those observed in CO. Further observations of the H_2 line emission, in particular spectra and profiles taken at several positions along the "jet," are required in order to determine the source of excitation.

V. CONCLUSIONS

We have measured the profile of the H_2 1–0 $S(1)$ line at high resolution in several sources where the emission is suspected of being fluorescent. In each case we find that the line is narrow and the emission velocity is close to that of the ambient cloud. This is as expected in sources excited by UV-radiation, but is not commonly observed in shock-excited sources. Thus in some cases high-resolution spectroscopy can provide an additional diagnostic for distinguishing between shocks and fluorescence. Taken together with other evidence, it is clear that the H_2 line emission in NGC 2023, Parsamyan 18, and the Orion Bar result from fluorescence. In OMC 2 we cannot distinguish yet between shocks and fluorescence.

We wish to thank the staff of the United Kingdom Infrared Telescope for friendly and able assistance. A. M. is supported by an SERC studentship. This work was done, in part, while M. G. B. held a National Research Council NASA Research Associateship at Ames Research Center.

REFERENCES

Aitken, D. K., Roche, P. F., Spenser, P. M., and Jones, B. 1979, *Astr. Ap.*, **76**, 90.
Beck, S. C., and Beckwith, S. 1983 *Ap. J.*, **271**, 175.
Becklin, E. E., Beckwith, S., Gatley, I., Mathews, K., Neugebauer, G., Sarazin, C., and Werner, M. W. 1976, *Ap. J.*, **207**, 770.
Beckwith, S., Persson, S. E., Neugebauer, G., and Becklin, E. E. 1978 *Ap. J.*, **223**, 464.
Black, J. H., and Dalgarno, A. 1976 *Ap. J.*, **203**, 132.
Black, J. H., and van Dishoeck, E. F. 1987 *Ap. J.*, **322**, 412.
Bragg, S. L., Brault, J. W., and Smith, W. H. 1982, *Ap. J.*, **263**, 999.
Brand, P. W. J. L., Moorhouse, A., Burton, M. G., Geballe, T. R., Bird, M., and Wade, R. 1988 *Ap. J. (Letters)*, **334**, L103.
Burton, M. G. 1987, *Quart. J.R.A.S.*, **28**, 269.
Burton, M. G., Garden, R. P., and Russell, A. P. G. R. 1990 in preparation.
Burton, M. G., Geballe, T. R., and Brand, P. W. J. L. 1989*a*, *M.N.R.A.S.*, **238**, 1513.
Burton, M. G., Hollenbach, D. J., and Tielens, A. G. G. M. 1989*b*, in *Infrared Spectroscopy in Astronomy*, 22 *ESLAB Symposium* (ESA SP-290), (Salamanca, Spain, 1988 December 7–9) ed. B. Kaldeich.

Burton, M. G., Moorhouse, A., Brand, P. W. J. L., Roche, P. F., and Geballe, T. R. 1989*c*, in *IAU Symposium 135, Interstellar Dust*, ed. L. J. Allamandolla and A. G. G. M. Tielens, in press (NASA CP 3036).
Chernoff, D. F., Hollenbach, D. J., and McKee, C. F. 1982, *Ap. J. (Letters)*, **259**, L97.
Cohen, M., Allamandola, L., Tielens, A. G. G. M., Bregman, J., Simpson, J. P., Witteborn, F. C., Wooden, D., and Rank, D. 1986, *Ap. J.*, **302**, 737.
Crawford, M. K., Genzel, R., Townes, C. H., and Watson, D. M. 1985 *Ap. J.*, **291**, 755.
Dinerstein, H. L., Lester, D. F., Carr, J. S., and Harvey, P. M. 1988, *Ap. J. (Letters)*, **327**, L27.
Draine, B. T., and Roberge, W. G. 1982, *Ap. J. (Letters)*, **259**, L91.
Fischer, J., Righini-Cohen, G., and Simon, M. 1980, *Ap. J. (Letters)*, **238**, L155.
Fischer, J., Sanders, D. B., Simon, M., and Solomon, P. M. 1985, *Ap. J.*, **293**, 508.
Garden, R., Geballe, T. R., Gatley, I., and Nadeau, D. 1986. *M.N.R.A.S.*, **220**, 203.
Gatley, I., Hasegawa, T., Suzuki, H., Garden, R., Brand, P., Lightfoot, J., Glencross, W., Okuda, H., and Nagata, N. 1987, *Ap. J. (Letters)*, **318**, L73.
Gautier, T. N., III, Fink, U., Treffers, R. R., and Larson, H. P. 1976, *Ap. J. (Letters)*, **207**, L29.
Hayashi, M., Hasegawa, T., Gatley, I., Garden, R. Gautier, and Kaifu, N. 1985, *M.N.R.A.S.*, **215**, 31P.
Hill, J. K., and Hollenbach, D. J. 1978, *Ap. J.*, **225**, 390.
Hippelein, H. H., and Münch, G. 1989, *Astr. Ap.*, **213**, 323.
Hollenbach, D. J., and McKee, C. F. 1989, *Ap. J.*, **342**, 306.
Hollenbach, D. J., and Shull, J. M. 1977, *Ap. J.*, **216**, 419.
Kwan, J. 1977, *Ap. J.*, **216**, 713.
Litzén, U. 1970, *Phys.-Scripta*, **2**, 103.
London, R., McCray, R., and Chu, S.-I. 1977, *Ap. J.*, **217**, 442.
Nadeau, D., and Geballe, T. R. 1979, *Ap. J. (Letters)*, **230**, L169.
Norlén, G. 1973, *Phys. Scripta*, **8**, 249.
Pankonin, V., and Walmsley, C. M. 1976, *Astr. Ap.*, **48**, 341.
Pendleton, Y., Werner, M. W., Capps, R., and Lester, D. 1986, *Ap. J.*, **311**, 360.
Rayner, J., McCaughrean, M., Aspin, C. and McLean, I. 1989, *M.N.R.A.S.*, **241**, 469.
Scoville, N. Z., Hall, D. N. B., Kleinmann, S. G., and Ridgway, S. T. 1982, *Ap. J.*, **253**, 136.
Sellgren, K. 1981, *Ap. J.*, **245**, 138.
———. 1986, *Ap. J.*, **305**, 399.
Shull, J. M., and Hollenbach, D. J. 1978, *Ap. J.*, **220**, 525.
Sternberg, A., and Dalgarno, A. 1989. *Ap. J.*, **338**, 197.
Tanaka, M., Hasegawa, T., Hayashi, S. S., Brand, P. W. J. L., and Gatley, I., 1989. *Ap. J.*, **326**, 207.
Tielens, A. G. G. M., and Hollenbach, D. J. 1985*a*, *Ap. J.*, **291**, 722.
———. 1985*b*, *Ap. J.*, **291**, 747.
Thronson, H. A., and Thompson, R. I. 1982. *Ap. J.*, **254**, 543.
Witt, A. N., and Schild, R. E. 1988, *Ap. J.*, **325**, 837.
Zuckerman, B., and Gatley, I. 1988, *Ap. J.*, **324**, 501.

P. W. J. L. Brand and A. Moorhouse: University of Edinburgh, Department of Astronomy, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland, UK

Michael G. Burton: NASA Ames Research Center, MS 245-6, Moffett Field, CA 94035

T. R. Geballe: Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720

Mon. Not. R. astr. Soc. (1981) 195, 353–360

Emission line ratios from H II regions in two nearby galaxies: NGC 7424 and M83

P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Edinburgh, Scotland*

I. M. Coulson *South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape, South Africa*

W. J. Zealey *Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

Received 1980 July 4; in original form 1979 October 4

Summary. Twenty-six H II regions in the two Sc galaxies M83 and NGC 7424 were observed, using the Wampler–Robinson image dissector scanner with a resolution of approximately 10 Å. A reduction program was applied to the data to derive intensity-calibrated parameters for the H II regions. In particular, the [O III]/Hβ excitation parameter, and [N II]/Hα ratio were correlated with distance from the centres of the galaxies. The results are consistent with what is known about abundance gradients and star formation rates in galaxies. In particular, the data indicate that more metals are present in the galaxy (M83) with the greater surface brightness.

1 Introduction

As part of a study of the structure of spiral arms, and the sizes of H II regions, spectra of H II regions were observed in two nearby galaxies M83 (NGC 5236) and NGC 7424. Plate 1(a) and (b) show enlargements from UK 1.2-m Schmidt plates of the two galaxies with the observed positions marked, and Table 1 lists their integral properties.

It has been established since the important work by Searle (1971) that gradients in line ratio versus radius in galaxies are due to gradients of chemical abundance. In this connection the major work has been described in Smith (1975), Webster (1977), Jensen, Strom & Strom (1976) and in Shields & Searle (1978) who modelled in detail the composition of both star and nebula to compare with observations of M101. The significance of these gradients (or lack of gradient) of chemical composition in galaxies is discussed by Pagel *et al.* (1978, 1979) and Alloin *et al.* (1979b). Results have been obtained for several H II regions in M83 by Dufour *et al.* (1980, hereafter DTJS), who found an O/H gradient.

This systematic effect shows clearly in our results for the galaxies M83 and NGC 7424. The main purpose of the observations was to compare the H II region excitation with size measurements (Coulson 1980), which will be published in due course.

Table 1. Properties of the two galaxies.

		M83	NGC 7424
RA (1950)	(d)	$13^h\ 34^m\ 12^s.0$	$22^h\ 54^m\ 28^s.1$
Dec (1950)	(d)	$-29°\ 36'\ 40''$	$-41°\ 20'\ 21''$
Type	(a)	SABc	SABc
Inclination	(a)	24°	27°
Distance (Mpc)	(c)	3.7	9.1
Isophotal diameter	(a)	329 arcsec ≡ 5.9 kpc	222 arcsec ≡ 8.0 kpc
	(b)	486 arcsec ≡ 8.7 kpc	–
Magnitude (B)	(ab)	8.2	10.98

Data from: (a) de Vaucouleurs *et al.* (1976); (b) DTJS; (c) de Vaucouleurs (1979b) or (a); (d) UK Schmidt plates.

2 Observations

The observations were made in 1976 May with the Wampler–Robinson image dissector scanner (IDS) on the Boller and Chivens Cassegrain spectrograph with a 600 lines mm^{-1} grating at the Anglo-Australian 3.9-m telescope, giving a spectral resolution of approximately 10 Å.

The decker and slit combination presented to the sky a pair of apertures 1.8 by 4.5 arcsec with their long axes parallel to the line joining their centres, which were separated by 20 arcsec. The apertures were oriented east–west. The dissector tube had poor blue response and was set with the blue limit of the recorded spectrum at approximately 3900 Å (unfortunately, therefore, having to miss the [O II] lines at 3726/9 Å) and the red limit at approximately 7500 Å.

An interactive program suite was created to perform the data reduction, which proceeds along the same general path as the Lick reduction program at the AAO.

The differential refraction throughout the wavelength range amounts to approximately 1 arcsec at the relatively small zenith distances at which we observed, and so is not a major contributor to error, although already at the margin of significance. More crucial was the fact that the seeing disc was comparable in size to the slit width, so variable size of the seeing disc produces large intensity variation. Thus absolute fluxes are not reliable.

In cases where blending of Hα and [N II] lines was significant, the profiles were carefully analysed in several different ways, and the error bars in Fig. 2 span all values so derived. The [S II] doublet λ 6716/31 was not clearly separated, but a crude density estimate was made by comparing its average wavelength with that of Hα.

3 Results

The results are given in Table 2 and Figs 1 to 3. Table 2 lists: (1) our identification letter for the object which is also marked in Plate 1. Note that spectra were obtained of four very closely spaced objects listed as N1 to N4. These will be discussed below. The identification Z is used for the nuclear region, and in NGC 5236 we observed two such regions. (2) The rectangular coordinates on the plane of the sky, in arcsec relative to the nucleus, the x-axis westward and the y-axis north. (3) The corresponding radial distance, corrected for tilt in the case of M83. (4) The optical depth at Hβ assuming an intrinsic Hα/Hβ emissivity of 2.86 appropriate for low density gas at 10 000 K, and also assuming all absorption in front of the source. This is further discussed below. (5) A code representing the density inferred from the [S II] line ratio, L, M, H, X indicating estimates of density respectively lower, approximately

Table 2. Values deduced from H II regions in two galaxies M83 and HGC 7424.

Object	Position (arcsec)		Radius (arcsec)	Optical depth at Hβ §	[S II] * density code	[O III]/Hβ ‡	Hα/Hβ	[N II]/Hα ‡	[S II]/Hα	Mean Hβ spectral index	Ratio of equivalent widths Hα/Hβ
(1)	(2)		(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
M83/L	−31	56	70	1.7 (0.03)	L	0.03	0.06	0.29	0.11	−3.9	8.5
K	−44	59	80	1.8 (0.02)	M	0.02	0.04	0.26	0.10	−3.9	9.8
I	−93	−15	97	2.8 (0.02)	L	0.19	0.10	0.33	0.13	−3.7	9.4
E	−83	−55	100	2.1 (0.12)	X	–	(0.30)†	0.44	0.19	−4.6	11.5
G	−90	−49	103	2.0 (0.06)	L	0.36	0.19	0.38	0.13	−3.6	12.5
H	−98	−35	106	1.9 (0.01)	L	0.04	0.12	0.34	0.12	−2.8	8.3
F	−90	−59	108	1.6 (0.01)	M	0.24	0.09	0.47	0.11	−3.7	9.2
N_1				2.2 (0.03)	L	–	0.03	–	0.14	−4.8	11.4
N_2	111	11	116	1.8 (0.02)	L	–	0.09	0.29	0.11	−3.6	7.3
N_3				1.4 (0.02)	M	0.19	–		0.16	−2.5	5.3
N_4				1.8 (0.04)	M	0.03	0.09		0.16	−4.7	5.3
J	−108	32	121	3.4 (0.16)	M	–	–	0.36	0.18	−5.4	15.7
D	−60	−109	125	1.4 (0.05)	M	0.44	–	0.42	0.09	−4.4	8.9
C	−29	−125	132	1.6 (0.01)	L	0.23	0.10	0.48	0.12	−5.0	9.6
M	111	−52	133	2.6 (0.02)	M	0.32	0.13	0.31	0.10	−4.2	11.5
B	8	−139	147	2.1 (0.07)	M	–	–	0.40	0.12	−2.5	6.4
A	29	−147	160	1.6 (0.03)	M	0.08	0.07	0.40	0.10	−5.7	10.0
P	234	−18	248	2.2 (0.01)	L	0.60	0.13	0.28	0.10	−3.2	8.9
Z_1		nucleus		9.0 (1.0)	M	(4.10)†	–	0.71	0.21	−9.4	(50)†
Z_2				4.8 (0.24)	M	0.33	–	0.57	0.14	−6.7	21.7
N7424/A	54	78	95	1.4 (0.04)	L	1.3	–	0.18	0.25	−3.9	7.7
E	100	−63	118	(4.7)† (0.5)	H	3.5	(0.62)†	0.10	(0.47)†	−8.7	(20)†
D	−50	153	161	0.9 (0.03)	L	3.5	0.21	0.16	0.15	−7.1	11.2
C	−177	−71	191	1.7 (0.04)	M	2.9	0.05	0.12	0.14	−4.9	9.1
B	−12	209	209	1.1 (0.02)	M	5.5	0.23	0.08	0.10	−4.0	13.7
Z		nucleus		–	X	–	–	(0.62)†	(0.38)†	–	–

* L: density less than 700 cm^{-3}.
† Values in brackets are uncertain.
‡ O III 5007 + 4959; N II 6583 only.
§ The second figure (in brackets) is the rms deviation of the line strength (averaged throughout the spectrum) divided by the Hβ line strength. To estimate the photometric accuracy of a particular line ratio, multiply this quantity by (1 + line ratio squared)$^{1/2}$.

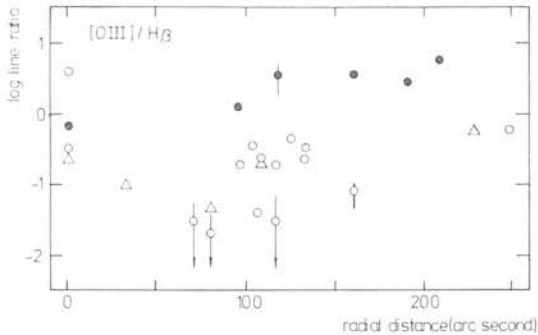


Figure 1. Relative line strength [O III] λ 5007/Hβ versus radial distance in arcsec in each galaxy. Solid circles represent data for NGC 7424, open circles for M83, and triangles for the data for M83 from DTJS.

equal to, or higher than 700 cm^{-3}; or indeterminate. (6) The line strength ratio [O III] λ (4959 + 5007)/Hβ. (7) The ratio He λ 5875/Hβ, formally 'de-reddened' by a factor $\exp(-0.17\tau_\beta)$. (8) The ratio [N II] λ 6583/Hα. (9) The ratio [S II] λ (6717 + 31)/Hα. (10) The mean stellar continuum spectral index $d \ln F_\lambda / d \ln \lambda$ at Hβ, corrected for reddening corresponding to the optical depth in column 4, and 1/λ absorption. (11) The ratio of Hα/Hβ equivalent widths. Figures in brackets are of doubtful accuracy.

Fig. 1 shows the ratio of [O III] λ 5007/Hβ line strengths as a function of radial distance. Since the two galaxies are nearly the same projected size on the sky, the results from both galaxies are plotted. The error bars are marked, unless they are smaller than the plotting symbol. Fig. 2 shows the ratio [N II] λ 6583/Hα as a function of radial distance. Because correlation between dust content and line ratios might be expected, the dust content was estimated using the ratio of Hα to Hβ line strengths. This ratio was assumed for the reduction to be 2.86 corresponding to a low density H II region at 10^4 K.

Errors will be caused by the assumption that all the dust extinction occurs outside the H II region. The optical depth at Hβ would be given, if this were the case, by

$$\tau_\beta = (1 - a_{\alpha\beta})^{-1} \ln (I_{\beta 0} I_\alpha / I_{\alpha 0} I_\beta) \qquad (1)$$

where I_α and $I_{\alpha 0}$ are line strengths at Hα, observed and intrinsic, respectively, and $a_{\alpha\beta}$ is the ratio of dust opacity (or optical depth) at Hα to that at Hβ. If the dust were uniformly

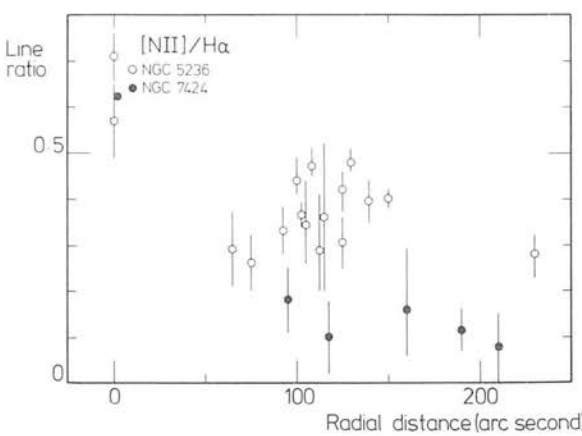
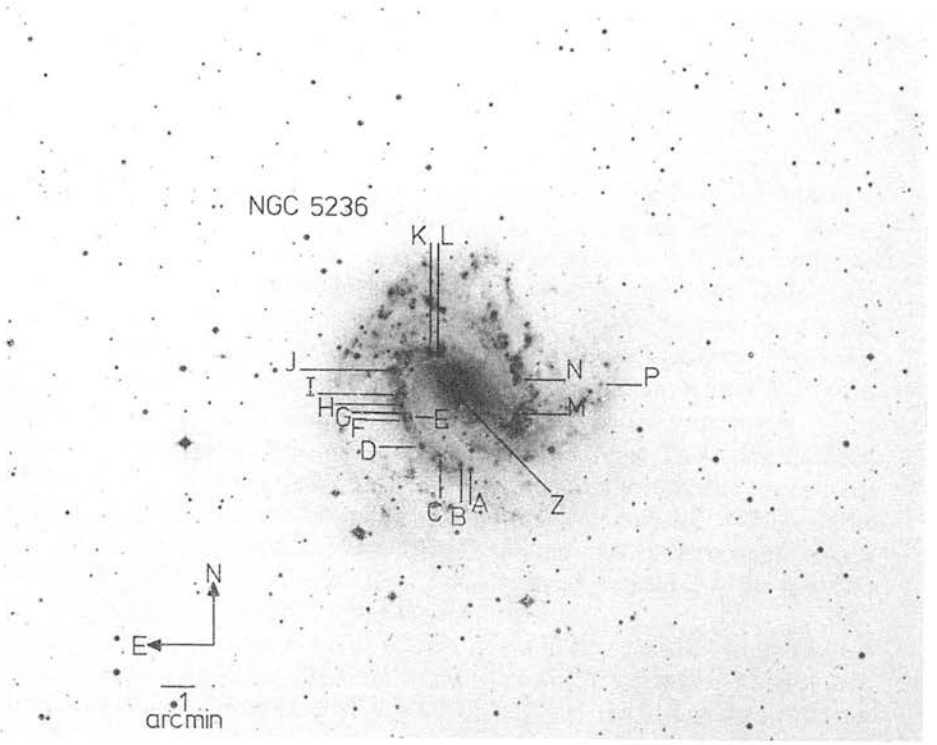
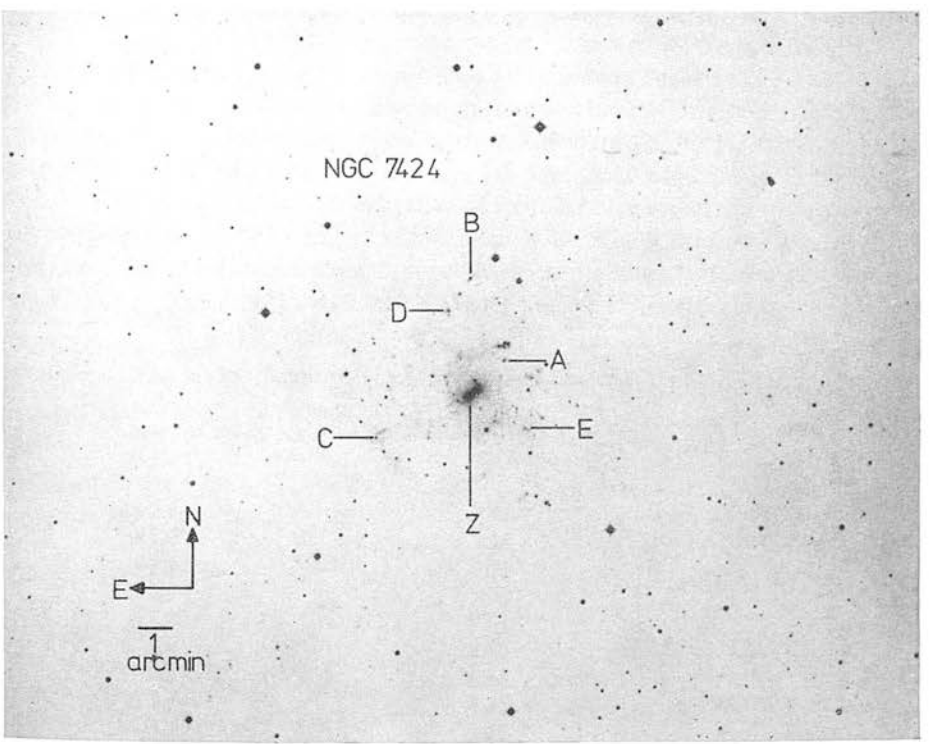


Figure 2. Relative line strength [N II] λ 6583/Hα versus radial distance. Symbols as in Fig. 1.



(a)



(b)

Plate 1. (a) NGC 5236 (M83) on a UK Schmidt Telescope Unit Hα plate with 120-min exposure. (b)NGC 7424 on a similar plate with 180-min exposure.

[*facing page* 356]

mixed in the emitting region and scattering is ignored, the optical depth $\bar{\tau}_\beta$ deduced from the observations of line strengths would be related to this value by

$$\tau_\beta = (1 - a_{\alpha\beta})^{-1} \ln \left\{ a_{\alpha\beta}{}^{-1} \frac{1 - \exp(-a_{\alpha\beta}\bar{\tau}_\beta)}{1 - \exp(-\bar{\tau}_\beta)} \right\}. \quad (2)$$

The largest value we would expect for τ_β would be (using interpolations from the reddening law given in Allen (1973) which gives $a_{\alpha\beta} = 0.65$) $\tau_\beta = 1.23$. From the results of Mathis (1970), with a dust albedo of 0.6 we find the largest value of τ_β to be 0.5. Thus scattering increases the constraint on internal reddening. Since values in column 4 of Table 1 are almost all in excess of these low values of τ_β, a large part of the reddening must be external to the source of the emission lines. The galactic optical depth in the directions of M83 and NGC 7424 at Hβ is 0.3 and 0.2 respectively (de Vaucouleurs, de Vaucouleurs & Corwin 1976). The mean effective optical depths are 2.0 ± 0.5 (1σ) and 1.3 ± 0.3 respectively.

There are three determinants of spectral type. Column 10 of Table 2 contains the mean spectral index ($d \ln F_\lambda / d \ln \lambda$) of the stellar continuum, approximately corrected for external reddening by subtracting τ_β from the index. By fitting blackbody curves at 4800 Å to the atmospheres of Kurucz, Peytremann & Avrett (1974) spectral indices for stars of various effective temperatures were calculated and are in Table 3. It can be seen that the observed spectral indices all indicate OB stars dominating the Hβ flux.

The second determinant, in column 11, is the ratio of Hα to Hβ equivalent widths, which can be translated to reddening-free stellar flux. This is listed in Table 3, where it can be seen that W_α/W_β increases from a value of 1.38 at 4000 K to 9.1 at the highest temperatures, when the lines are sitting on the Rayleigh–Jeans tail.

Fig. 3 shows the third determinant of stellar spectrum. This is the Hβ equivalent width which can be used to determine the ratio of UV photons emitted to stellar flux, at Hβ. From the models of Kurucz *et al.* (1974) one can determine the flux at 4800 Å as a fraction of the total flux. For a slightly different but comparable set of models Panagia (1974) determined the number of ionizing photons per unit luminosity for stars of different effective temperature. Assuming that all the ionizing photons are absorbed by gas, the Hβ flux can be calculated (Osterbrock 1974), and hence we find the Hβ equivalent width as in Table 3. The ratios in column 11 of Table 2 are all indicative of extreme blue stellar continua, and show no particular correlation with radius, and neither is there any marked difference between the galaxies. The third determinant is most likely to be abundance sensitive (as pointed out by Shields & Searle 1978) via the line blocking of the UV stellar continuum by

Table 3. Parameters from stellar models (Kurucz *et al.* 1974; Carbon & Gingerich 1969).

$T_{\rm eff}$ (K)	Spectral index at Hβ	W_α/W_β ($2.86 \times F_\beta/F_\alpha$)	W_β (Å)
4000	+2.8	1.38	–
6000	+1.0	2.33	–
8000	−1.5	5.94	–
10 000	−2.9	7.28	–
15 000	−3.3	8.06	–
20 000	−3.4	8.67	0.4
25 000	−3.6	9.12	5.7
30 000	−3.7	8.96	75.0
35 000	−3.7	8.58	332.0
40 000	−3.7	9.1	695.0
50 000	−3.7	9.1	2086.0

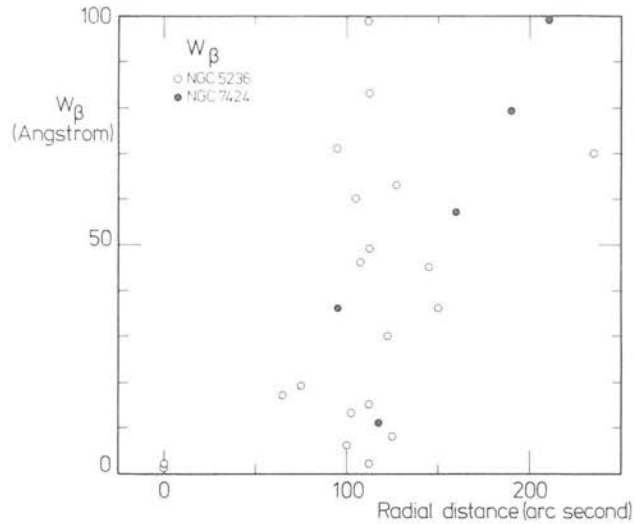


Figure 3. Equivalent width in Angstrom units of the Hβ line, versus radial distance in each galaxy.

heavy elements. However, there is no marked difference in the mean value between the two galaxies, the value for M83 is 40 ± 29, that for NGC 7424 is 57 ± 36. In the data for NGC 7424 there is an evident trend, with the largest values for equivalent width at the greatest distances, which would be consistent with other trends, but is not consistent with the small difference in mean values between the galaxies.

The four closely spaced regions N1–N4 observed in M83 differ markedly in their Hα equivalent widths (i.e. presumably the Zanstra temperature of their stellar background) and also slightly in their [O III]/Hβ line ratio, but this is almost certainly due to noise as the [O III] lines are weak.

4 Discussion

Since we were unable to observe [O II] lines, we cannot complete the analysis of O/H gradients directly, but we will use the results of Pagel *et al.* (1979) and Alloin *et al.* (1979b) as discussed in DTJS. Fig. 4 of DTJS is a plot of observations of [O II]/Hβ versus [O III]/Hβ. We use the symbols II and III to denote the two ratios. At low values of III, II increases with III as the temperature of the gas rises, and then falls as the abundance of III increases at the expense of II. For III less than unity, this trend is fit by a line

$$\log \mathrm{II} = 0.4 + 0.4 \log \mathrm{III} \pm 0.25. \tag{3}$$

Now, as shown in the Appendix of DTJS by model-fitting, following Pagel *et al.* (1979),

$$\Delta \log (\mathrm{O/H}) \doteqdot -\Delta \log (\mathrm{II} + \mathrm{III}). \tag{4}$$

For small III, this gives approximately

$$\Delta \log (\mathrm{O/H}) \doteqdot -0.4\, \Delta \log \mathrm{III}. \tag{5}$$

For M83, a straight line fitted to our values of log III versus R/R_0 gives a slope of 2.2. Here R_0 is the isophotal radius of de Vaucouleurs *et al.* (1976) as corrected by DTJS. A point of confusion is that although they use this value to plot their data, they use a different value, obtained from their previous paper (Talbot, Jensen & Dufour 1979), in their Table 4 which

contains the abundance gradient. It is a good approximation to treat III ≪ II in M83, and so $O/H \propto II^{-1}$, and this gives their slope -0.8 for $\log(O/H)$ versus R/R_0 if the corrected de Vaucouleurs R_0 is used. Their values for III, multiplied by 4/3 since they quote the λ 5007 line strength only, are plotted in Fig. 1 and are seen to be entirely in accord with our values. However, the slope of log III versus R/R_0 for their points is half that deduced from our values. A large part of the discrepancy is due to their region VI, which runs against the trend of a simple gradient with R/R_0. This kind of variation should be borne in mind when discussing gradients. In any event, using our gradient, and the relationship deduced in equation (5) above, we estimate the gradient of $\log(O/H)$ versus R/R_0 to be -0.9, close to the value deduced by DTJS. This is fortuitously close, and due to the fact that we compensate for the larger slope in log III versus R/R_0 of our data by a smaller coefficient in equation (5) than would be deduced from the DTJS data alone.

For NGC 7424, we have III > II. In this case the trend from DTJS Fig. 4 is

$$\log II = 0.4 - 0.6 \log III \pm 0.25. \tag{6}$$

Fitting a straight line over the range $1 < III < 5$ gives again from (4)

$$\Delta \log(O/H) = -0.4\, \Delta \log III.$$

Our data is fit by a line of slope (log III versus R/R_0) 0.9, and so our estimated slope for O/H is

$$d \log(O/H)/d(R/R_0) = -0.4,$$

half that of M83.

The nitrogen line ratio is consistent with a nearly constant ratio in M83 equal to 0.36 ± 0.07, except for the nucleus with nearly twice this value. The [S II]/Hα ratio is also gradientless, value 0.13 ± 0.03, with a significantly higher nuclear value. This run of [N II] with [S II] confirms the suggestion in DTJS that nitrogen is a product of primary nucleosynthesis.

In NGC 7424, the gradient of log [N II]/Hα versus R/R_0 is 0.4, the same as the inferred log (O/H) gradient, while the gradient in log [N II]/[S II] is 0.1, again indicative of primary nucleosynthesis for nitrogen. The mean value of [N II]/Hα is 0.13 ± 0.04.

Finally, we note that the ratio of the mean value of [O III]/Hβ in M83 to that in NGC 7424 is 0.06, the corresponding nitrogen ratio is 2.8, and the ratio of surface brightness (Table 1) is approximately 10. Comparison of line ratios with results by Shields & Searle (1978) suggest that NGC 7424 is not very different in its abundances from M101, region 10.

Acknowledgments

PWJLB thanks Dr Barlow for prompt assistance in providing a correct calibration source scan. We thank PATT for awarding telescope time, the UK Schmidt Telescope Unit for the use of plates, and the staff of the AAO for their friendly help. We also thank a referee for detailed and helpful criticisms.

References

Allen, C. W., 1973. *Astrophysical Quantities*, Athlone, London.
Alloin, D., Collin-Souffrin, S. & Joly, M., 1979a. *Astr. Astrophys. Suppl.*, 37, 361.
Alloin, D., Collin-Souffrin, S., Joly, M. & Vigroux, L., 1979b. *Astr. Astrophys.*, 78, 200.
Carbon, D. F. & Gingerich, O., 1969. In *Theory and Observation of Normal Stellar Atmospheres*, ed. Gingerich, O., MIT, Cambridge, Massachusetts.

Coulson, I. M., 1980. *Thesis*, University of Edinburgh.
de Vaucouleurs, G., 1979a. *Astrophys. J.*, **227**, 729.
de Vaucouleurs, G., 1979b. *Astr. J.*, **84**, 1270.
de Vaucouleurs, G., de Vaucouleurs, A. & Corwin, H., 1976. *Second Reference Catalogue of Bright Galaxies*, University of Texas Press, Austin.
Dufour, R. J., Talbot, R. J., Jensen, E. B. & Shields, G. A., 1980. *Astrophys. J.*, **236**, 119 (DTJS).
Jensen, E. B., Strom, K. M. & Strom, S. E., 1976. *Astrophys. J.*, **209**, 748.
Kurucz, R. E., Peytremann, E & Avrett, E. H., 1974. Smithsonian Institution, Washington.
Mathis, J. S., 1970. *Astrophys. J.*, **159**, 263.
Osterbrock, D. E., 1974. *Astrophysics of Gaseous Nebulae*, Freeman, San Francisco.
Pagel, B. E. J., Edmunds, M. G., Fosbury, R. A. E. & Webster, B. L., 1978. *Mon. Not. R. astr. Soc.*, **184**, 569.
Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S. & Smith, G., 1979. *Mon. Not. R. astr. Soc.*, **189**, 95.
Panagia, N., 1974. *Astr. J.*, **78**, 929.
Searle, L., 1971. *Astrophys. J.*, **168**, 327.
Shields, G. A. & Searle, L., 1978. *Astrophys. J.*, **222**, 821.
Smith, H. E., 1975. *Astrophys. J.*, **199**, 591.
Talbot, R. J., Jensen, E. B. & Dufour, R. J., 1979. *Astrophys. J.*, **229**, 91.
Webster, B. L., 1977. In *Topics in Interstellar Matter* ed. van Woerden, H., Reidel, Dordrecht.

Mon. Not. R. astr. Soc. (1987) **228**, *Short Communication*, 11P–16P

The origin of the far-infrared flux from spiral galaxies

S. K. Leggett and P. W. J. L. Brand *Astronomy Department, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

C. M. Mountain *Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ*

Accepted 1987 July 15. Received 1987 July 10; in original form 1987 June 18

Summary. Studies of the Virgo cluster spiral galaxies have shown them to have the same mid- and far-infrared properties as redshift-limited samples of field galaxies. This is true even of the galaxies in the core of the cluster which have been stripped of neutral hydrogen. These findings appear to disagree with current models of the far-infrared emission from galaxies which conclude that it is dominated by disc emission from cold H I-associated dust. As the molecular content of the Virgo galaxies is normal, we suggest that the far-infrared flux from spirals is in fact dominated by emission from dust associated with the more centrally concentrated molecular gas.

1 Introduction: the problem

This work arose from an infrared–optical study by Leggett *et al.* (1987) of sources found in the Virgo region by the *Infrared Astronomical Satellite*, *IRAS* (Beichman *et al.* 1985). That study is part of a large-scale programme, the optical identification of *IRAS* point sources using the plate measuring machine COSMOS (Wolstencroft *et al.* 1986).

Leggett *et al.* showed that the majority of *IRAS* galaxies have the same properties as typical Virgo cluster galaxies. These are bright ($B<14$) spiral galaxies of type Sa–d, with far-infrared (FIR) luminosity about equal to optical luminosity, $F_\nu(100\,\mu\text{m})/F_\nu(60\,\mu\text{m})\sim 3$ and $L(\text{FIR})\sim 10^9\,L_\odot$. It is apparent that the cluster environment has no effect on FIR properties; the cluster galaxies are indistinguishable in the FIR from field galaxies at similar redshifts.

The sample of field galaxies used for this comparison was the unbiased, volume-limited sample of Rieke & Lebofsky (1986). Rieke & Lebofsky find that $L(\text{FIR})$ correlates with $F(60\,\mu\text{m})/F(100\,\mu\text{m})$ and $L(\text{FIR})/L(B)$, and that Sb–c spiral types are the most infrared-luminous, with $L(\text{FIR})\sim 10^{10}\,L_\odot$.

Comparisons between Virgo cluster galaxies and redshift-limited samples of field galaxies have also been made by Devereux, Becklin & Scoville (1987). They find no difference in either 10-μm or FIR ($60\,\mu\text{m}+100\,\mu\text{m}$) luminosity between the two samples. A study of seven other clusters by Bicay & Giovanelli (1987) finds that these galaxies are also 'IR normal'.

These results are surprising, as recent models of the far-infrared emission from disc galaxies associate this emission with the infrared 'cirrus' which was first detected in our own Galaxy at 100 μm by *IRAS*. This diffuse emission from dust heated by the interstellar radiation field has been associated with the interstellar atomic gas. As many of the cluster galaxies are strongly H I-deficient, the models imply that they should be distinguishable in the FIR from field spirals.

2 Discussion

2.1 COMPARISON OF THE *IRAS* DATA WITH THE MODELS

The 100-μm flux from Galactic cirrus has been successfully reproduced using composite models for interstellar grains and an interstellar radiation field similar to that of the solar neighbourhood (e.g. Draine & Lee 1984; Draine & Anderson 1985; Rowan-Robinson 1986). These models assume that interstellar gas and dust are well mixed and that there is a fairly constant gas-to-dust ratio. In the Galaxy the neutral hydrogen maps and infrared cirrus do not correspond perfectly (e.g. Low *et al.* 1984; Persson & Helou 1987; Rowan-Robinson 1986), although correlations at high latitudes have been observed (e.g. Fong *et al.* 1987; Terebey & Fich 1986).

Recent interpretations of the far-infrared flux from our Galaxy and other galaxies (e.g. Chini *et al.* 1986; Cox, Krügel & Mezger 1986; Persson & Helou 1987; Rowan-Robinson & Crawford 1986) consist of multi-temperature dust components, simplified into warm (~40 K) and cool (~20 K) components. The warm component is associated with star-forming regions, and the cool component with grains heated in the diffuse interstellar radiation field, the infrared 'cirrus'. In this model the flux can be predicted by

$$F_\nu = \frac{3MB(\nu, T)\, Q(\nu)}{4a\varrho D^2} \tag{1}$$

if the total dust mass M, blackbody intensity $B(\nu, T)$, grain emissivity $Q(\nu)$, grain size a, grain density ϱ, and distance D, for optically thin emission (Hildebrand 1983) are known.

Persson & Helou (1987) find that Draine & Anderson's models of the far-infrared emissivity per H-atom agree with the observed ratios of 100-μm flux to H I flux for their field sample. Their sample consists of normal disc galaxies (studied in Hα by Kennicutt & Kent 1983) and they

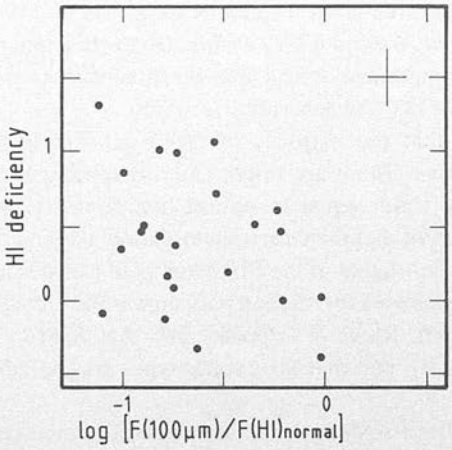


Figure 1. H I deficiency [$\Delta(\log M)$] from Haynes & Giovanelli (1986) against *IRAS* $F_\nu(100\,\mu\text{m})$ normalized by dividing by the expected H I flux from a galaxy of the same morphological type and diameter (units km^{-1}s). A typical error bar is also shown.

conclude that such galaxies have 50–70 per cent of their far-infrared emission due to cirrus. These findings are consistent with the other models mentioned above.

Fig. 1 shows a plot of H I deficiency [$\Delta(\log M)$] against normalized 100-μm flux for the 23 Virgo *IRAS* galaxies for which Haynes & Giovanelli (1986) give deficiencies. The 100-μm flux has been normalized by dividing by the expected H I flux from a galaxy of the same morphological type and size, taken from Haynes & Giovanelli. The Spearman's ϱ correlation coefficient is -0.25, and so, although there is a large spread in the normalized 100-μm flux, it appears to be independent of H I deficiency; galaxies which are deficient by more than a factor of 10 in H I mass are still detected by *IRAS*. There is no evidence of a decrease in flux of a factor corresponding to the deficiency, as would be expected from equation (1) for emission from H I-associated dust.

The infrared flux will increase if the interstellar radiation field increases due to an increase in dust temperature, as F is proportional to $B(\nu, T)$ (equation 1). However, it can be demonstrated that the H I-deficient galaxies do not have a higher radiation field; a plot of the ratio of the 100-μm flux to 60-μm flux (a reasonable measure of the dust temperature) against the ratio of 100-μm flux to observed H I flux, Fig. 2, shows a uniform distribution in colour temperature. The Spearman's ϱ correlation coefficient is -0.01. Thus radiation excess is not compensating for H I deficit.

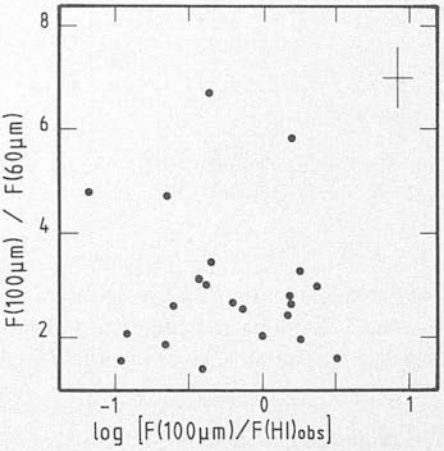


Figure 2. The flux ratio $F_\nu(100\,\mu\text{m})/F_\nu(60\,\mu\text{m})$, from *IRAS*, against the ratio of $F_\nu(100\,\mu\text{m})/F$[H I(observed) km^{-1}s], using data from *IRAS* and Haynes & Giovanelli (1986). A typical error bar is also shown.

2.2 COMPARISON OF THE *IRAS* DATA WITH OBSERVATIONS

Observations of the molecular content of Virgo galaxies show them to be normal. The observed CO luminosity is normal and independent of H I deficiency, as shown for example by Kenney & Young (1986) and Stark *et al.* (1986). These authors find that the core of the cluster is H I/CO deficient, and that these galaxies have smaller $M(\text{H I})/L(B)$ than field galaxies, but the same $M(\text{H}_2)/L(B)$. Stark *et al.* find that the total CO emission correlates with radio continuum emission and far-infrared (60 μm + 100 μm) emission. The H I flux, on the other hand, shows no correlation with FIR flux (*cf.* our Fig. 1), and the CO flux correlates only poorly with blue luminosity. The CO diameter is also normal, while the H I diameter for the deficient galaxies is smaller than that of field galaxies. Kenney & Young suggest that the events which remove atomic gas (ram-pressure stripping, thermal evaporation, or turbulent viscous stripping) have taken place over at least a cluster crossing time ($\sim 10^9$ yr), and that the molecular clouds in galaxies may be long-lived.

Table 1. Various data for the sample of Virgo spiral galaxies: for a full description and details of sources see text.

Name NGC/ other	HI Def.	CC 60μm 100μm	PS F Jy 60μm	PS F Jy 100μm	SSS F Jy 60μm	LOG L_{IR} /L_B	LOG F(100μm) /F(HI) $Km^{-1}s$	F(CO) 10^{-19} Wm^{-2}	SSS D(60μm) arcmin	D(opt) arc- min	D(HI)/ D(opt)
4178	-0.13	KB	1.52	8.10	3.92	-0.33	--0.92	4.2(3.3)	2.3	4.2	1.5
4192	0.09	BA	5.02	18.38	9.86	-0.78	-0.66	20(2)	2.0	7.8	1.3
4206	-0.09	DA	0.52	2.48		-0.43	-1.19	3.6(1.8)		3.9	
4216	0.35	BF	0.93L	7.42	3.51[a]	>-0.87	-0.66	8(5)	a	6.3	0.9
4254	0.02	CA	22.83	71.28	35.20	-0.31	-0.00	100(20)	1.8	5.5	
4316	0.38	AA	1.60	5.49		0.28	-0.36				
4321	0.52	BA	17.98	56.55	23.40	-0.65	0.17	169(40)	1.6	7.1	1.0
4383	-0.39	AA	8.53	11.98		0.22	-0.41			2.0	
4388	1.06	AA	10.73	17.27		0.21	0.51	38(3.5)		4.0	0.3
4402	0.61	AA	5.69	16.94		-0.12	0.37	30(1.8)		3.2	0.6
4450	1.31	CA	1.18	6.85		-0.66	0.19	2.6(1.9)		4.7	0.8
4498	0.47	BA	1.19	3.70		-0.20	-0.44			2.9	
IC 797	0.20	BA	0.95L	2.05		>-0.38	-0.28				
4501	0.47	CA	13.96	54.38	16.70	-0.02	0.25	158(45)	1.6	6.3	0.9
4519	-0.33	AA	3.71	5.78		-0.46	-0.96			3.0	
4522	0.51	BA	1.29	3.89		-0.59	-0.39			3.0	
4535	0.17	DB	6.42	20.88	9.34[b]	~-0.58	-0.61	83(10)	1.7[b]	6.6	1.1
4548	0.86	CA	1.21	7.76	3.05	-0.70	-0.14	14(3)	2.3	5.4	1.0
4569	0.99	CA	7.08	22.72	11.60	-0.37	0.26	170(90)	1.8	8.3	0.3
4571	0.44	DA	0.94	6.28		-0.54	-0.38			3.8	
4579	1.01	CA	4.46	17.38	6.64	-0.93	0.19	30(15)	1.8	5.4	0.6
4607	0.72	AA	2.71	7.55		-0.03	0.18			2.4	
4654	0.00	AA	12.96	34.42		-0.33	-0.21	42(4)		4.4	1.3

Notes:

(a) NGC 4216 SSS $F(100\mu m) = 16.50$ Jy, $D(100\mu m) = 4.7'$.

(b) NGC 4535 SSS $F(100\mu m) = 24.30$ Jy, $D(100\mu m) = 3.4'$.

Table 1 summarizes some of these data for the 23 Virgo spirals in our sample. These spirals are of type a–cd, and the sample contains both barred and unbarred galaxies (for more details and finding charts see Leggett *et al.* 1987). The table gives the following information:

Column 1: Name;

Column 2: H I deficiency (Haynes & Giovanelli 1986);

Column 3: *IRAS* point-source correlation coefficient (Beichman *et al.* 1985), which can have a value between 87 and 100 per cent, encoded alphabetically with $A = 100$, $B = 99$, $N = 87$;

Column 4: *IRAS* point source 60-μm flux (Beichman *et al.* 1985);

Column 5: *IRAS* point source 100-μm flux (Beichman *et al.* 1985);

Column 6: IRAS small-scale structure 60-μm flux (Helou & Walker 1986);

Column 7: Log of the ratio of infrared to blue luminosity (Leggett *et al.* 1987);

Column 8: Log of the ratio of 100-μm flux to H I 21-cm line flux ($\int S dV$, Haynes & Giovanelli 1986);

Column 9: CO flux (Stark *et al.* 1986);

Column 10: *IRAS* small-scale structure 60-μm diameter (Helou & Walker 1986);

Column 11: Optical diameter (the *Second Reference Catalogue of Bright Galaxies*, de Vaucouleurs, de Vaucouleurs & Corwin 1976);

Column 12: Ratio of H I diameter (van Gorkom & Kotanyi 1985) to optical diameter.

The full-width half-maximum of the *IRAS* 60-μm beam is 1.44 arcmin, and of the 100-μm beam is 3.14 arcmin. Table 1 shows that 9 of the 23 Virgo *IRAS* galaxies are extended at 60 μm, but only 2 were detected beyond the larger 100-μm beam. This implies that the 100-μm diameter for these galaxies is not much larger than the 60-μm diameter.

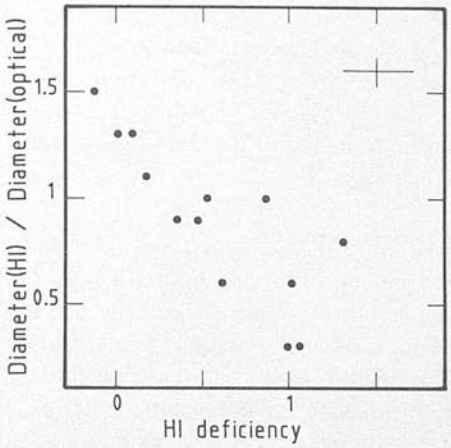


Figure 3. The ratio of the H I diameter (van Gorkom & Kotanyi 1985) to optical diameter (*Second Reference Catalogue of Bright Galaxies*) against H I deficiency [$\Delta(\log M)$, Haynes & Giovanelli 1986]. A typical error bar is also shown.

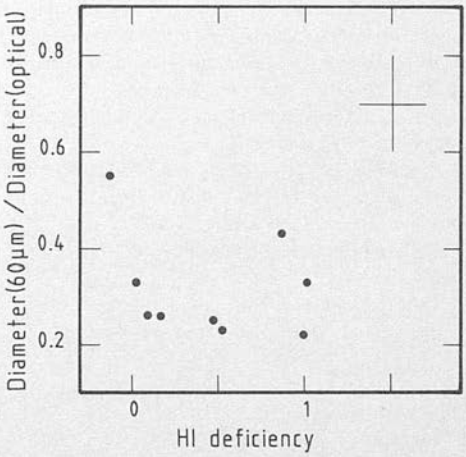


Figure 4. The ratio of the *IRAS* 60 μm diameter to optical diameter (*Second Reference Catalogue of Bright Galaxies*) against H I deficiency [$\Delta(\log M)$, Haynes & Giovanelli 1986]. A typical error bar is also shown.

Fig. 3 shows that the H I diameter (normalized by dividing by the optical diameter) decreases with increasing H I deficiency, as previously stated. The Spearman's ϱ correlation coefficient is -0.86. Fig. 4 plots the ratio of the 60-μm diameter to the optical diameter against H I deficiency; this diagram indicates that $D(60\,\mu\text{m})/D(\text{optical})$ is approximately constant and independent of H I deficiency (Spearman's ϱ correlation coefficient -0.37). Upper limits for the ratio $D(60\,\mu\text{m})/D(\text{optical})$ can be derived for the galaxies that were not seen to be extended by *IRAS*, using the *IRAS* beamsize. These limits are consistent with the conclusion that the far-infrared diameters of the Virgo spirals are approximately constant and about one-quarter to one-third of the size of the optical diameter.

Kenney & Young show graphically that the ratio $D(\text{CO})/D(\text{H I})$ varies from ~ 1.0 at the cluster core to ~ 0.5 for galaxies $\geqslant 3°$ from M87. It appears therefore that both the 60-μm diameter and the CO diameters are approximately constant and independent of H I deficiency. Moreover $D(\text{CO})$ and $D(60\,\mu\text{m})$ are approximately equal, with values between one-quarter and one-half of

the optical diameter.

As the FIR flux correlates well with CO emission and only poorly with H I emission, and the CO and 60-μm diameters of the Virgo spirals are measured to be approximately equal, we suggest that the far-infrared flux from normal disc galaxies is not dominated by emission from dust associated with H I, but in fact by emission from dust associated with the molecular gas.

3 Conclusion

We have shown that current models of the far-infrared emission from spiral galaxies cannot account for the *IRAS* observations of Virgo galaxies. On the other hand there are strong similarities between the FIR and molecular properties of both Virgo cluster and field spirals. We suggest that although the FIR flux from spiral galaxies may arise in part from diffuse cirrus emission as in our Galaxy, it originates predominantly from the star forming regions of the CO-rich disc. Sub-millimetre and millimetre observations are needed to confirm this.

References

Beichman, C. A., Neugebauer, G., Habing, H. J., Clegg, P. E. & Chester, T. J., 1985. *Explanatory Supplement to the IRAS Catalogs and Atlases*, US Government Printing Office, Washington DC.

Bicay, M. D. & Giovanelli, R., 1987. *Proc. Conf. Star Formation in Galaxies*, in press.

Chini, R., Kreysa, R., Krügel, E. & Mezger, P. G., 1986. *Astr. Astrophys.*, **166,** L8.

Cox, P., Krügel, E. & Mezger, P. G., 1986. *Astr. Astrophys.*, **155,** 380.

de Vaucouleurs, G., de Vaucouleurs, A. & Corwin, H. G., 1976. *Second Reference Catalogue of Bright Galaxies*, Monographs in Astronomy No. 2, University of Texas.

Devereux, N. A., Becklin, E. E. & Scoville, N., 1987. *Astrophys. J.*, **312,** 529.

Draine, B. T. & Anderson, N., 1985. *Astrophys. J.*, **292,** 494.

Draine, B. T. & Lee, H. M., 1984. *Astrophys. J.*, **285,** 89.

Fong, R., Jones, L. R., Shanks, T., Stevenson, P. R. F., Strong, A. W., Dawe, J. A. & Murray, J. D., 1987. *Mon. Not. R. astr. Soc.*, **224,** 1059.

Haynes, M. P. & Giovanelli, R., 1986. *Astrophys. J.*, **306,** 466.

Helou, G. & Walker, D. W., 1986. *IRAS Small Scale Structure Catalog*, US Government Printing Office, Washington DC.

Hildebrand, R. H., 1983. *Q. Jl R. astr. Soc.*, **24,** 267.

Kenney, J. D. & Young, J. S., 1986. *Astrophys. J.*, **301,** L13.

Kennicutt, R. C. & Kent, S. M., 1983. *Astr. J.*, **88,** 1094.

Leggett, S. K., Clowes, R. G., Kalafi, M., MacGillivray, H. T., Puxley, P. J., Savage, A. & Wolstencroft, R. D., 1987. *Mon. Not. R. astr. Soc.*, **227,** 563.

Low, F. J. *et al.*, 1984. *Astrophys. J.*, **278,** L19.

Persson, C. J. L. & Helou, G., 1987. *Astrophys. J.*, **314,** 513.

Rieke, G. H. & Lebofsky, M. J., 1986. *Astrophys. J.*, **304,** 326.

Rowan-Robinson, M., 1986. *Mon. Not. R. astr. Soc.*, **219,** 737.

Rowan-Robinson, M. & Crawford, J., 1986. *Light on Dark Matter*, ed. Israel, F. P., Reidel, Dordrecht, Holland.

Stark, A. A., Knapp, G. R., Bally, J., Wilson, R. W., Penzias, A. A. & Rowe, H. E., 1986. *Astrophys. J.*, **310,** 660.

Terebey, S. & Fich, M., 1986. *Astrophys. J.*, **309,** L73.

van Gorkom, J. & Kotanyi, C., 1985. *ESO Workshop on the Virgo Cluster of Galaxies*, p. 61, eds Richter, O.-G & Binggeli, B., European Southern Observatory, Garching.

Wolstencroft, R. D., Savage, A., Clowes, R. G., MacGillivray, H. T., Leggett, S. K. & Kalafi, M., 1986. *Mon. Not. R. astr. Soc.*, **223,** 279.

[ISSN 0911-5501]

NRO

Nobeyama
Radio
Observatory
Report
No. 223

37

Detection of H53α Emission from M82: A Reliable Measure of the Ionization Rate and its Implications

P.J. Puxley, P.W.J.L. Brand, T.J.T. Moore, C.M. Mountain,
N. Nakai, and T. Yamashita

NOBEYAMA RADIO OBSERVATORY OF THE NATIONAL ASTRONOMICAL OBSERVATORY
NOBEYAMA, MINAMISAKU, NAGANO-KEN 384-13, JAPAN

To appear in Oct 1 1989 Ap. J.

Detection of H53α Emission from M82 : A Reliable Measure of the Ionization Rate and its Implications

P.J. Puxley [1], **P.W.J.L. Brand & T.J.T. Moore**

Dept. of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ.

C.M. Mountain

Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ.

N. Nakai & T. Yamashita

Nobeyama Radio Observatory, Nobeyama, Minamisaku, Nagano 384-13, Japan.

Received 1988 August 31; accepted 1989 March 9

Abstract

Emission in the hydrogen recombination line H53α has been detected, for the first time, in another galaxy : the nearby 'starburst' system M82. This line is produced primarily by spontaneous emission and provides the most direct, extinction independent, estimate of the ionization rate in the star-forming complex. The line strength implies an ionization rate of 1.1×10^{54} s^{-1}, approximately 5 times larger than that inferred from Brα observations, and indicates a dust extinction at 4 μm of more than 1 mag ($A_V \simeq 27$ mag). Comparison of

[1]Present address : Institute for Astronomy, 2680 Woodlawn Drive, Honolulu, Hawaii 96822

the 3.3-mm free–free continuum and H53α fluxes implies an average electron temperature of $T_e \sim 5 \times 10^3$ K. Analysis of the line excitation conditions, using the H53α emission in conjunction with that in the far–IR O^{2+} and N^{2+} forbidden lines, together with the IR luminosity, suggests that only a very restricted range of stellar masses are formed. Less plausible interpretations are (1) that approximately two thirds of the direct stellar emission escapes from the central $40\widehat{''}$ without encountering dust, and (2) that much of the line emission and IR luminosity arises via fast ($\sim$100 km s^{-1}) shocks.

Subject headings: galaxies : individual (M82) – galaxies : interstellar matter – interstellar : grains

1 Introduction

The galaxy M82 has been well–studied throughout the observable electromagnetic spectrum because of its propinquity. Infra–red observations, which are less affected by the heavy extinction within this edge–on system, have been instrumental in revealing that the IR luminosity of the galaxy is dominated by a population of young stars in the central regions; see Telesco (1988) for a review. However, even in the near–IR, attenuation by dust cannot be ignored. In this paper we present the first detection in another galaxy of the n=54→53 transition of atomic hydrogen (H53α) which occurs at a rest frequency of 42.952 GHz. At this frequency the line emission is dominated by spontaneous emission and thus provides a reddening–independent estimate of the photo–ionization rate and physical conditions within the ionized regions (§3). In Section 4 we construct a simple model for the far–IR and

millimetre–wavelength line emission and IR luminosity in terms of an ensemble of dust–enshrouded photo–ionized regions, and investigate the implications of this model for the stellar initial mass function, elemental abundances and source geometry.

2 Observations

Observations of M82 were made during 1988 April using the 45m telescope of the Nobeyama Radio Observatory, simultaneously in the 53α and 40α transitions of hydrogen at frequencies of 43 and 99 GHz respectively. The instrumentation comprised SIS (H53α) and cooled Schottky diode (H40α) frontends together with 2048–channel wide–band acousto–optical spectrometer backends, which gave a frequency resolution of 250 kHz and an instantaneous band–width of 256 MHz. The system noise temperatures (T_{SSB}) during the observations were ~300 and ~670 K at 43 and 99 GHz respectively. Aperture efficiencies, derived from measurements of planetary antenna temperatures (assuming brightness temperatures respectively for Mars, Venus and Saturn of 200, 400 and 135 K, at 43 GHz, and 210, 300 and 150 K, at 115 GHz) and half–power beam widths obtained from cross scans of the point sources 3C84 and 3C273, were η_A=0.60 and 0.34 at 43 and 99 GHz, respectively, with a statistical error of $\pm$0.03 (see Kraus 1986 for details). The half–power beam widths were 41 $\hat{''}$ and 19$\hat{''}$.

The telescope was pointed at $\alpha(1950) = 9^h\,51^m\,42.5^s$, $\delta(1950) = +69^\circ\,55\hat{'}\,00\hat{''}$, *i.e.* 5$\hat{''}$ West of the 2- μm peak and HI dynamical centre (Joy, Lester & Harvey 1987; Weliachew, Fomalont & Greisen 1984) and coincident with the 10- μm centroid (Rieke *et al.* 1980). The extent

of the central star-forming complex is about 25″ as determined from Brγ recombination line observations by Lester and co-workers (quoted in Telesco 1988), and thus we expect our H53α beam to encompass all of the detectable emission. The pointing was checked regularly using a nearby SiO maser source, and we estimate the overall accuracy to be ±3 ″. Calibration was performed using a warm chopper in front of the receiver, and converted to flux density using the efficiencies quoted above. Linear baselines were removed and the data re-binned into channels of width 8 MHz. The resulting spectra are shown in Figure 1, together with the [OIII]λ88 μm profile from Duffy *et al.* (1987).

The line fluxes integrated between -100 and +500 km s^{-1} are $F(H53\alpha) = (1.87 \pm 0.28) \times 10^{-20}$ W m^{-2} and $F(H40\alpha) = (9.22^{+1.38}_{-1.84}) \times 10^{-20}$ W m^{-2}. The error in the former value is primarily the result of uncertainty in the exact aperture efficiency, while the larger asymmetric error of the latter reflects uncertainty in the baseline, which may not be linear. Detection of H53α emission from M82 was confirmed the following night, although the weather conditions were poor and therefore these data are not included in the following analysis. Comparison with measurements of the radio hydrogen recombination lines (*e.g.* Bell & Seaquist 1978; Shaver, Churchwell & Walmsley 1978), which are due to stimulated emission and whose strength decreases with increasing frequency (mainly due to the spectral index of the non-thermal, stimulating continuum radiation), indicates that less than 3% of the H53α flux is produced in this way : the line is generated primarily by spontaneous emission. We expect negligible contributions to the integrated line fluxes from the equivalent helium transitions He40α and He53α ($\lesssim$5%, Peimbert & Spinrad 1970; velocity shift -122 km s^{-1}), and from other high n-level hydrogen lines.

3 Hydrogen Recombination Line Emission

Examination of Figure 1 shows that the H53α profile, in particular, bears a strong resemblance to the far–IR line [OIII]λ88 μm (and also [OIII]λ52 and [NIII]λ57 μm; Lugten *et al.* 1986; Duffy *et al.* 1987), as expected due to their common origin in regions of ionized gas. The shapes of the H40α and H53α lines appear somewhat different, but are probably consistent within the errors. In Table 1, rows 1 & 2, we present theoretical intensities in the H40α and H53α transitions relative to Hβ for a range of electron temperatures and densities, and accurate to about 1% (P.J. Storey - personal communication; see Hummer & Storey (1987) for details of the methods used). Case B recombination was assumed (Baker & Menzel 1938). The observed line ratio $H40\alpha/H53\alpha = (4.9^{+1.0}_{-1.2})$ is consistent with the theoretical values for the range of electron densities and temperatures considered here. However, the measured H40α flux is likely to be an underestimate since the telescope beam at 99 GHz ($19\widehat{''}$) is somewhat smaller than the extent of the recombination line emission ($\sim 25\widehat{''}$). This suggests that the electron density may lie towards the lower end of the range, although we cannot provide much constraint on its value. Given the lower signal–to–noise ratio of the H40α detection, and its less certain baseline, we will consider only the H53α measurement in the rest of this paper.

The electron density in M82 has been determined using a variety of techniques, but those involving far–IR forbidden or radio recombination lines are much preferred since they are less affected by extinction. Observations of the density–sensitive ratios [OIII]λ88/λ52 and [SIII]λ33/λ18 imply an electron density of $n_e \simeq 200$ cm^{-3} (Duffy *et al.* 1987; Houck *et*

al. 1984), which is consistent with *some* analyses of the radio hydrogen recombination lines (*e.g.* Bell & Seaquist 1978; Rodriguez & Chaisson 1980). However, Lugten *et al.* (1986) have also measured the ratio [OIII]λ88/λ52 and, whilst agreeing with Duffy *et al.*'s value of n_e for a single–density model, they find that the data can also be decomposed into two components. One component, having $V_{LSR} \sim 150$ km s^{-1} and contributing of the order of two–thirds of the integrated line flux (see also Figure 1), implies $n_e \sim 600$ cm^{-3}, whereas the other gives $n_e \sim 60$ cm^{-3}. This two–component density model is supported by the radio hydrogen recombination line results of Seaquist, Bell & Bignell (1985). This disagreement probably arises in part from the different flux calibrations and beam sizes used for the [OIII] measurements, and also from low signal–to–noise ratios in the higher frequency radio recombination line data. In this paper we adopt a single–component electron density model with n_e=100 cm^{-3} for simplicity, but, where appropriate, we also investigate the effects of employing a two–component model with n_e=100 and 1000 cm^{-3}, comprising one–third and two–thirds of the total line fluxes, respectively.

Together with observations of the 3.3-mm continuum emission, our measure of the H53α flux allows calculation of the average electron temperature in the ionized gas. Following Rubin (1968) we can write:

$$\left(\frac{F(H53\alpha)}{Wm^{-2}}\right) = 1.78 \times 10^{-16} \left(\frac{N_{H53\alpha}}{N_{Lyc}}\right)_{CaseB} \left(\frac{S_\nu^{th}}{Jy}\right) \left(\frac{\nu}{GHz}\right)^{0.1} \left(\frac{T_e}{10^4 K}\right)^{-0.45}$$

where $N_{H53\alpha}/N_{Lyc}$ is the predicted number of H53α photons produced per recombination (see Table 1, row 3) and S_ν^{th} is the free–free flux density at a frequency ν. This relation is insensitive to the electron density. Integration over the aperture–synthesis map by Carl-

strom (1988) gives $S(91GHz) = 0.55$ Jy, which is consistent with the large beam ($\sim 75''$) measurement by Jura, Hobbs & Maran (1978) of $S(87GHz) = (0.54 \pm 0.08)$ Jy. Assuming that this emission is entirely due to bremsstrahlung, we derive an electron temperature of $T_e \sim 5 \times 10^3 K$. Thence we conclude that contributions to the 90 GHz flux due to thermal emission from cold dust and from synchrotron emission must be negligible (*c.f.* Klein, Wielebinski & Morsi 1988), otherwise the inferred electron temperature would be unfeasibly low. For example, if 10% of the observed 90 GHz continuum emission was non–thermal (*c.f.* the predicted contribution from known supernova remnants of about 3%, assuming a spectral index for the discrete sources studied by Kronberg, Biermann & Schwab (1985) of -0.9), then the equation above implies an electron temperature of less than 2900 K.

The ionization rate implied by the H53α flux is also insensitive to the electron density. Using the recombination coefficients given by Hummer & Storey (1987) and a distance to M82 of 3.25 Mpc (Tammann & Sandage 1968), we obtain $N_{Lyc} = (1.08 \pm 0.16) \times 10^{54}$ s^{-1} for n$_e$=100 cm^{-3} and T$_e$ = 5×10^3 K. For the two–component density model we calculate $N_{Lyc} = 1.03 \times 10^{54}$ s^{-1}. These values are more than 5 times larger than the ionization rate derived from the large–aperture Brα measurement by Willner *et al.* (1977), and imply significant extinction by dust even at 4 μm. Assuming an intrinsic Case B line ratio of $I(Br\alpha)/I(H53\alpha) = 2.29 \times 10^6$ (see Table 1) and that the attenuation at 4 μm is given by $A_{4-\,\mu\mathrm{m}} = 0.04 A_V$ (Rieke & Lebofsky 1985), the measured Brα flux of 1.6×10^{-14} W m^{-2} implies $A_V = 27 \pm 4$ mag. This is consistent within errors with estimates of the visual extinction deduced from the depth of the 10- μm silicate feature after correction for the narrow dust emission features ($A_V = 20 \pm 4$ mag; Aitken & Roche 1984).

The visual extinction derived from the H53α and Brα lines is considerably greater than that inferred from the optical Hβ and Hα lines, *e.g.* $A_V \sim 4$ mag (Peimbert & Spinrad 1970). Furthermore, the Hα flux of 1.2×10^{-12} W m^{-2} measured through a $35''$ aperture centred on the nucleus of M82 (J. Bland, personal communication, from data in Bland & Tully 1988), implies an ionization rate of 1.4×10^{54} s^{-1} *without reddening corrections* and using the same assumptions as those adopted above. This is clearly inconsistent with the H53α and Brα results, and the problem is exacerbated when one includes the effects of dust extinction. This disagreement cannot be caused by patchy extinction. To explain this discrepancy we suggest that much of the observed hydrogen line emission from the low n–levels arises in front of the star formation region and is generated by collisional excitation via shocks (see also McCarthy, Heckman & van Breugel 1987).

4 Modelling the Emission from M82

4.1 The Far–IR Forbidden and Millimetre–Wavelength Lines

The recent availability of high quality long–wavelength spectral line measurements of M82, made through similar large apertures, allows us to investigate quantitatively the global properties of the central star formation complex without the usual problems associated with dust obscuration and beam–size effects. Therefore, given the spectral similarity of the [OIII], [NIII] and H53α lines, and the expectation that the dominant sources of emission are the same (namely, ionized regions around hot stars), we have constructed a simple model of the far–IR and mm line emission from M82 in terms of an ensemble of such regions using

the numerical calculations of Rubin (1985, henceforth R85). Below, we describe in detail the values adopted for the model parameters.

Using the compilation of stellar properties in Landolt-Bornstein (1982), ionization rates (N_{Lyc}) were calculated by integrating under the Kurucz (1979) model atmospheres of stars with effective temperatures in the range $31,000 \leq T_{eff} \leq 45,000$ K. The [OIII]$\lambda 88\,\mu$m and [NIII]$\lambda 57\,\mu$m line intensities were scaled from R85 (where fixed values of N_{Lyc} were adopted) to be consistent with the derived ionization rates (see Puxley, Hawarden & Mountain 1988a for the values of $N_{Lyc}(mass)$ employed). The H53α fluxes were calculated from the adjusted tabulated Brγ strengths using the predicted Case B ratio given in Table 1 for n_e=100 cm^{-3} and T_e=5 $\times 10^3$ K. Stars with temperatures outside the range considered here do not affect the results; the UV fluxes from those with lower T_{eff} are insufficient to provide significant H^+, O^{2+} or N^{2+} abundances, while there are no stars with higher T_{eff} (see below).

Analysis of the argon and neon mid-IR line emission (Willner *et al.* 1977) suggests that Ar^+ and Ne^+ are the dominant ionization states. Following Willner *et al.*, and setting A_V=27 mag and $T_e = 5 \times 10^3$ K, the inferred abundances of these two elements are within $\pm$50% of their solar values. This result is insensitive to the range of electron densities considered here, and a similar value is obtained using $T_e = 10^4$ K. Initially, therefore, we have adopted solar abundances of oxygen and nitrogen. We note that if the elemental abundances were significantly less than their solar values (*e.g.* Z= $Z_\odot/10$), then the electron temperature within the nebulae would be more than 10^4 K due to the decreased abundance of the coolant species (R85). This would be inconsistent with the value of $T_e \sim 5 \times 10^3$

K found in Section 3, although we must be slightly cautious since the free–free emission coefficient decreases with increasing T_e and thus the observed emission is weighted to lower temperature regions. McCarthy, Heckman & van Breugel (1987) have analysed the optical line emission towards the nucleus of M82 and also conclude that the elemental abundances have roughly their solar values.

The values for the line emission arising in photo–ionized regions around individual stars together with the compilation of stellar main–sequence lifetimes and the initial mass function (IMF) given by Scalo (1986) were used to calculate the total [OIII]λ88, [NIII]λ57 and H53α fluxes from an ensemble of stars having masses distributed according to various present day mass functions; see Gehrz, Sramek & Weedman (1983); Telesco (1985); Puxley, Hawarden & Mountain (1988b) for further details of the methods used. The results of this modelling are shown in Figure 2.

The cross in the lower left–hand corner of Figure 2 shows the observed line ratios, and 1σ uncertainties, calculated using the forbidden line fluxes in Duffy *et al.* (1987): $F([OIII]) = (9.2 \pm 0.3) \times 10^{-14}$ W m^{-2} and $F([NIII]) = (4.3 \pm 0.3) \times 10^{-14}$ W m^{-2}, both measured through a $48\widehat{''}$ aperture. Representing the star formation complex as a single object yields an effective temperature for the ionizing source of T$\simeq$35,500 K, equivalent to an O8 star (Landolt–Bornstein 1982). This is in agreement with the upper limit of T<37,000 K implied by the non–detection of [ArIII]λ9 μm (Gillett *et al.* 1975; Willner *et al.* 1977; R85), and the value T$\sim$30,000 K from analysis of the ionization state of helium (Peimbert & Spinrad 1970).

In Figure 2 the point marked O represents the predicted [OIII]λ88/H53α and [NIII]λ57/

H53α flux ratios for an unevolved ("zero-age") starburst with stellar masses distributed according to the initial mass function (IMF) of Scalo (1986) (effectively $\alpha = -2.85$, $\psi(m) \propto m^{\alpha}$), and an upper mass limit of m_u=64 $M_\odot$ corresponding to T_{eff}=45,000 K. Point S shows the effect of using a flatter IMF ($\alpha = -2.35$, Salpeter 1955), and keeping the other model parameters constant. The connecting line O–E represents evolution with a constant rate of star formation. After 10^7 yr the stellar birth rates and death rates become equal and the line fluxes do not change (point E). It is immediately apparent that neither the original model, nor evolution, nor reasonable values for the IMF slope can reproduce the observed line ratios. Even employing a model in which the star formation rate varies as e^{-t/τ_0} (with $\tau_0 \sim 10^7 - 10^8$ yr; Rieke *et al.* 1980) can not reproduce the observed values - in fact this model is very similar to that with constant evolution. The reason for this, and for why evolution and IMF variations have only a small effect, is that the predicted line fluxes, and thus the line ratios, are dominated by the most massive stars. For example, the predicted [OIII] and [NIII] fluxes vary by roughly a factor of 10^5 over the range 18–64 $M_\odot$. Thus an exponential decrease in the rate of star formation having a time constant one order of magnitude larger than the lifetime of the most massive stars simply does not occur quickly enough to affect the line ratios dramatically. The presence of dust within the ionized regions can also be rejected as an explanation for the low excitation conditions, since this would require an unfeasible increase in the dust photo–absorption cross–section from the ionization potential of hydrogen to that of O^+ and N^+ (~35 eV); see Mathis (1986).

Two other vectors of note in Figure 2 are those ending in the points labelled D and N. The former illustrates the predicted line ratios for the unevolved model with m_u=64 $M_\odot$ but

adopting the two–component density model described in Section 3. The line ratios decrease by about 0.2 dex due to collisional de–excitation of the excited states of [OIII] and [NIII] in the n_e=1000 cm^{-3} component. Point N shows the effect on the single–component (n_e=100 cm^{-3}) model of reducing the oxygen and nitrogen abundances to 0.6× and 0.4× their solar values, respectively (*i.e.* model series N from R85). This decrease in the abundances dominates over the slight increase in the fractional abundances of O^{2+} and N^{2+}, which is due to a hardening of the stellar UV spectra with lower metallicity. The net result is a decrease in the line ratios in proportion to the change in the elemental abundances.

Figure 2 shows that dramatic variations in the predicted line ratios are due to changes in the upper mass limit (m_u) of the IMF. The solid line represents an unevolved starburst having $\alpha = -2.85$, and has various values of m_u marked along it. This line essentially also illustrates the evolution of the line ratios after a rapid switch–off of star formation (*i.e.* small τ_0 in Rieke *et al.'s* model), so that stars of progressively lower mass dominate the observed line emission. *This locus clearly demonstrates that there cannot currently be any stars in the central star–formation complex of M82 which are more massive than about 24* $M_\odot$ *(n_e=100 cm^{-3}, Z=*$Z_\odot$*) or 30* $M_\odot$ *if we assume Z*$\sim Z_\odot/2$ *and the two–component density model.* The absence of stars with masses greater than 24–30 $M_\odot$ implies that, unless there has been a dramatic decrease in the star–formation rate during the last $\sim 10^7$ yr [$\tau(25\,M_\odot) \sim 8 \times 10^6$ yr], the stellar *initial* mass function in M82 must also terminate at about this mass. In the following section we investigate the constraints on the lower IMF bound (m_l) implied by this value of m_u.

4.2 The Luminosity and Ionization Rate

Stellar luminosities and ionization rates were calculated for stars in the mass range 0.8 – 100 $M_\odot$ by integrating under the same Kurucz (1979) model atmospheres, for consistency. The total IR luminosity from the ensemble of stars was derived following the method of Section 4.1 and assuming that the entire stellar output is thermalised by dust within the effective IR beam used for the observations ($\sim 40''$). For any star formation model, the maximum value of the ratio of the ionization rate to IR luminosity occurs before any stars have started dying (provided, of course, that the IMF is constant in time), and is thus independent of the star formation rate. Since the ionization rate inferred from the H53α measurement is considerably larger than previously believed, we have used this maximum value of N_{Lyc}/L_{IR} to characterise a particular lower–upper mass (m_l,m_u) model. The observed ratio for the central $40''$ of M82 is $\log(N_{Lyc}/L_{IR}) = (43.56 \pm 0.11)$ assuming an IR luminosity of 3×10^{10} $L_\odot$ (Telesco & Harper 1980). The dependence of the predicted ratio on the value of the IMF slope is small: $\Delta\log(N_{Lyc}/L_{IR}) \sim 0.05$ for α=–2.85 and α=–2.35. For a given value of the upper mass limit there is a minimum value of m_l which can reproduce the observed ratio; a smaller m_l produces too much relative luminosity since the lower mass stars make no contribution to the integrated ionization rate. The permitted ranges of m_l and m_u are explored in Figure 3, where constraints on m_u imposed by the line excitation conditions are also illustrated.

If the assumptions outlined above for this model hold, then it is apparent that *star formation in the centre of M82 is restricted to only a narrow range of stellar masses.* For n_e=100 cm^{-3} and $Z = Z_\odot$ we see that only stars having $m_* = 20 - 24$ $M_\odot$ are allowed (for

2σ uncertainty in the ratio N_{Lyc}/L_{IR}). Taking values of the metallicity ($Z_{\odot}/2$) and electron density (two–component model; §3) which permit the widest range of stellar masses, we calculate $m_* = 17 - 30\ M_{\odot}(2\sigma)$. Even if the ratio were $N_{Lyc}/L_{IR} = 10^{43.23}$ (reflecting 3σ uncertainty), the allowed range would still only be 17 – 24 or 12 – 30 $M_{\odot}$, depending on Z and n_e. Further relaxation of these stringent constraints on the limits of the mass function would require greater uncertainty in the infrared luminosity (since this dominates the error quoted for N_{Lyc}/L_{IR}), a metallicity of one half the solar value, adoption of the two–component density model and the condition that the starburst be unevolved. The evolutionary phase of the starburst affects the inferred lower mass limit because, as time progresses, the longer–lived lower mass stars, which exhibit relatively small values of the ratio N_{Lyc}/L_{IR}, accumulate on the main sequence thereby lowering the integrated ratio. The high value of N_{Lyc}/L_{IR} observed at the present epoch would therefore imply an even larger ratio, and thus greater predominance of massive stars, in the past.

The range of stellar masses implied by this analysis is uncomfortably small, particularly since the 2- μm spectrophotometry by Rieke *et al.* (1980) of the stellar CO absorption feature produced in the atmospheres of red giant and supergiant stars may require a mass function extending to considerably lower mass than identified above, although the exact value is not specified by our simple model. Assuming that the stellar–atmosphere and ionized–region models are correct, we suggest two other possible explanations for this discrepancy.

(i) The assumption of complete thermalisation of the stellar radiation by dust within the central $40''$ may be incorrect. If sufficient stellar luminosity ($L_{ESC} \gtrsim 5 \times 10^{10}\ L_{\odot}$) can escape from the star–forming complex without encountering dust, then the ratio $N_{Lyc}/(L_{IR}+$

L_{ESC}) provides no constraint on the lower mass limit of the IMF. In other words, increasing the total luminosity ($L_{IR} + L_{ESC}$) allows the permitted range of stellar masses to extend to lower values.

Rieke *et al.* (1980) speculated that a luminosity of 10^{10} $L_\odot$, emanating directly from the photospheres of red giants, escapes perpendicular to the plane. Assuming their value for the total intrinsic luminosity ($L_{IR} = 4 \times 10^{10}$ $L_\odot$) broadens the allowed range of stellar masses to 17 - 24 $M_\odot$ (2σ) for n_e=100 cm^{-3} and Z= $Z_\odot$. This range is still very narrow, despite having allowed one quarter of the stellar radiation to escape. At least 60% of the total luminosity must escape for complete relaxation of the constraints on m_l (*i.e.* $L_{TOT} \gtrsim 8 \times 10^{10}$ $L_\odot$, which implies $L_{ESC} \gtrsim 5 \times 10^{10}$ $L_\odot$).

The central hollow cylinder geometry proposed by Nakai *et al.* (1987) may provide a dust–free escape route, although this would further require that the stars responsible for L_{ESC} be situated away from the dusty molecular clouds in which they were born, or have since dispersed.

(ii) The dominant excitation mechanism of the far–IR and mm–wavelength line emission is photo–excitation not by stars but by the UV–precursors of fast shocks. The theoretical calculations by Shull & McKee (1979) are able to reproduce approximately the observed [OIII]λ52 and 88 μm and H53α relative strengths for shock velocities of $\sim$100 km s^{-1}. Exploring this model further, the Hβ line intensity from Shull & McKee ($\sim 3 \times 10^{-9}$ W m^{-2} sr^{-1}), together with Table 1, imply a source emitting area of $\sim$1000 kpc^2 if the entire measured H53α flux is produced by shocks of velocity 100 km s^{-1}. Further, if the shocks were due to identical supernova remnants with diameters of about 50 pc (*i.e.* $V_s \sim$ 100 km

s^{-1}), it would require the existence of $\sim 2 \times 10^5$ such remnants, with a mean separation of ~ 5 pc (assuming that they are equally distributed within a pillbox of 500 pc diameter and 150 pc thickness). This calculation assumes an ambient density for the ISM of 1 cm^{-3}. Moreover, if the time dependences of the star–formation and supernova rates are equal then the predicted SN rate is $\sim 0.1 - 1$ yr^{-1} (see Gehrz, Sramek & Weedman 1983), depending on the stellar mass range and assuming that each star with mass $m_* > 8$ $M_\odot$ produces a SN remnant. Thus the lifetime of a remnant in the far–IR line–emitting phase would have to be from 2×10^5 to 2×10^6 yr to match the required number of remnants. This seems implausibly long, particularly when one considers that only a relatively narrow range of remnant velocities ($\sim 80 - 120$ km s^{-1}, depending on ambient density) can reproduce the observed low–excitation conditions. Furthermore, the required number of line–emitting supernova remnants is many orders of magnitude larger than that actually observed using VLBI techniques (~40; Unger *et al.* 1984; Kronberg, Biermann & Schwab 1985). The supernova rate inferred from observations of the radio–emitting remnants (~0.2 yr^{-1}) is in agreement with that deduced above, using the ionization rate, only if the line emission arises primarily from ionized regions around stars.

A more detailed study of the line excitation in M82 is required to help examine the two possibilities described above; however we regard both explanations as unlikely due to the extreme conditions they imply.

Conclusions

The first detection of H53α emission from another galaxy has been made. At the frequency of this transition (43 GHz), spontaneous emission dominates and extinction by dust can be totally ignored. Thus, this detection provides a reliable measure of the photo–ionization rate ($N_{Lyc} = 1.1 \times 10^{54}$ s^{-1}) and electron temperature ($T_e \sim 5 \times 10^3$ K). Comparison of the H53α line strength with large aperture Brα observations shows there to be significant attenuation by dust even in the near–infrared ($A_{4\,\mu\mathrm{m}} \sim 1$ mag).

In combination with the observed IR luminosity and far–IR forbidden line measurements of O^{2+} and N^{2+}, the H53α emission has been used to constrain the allowed mass–range of stars present. This is found to be very narrow : approximately 20 – 30 $M_\odot$ depending on the electron density and elemental abundances. This restriction is lifted if a large fraction (~60%) of the stellar light escapes from the central regions without being absorbed by dust, or if significant amounts of the line radiation and luminosity are produced in fast shocks. Neither of these two situations is regarded as likely.

Acknowledgements

We are grateful to Tetsuo Hasegawa for assistance with tuning of the receivers and to the staff of NRO for their support and hospitality. We also thank Peter Storey for calculating the H40α and H53α recombination line strengths, Joss Bland for providing the Hα data, Pat Roche for stimulating discussions, Marjorie Fretwell for drawing the diagrams, and James Houck for improving the presentation of this work. PJP and TJTM acknowledge SERC

studentships for financial support.

Table 1

Theoretical recombination line ratios for a range of electron temperatures and densities assuming case B.

	$T_e = 5 \times 10^3$ K		$T_e = 10^4$ K	
	$n_e = 100cm^{-3}$	$n_e = 10^4cm^{-3}$	$n_e = 100cm^{-3}$	$n_e = 10^4cm^{-3}$
$I_{H53\alpha}/I_{H\beta}$	4.54×10^{-8}	5.16×10^{-8}	3.16×10^{-8}	3.38×10^{-8}
$I_{H40\alpha}/I_{H\beta}$	2.58×10^{-7}	2.28×10^{-7}	1.76×10^{-7}	1.59×10^{-7}
$N_{H53\alpha}/N_{Lyc}$	7.76×10^{-5}	8.82×10^{-5}	5.31×10^{-5}	5.68×10^{-5}
$I_{Br\alpha}/I_{H53\alpha}$	2.29×10^6	1.92×10^6	2.53×10^6	2.30×10^6
$I_{Br\gamma}/I_{H53\alpha}$	7.72×10^5	6.28×10^5	8.80×10^5	8.14×10^5

References

Aitken, D.K. & Roche, P.F., 1984. *M.N.R.A.S.*, **208**, 751.

Baker, J.G. & Menzel, D.H., 1938. *Ap. J.*, **88**, 52.

Bell, M.B. & Seaquist, E.R., 1978. *Ap. J.*, **223**, 378.

Bland, J. & Tully, R.B., 1988. *Nature*, **334**, 43.

Carlstrom, J.E., 1988. In *Galactic and Extragalactic Star Formation*, p. 571, eds. Pudritz, R.E. & Fich, M., Kluwer Acad. Publ., Dordrecht.

Duffy, P.B., Erickson, E.F., Haas, M.R. & Houck, J.R., 1987. *Ap. J.*, **315**, 68.

Gehrz, R.D., Sramek, R.A. & Weedman, D.W., 1983. *Ap. J.*, **267**, 551.

Gillett, F.C., Kleinmann, D.E., Wright, E.L. & Capps, R.W., 1975. *Ap. J.*, **198**, L65.

Houck, J.R., Shure, M.A., Gull, G.E. & Herter, T., 1984. *Ap. J.*, **287**, L11.

Hummer, D.G. & Storey, P.J., 1987. *M.N.R.A.S.*, **224**, 801.

Joy, M., Lester, D.F. & Harvey, P.M., 1987. *Ap. J.*, **319**, 314.

Jura, M., Hobbs, R.W. & Maran, S.P., 1978. *A. J.*, **83**, 153.

Klein, U., Wielebinski, R. & Morsi, H.W., 1988. *Astr. Ap.*, **190**, 41.

Kraus, J.D., 1986. *Radio Astronomy (2nd Ed.)*, Cygnus–Quasar Books, Ohio.

Kronberg, P.P., Biermann, P. & Schwab, F.R., 1985. *Ap. J.*, **291**, 693.

Kurucz, R.L., 1979. *Ap. J. Suppl.*, **40**, 1.

Landolt–Bornstein 1982. *Astronomy & Astrophysics*, **2b**, Stars & Star Clusters. Eds. Schiaffers, K. & Voigt, H.H.

Lugten, J.B., Watson, D.M., Crawford, M.K. & Genzel, R., 1986. *Ap. J.*, **311**, L51.

McCarthy, P.J., Heckman, T. & van Breugel, W., 1987. *A. J.*, **93**, 264.

Mathis, J.S., 1986. *Pub. A.S.P.*, **98**, 995.

Nakai, N., Hayashi, M., Handa, T., Sofue, Y. & Hasegawa, T., 1987. *Pub. A.S.J.*, **39**, 685.

Peimbert, M. & Spinrad, H., 1970. *Ap. J.*, **160**, 429.

Puxley, P.J., Hawarden, T.G. & Mountain, C.M., 1988a. To be submitted to *M.N.R.A.S.*.

Puxley, P.J., Hawarden, T.G. & Mountain, C.M., 1988b. *M.N.R.A.S.*, **234**, 29P.

Rieke, G.H. & Lebofsky, M.J., 1985. *Ap. J.*, **288**, 618.

Rieke, G.H., Lebofsky, M.J., Thompson, R.I., Low, F.J. & Tokunaga, A.T., 1980. *Ap. J.*, **238**, 24.

Rodriguez, L.F. & Chaisson, E.J., 1980. *Ap. J.*, **238**, 41.

Rubin, R.H., 1968. *Ap. J.*, **153**, 761.

Rubin, R.H., 1985. *Ap. J. Suppl.*, **57**, 349. R85.

Salpeter, E.E., 1955. *Ap. J.*, **121**, 161.

Scalo, J.M., 1986. *Fund. Cosm. Phys.*, **11**, 1.

Seaquist, E.R., Bell, M.B. & Bignell, R.C., 1985. *Ap. J.*, **294**, 546.

Shaver, P.A., Churchwell, E. & Walmsley, C.M., 1978. *Astr. Ap.*, **64**, 1.

Shull, J.M. & McKee, C.F., 1979. *Ap. J.*, **227**, 131.

Tammann, G.A. & Sandage, A., 1968. *Ap. J.*, **151**, 825.

Telesco, C.M., 1985. In: *Extragalactic Infrared Astronomy*, p87, Rutherford Appleton Laboratory. Ed. Gondhalekar, P.M.

Telesco, C.M., 1988. *Ann. Rev. Astr. Ap.*, **26**. 343.

Telesco, C.M. & Harper, D.A., 1980. *Ap. J.*, **235**, 392.

Unger, S.W., Pedlar, A., Axon, D.J., Wilkinson, P.N. & Appleton, P.N., 1984. *M.N.R.A.S.*, **211**, 783.

Weliachew, L., Fomalont, E.B., Greisen, E.W., 1984. *Astr. Ap.*, **137**, 335.

Willner, S.P., Soifer, B.T., Russell, R.W., Joyce, R.R. & Gillett, F.C., 1977. *Ap. J.*, **217**, L121.

Figure Captions

Figure 1.

(a) H53α and (b) H40α recombination–line spectra of the central star–forming complex in M82, together with (c) the [OIII]λ88 μm profile from Duffy *et al.* (1987) for comparison.

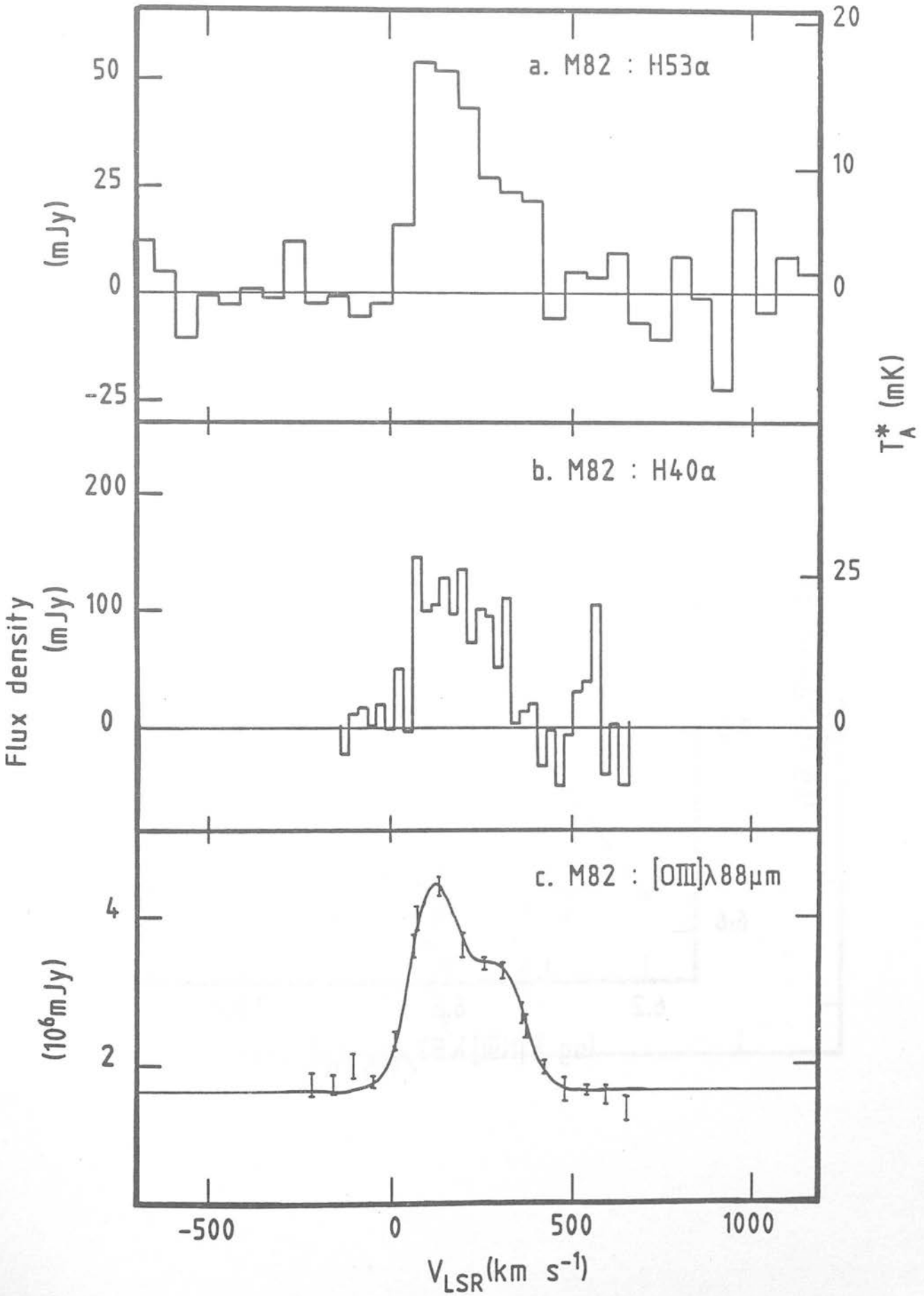
Figure 2.

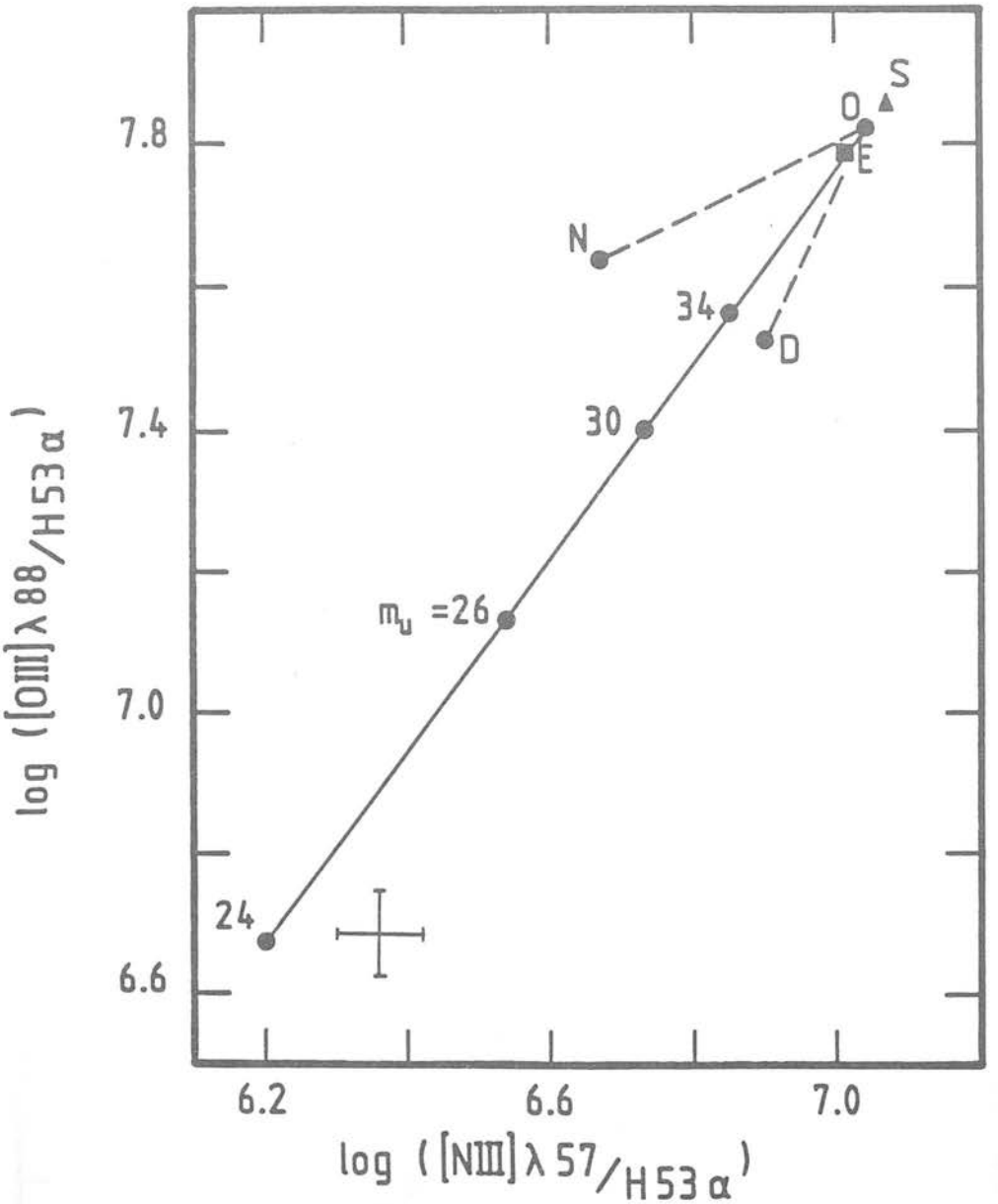
Theoretical and observed [OIII]λ88/H53α and [NIII]λ57/H53α line ratios for M82. The cross in the lower–left corner indicates the observed values and 1σ uncertainties. The points labelled O (unevolved starburst model with upper mass limit m_u=64 $M_\odot$, electron density n_e=100 cm^{-3} and solar abundances), N (as O but with lower elemental abundances), S (filled triangle; as O but with a flatter stellar mass function), D (as O but assuming a two–component density model), and E (filled square; as O but assuming constant evolution for $\sim 10^7$ yr) are described in detail in the text. The solid line represents the basic model (O) but adopting various values for the upper mass limit, which are marked along it.

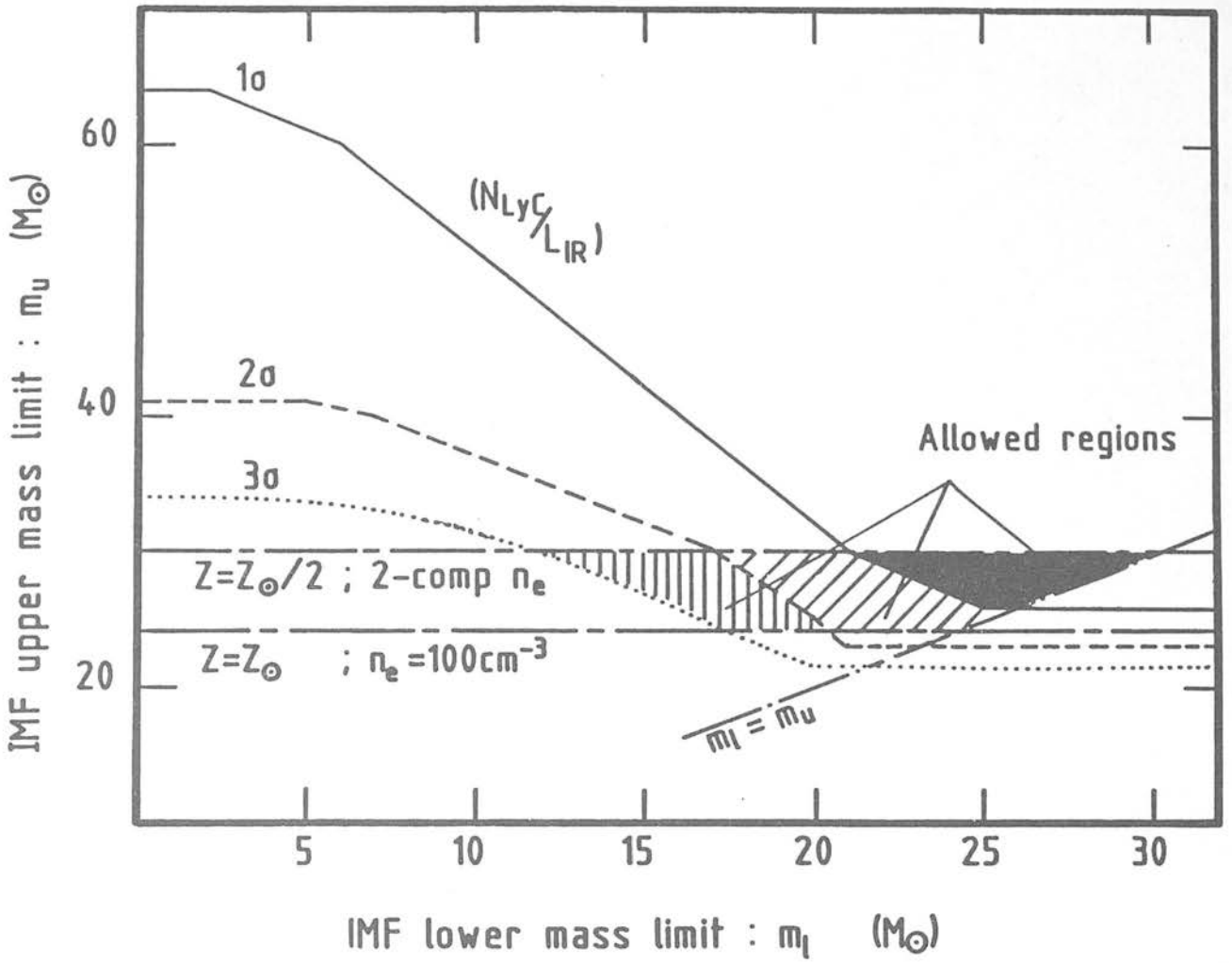
Figure 3.

Constraints on the lower (m_l) and upper (m_u) mass limits of the initial mass function in M82 identified using the models described in the text. The allowed ranges of m_l and m_u are indicated by the solid (1σ uncertainty), diagonally hatched (2σ), and vertically ruled (3σ) regions. Areas of the (m_l,m_u) parameter space above the short dash–long dash line labelled $Z=Z_\odot/2$ and below that labelled $Z=Z_\odot$ are forbidden by the line–excitation conditions. The region below the solid line (1σ uncertainty, short dashed line for 2σ, or dotted line

representing 3σ) is forbidden by the observed ratio of the ionization rate to IR luminosity. A line indicating equivalence of m_l and m_u is also shown.


a. M82 : H53α
b. M82 : H40α
c. M82 : [OIII]λ88μm
Flux density
(mJy)
(mJy)
(10⁶mJy)
T*A (mK)
50
25
0
-25
200
100
0
4
2
20
10
0
25
0
-500
0
500
1000
VLSR(km s⁻¹)


7.8
7.4
7.0
6.6
6.2
6.6
7.0
log ([OIII]λ88/H53α)
log ([NIII]λ57/H53α)
O
S
E
N
34
D
30
m_u =26
24


1σ
2σ
3σ
(N_{LyC}/L_{IR})
Allowed regions
$Z=Z_\odot/2$; 2-comp n_e
$Z=Z_\odot$; $n_e=100cm^{-3}$
$m_l = m_u$
60
40
20
5
10
15
20
25
30
IMF upper mass limit : m_u $(M_\odot)$
IMF lower mass limit : m_l $(M_\odot)$

Mon. Not. R. astr. Soc. (1990) 243, 97–102

A 450-μm continuum map of M82: comparison with the CO emission

P. A. Smith, P. W. J. L. Brand and P. J. Puxley*
University of Edinburgh, Department of Astronomy, Blackford Hill, Edinburgh EH9 3HJ

C. M. Mountain
Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ

N. Nakai
Nobeyama Radio Observatory, Nobeyama, Minamisaku Nagano 384-13, Japan

Accepted 1989 August 23. Received 1989 August 23; in original form 1989 April 13

SUMMARY

We have mapped the 450-μm dust continuum emission from the central 40×40 arcsec region of M82, and find that the total mass of gas in this region is $\simeq 2 \times 10^8$ $M_\odot$. The 450-μm map has one peak, in contrast to existing ^{12}CO maps which show two distinct peaks. The difference in structure between the two maps is probably due to the sensitivity of CO and dust emission to the physical and excitation conditions in molecular clouds. In particular, ^{12}CO and ^{13}CO observations suggest significant optical depth variations across the galaxy, and the intense UV field in the central 'starburst' region implies the existence of photodissociation regions from which the CO emission may originate. We conclude that the ^{12}CO and sub-millimetre emission cannot both be reliable tracers of molecular hydrogen in the extreme environments of starburst galaxies, and advise caution when using conversion factors determined in the Milky Way to find the H_2 mass in extragalactic molecular clouds.

1 INTRODUCTION

An important step towards an understanding of the global processes of star formation in galaxies is the ability to determine accurately the mass of available material in molecular clouds. Unfortunately the major constituent, H_2, is difficult to detect directly, and so it is necessary to observe a tracer to calculate the molecular hydrogen column density $N(H_2)$. The $J=1$–0 rotational line of ^{12}CO is commonly employed. Under the assumption that the ^{12}CO emission is optically thick, as is nearly always the case in galactic clouds (e.g. Lada 1985), and with excitation temperature $T_{ex} \simeq 10$ K, $N(H_2)$ follows from the ^{12}CO integrated intensity via $N(H_2)$ $(cm^{-2}) = 4 \times 10^{20}$ I_{CO} (K km s^{-1}) (Young & Scoville 1982). However, doubts have arisen over the constancy of this conversion factor, particularly in so-called 'starburst' galaxies where the presence of large numbers of hot young stars may invalidate the principal assumptions of optically thick ^{12}CO and $T_{ex} \simeq 10$ K. An alternative method of H_2 mass determination is to observe the sub-millimetre continuum emission due to re-radiation by dust grains, at characteristic temperatures ~20–50 K, of the stellar UV, and convert to H_2 mass via the ratio $N(H_2)/\tau(\nu)$ where $\tau(\nu)$ is the continuum optical depth. The method has been discussed in detail by Hildebrand (1983) and Gear (1988). The sub-millimetre emission is usually optically thin, and so all the dust is sampled.

To study the applicability of these two methods to extragalactic star-forming regions, we have mapped the sub-millimetre dust continuum of M82, a nearby galaxy already extensively observed in CO. M82 is a 'classic' starburst galaxy, the bulk of its energy output being in the IR, with a far-IR luminosity of 3×10^{10} $L_\odot$ (Telesco & Harper 1980). The multiwavelength observations are consistent with an intense episode of star formation in the central 1 kpc, which started some 5×10^7 yr ago (Rieke *et al.* 1980). Radio observations reveal numerous supernova remnants within 300 pc of the nucleus (Kronberg, Biermann & Schwab 1985). Hα filaments and X-ray emission extend several kpc perpendicular to the galactic plane (Lynds & Sandage 1963; Watson, Stanger & Griffiths 1984; Kronberg *et al.* 1985). CO and H I maps show a double-lobed structure, which is interpreted as a rotating molecular ring, 400 pc in diameter, enclosing the central starburst, with molecular and ionized gas being driven out of the galactic plane by the associated stellar winds and supernovae (Weliachew, Fomalont & Greisen 1984; Lo *et al.* 1987; Nakai *et al.* 1987). In contrast to galactic molecular clouds, ^{12}CO $J=2$–1/$J=1$–0 antenna temperature ratios and ^{12}CO $J=1$–0/^{13}CO $J=1$–0 integrated intensity ratios suggest that the CO in the central

*Present address: Institute for Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, Hawaii 96822, USA.

regions of M82 is optically thin, perhaps in the form of small (few pc) hot (40 K), $6 \times 10^3\ M_\odot$ clouds, quite unlike the giant molecular clouds (40 pc, 10 K, $5 \times 10^5\ M_\odot$) in the Milky Way (Knapp *et al.* 1980; Sofue 1988; Stark & Carlson 1982; Young & Scoville 1984; Nakai *et al.* 1987).

2 OBSERVATIONS

The central 40×40 arcsec of M82 was mapped at 450 μm in the dust continuum on the night of 1988 March 20, using the James Clerk Maxwell Telescope (JCMT) on Mauna Kea, Hawaii, and the common-user bolometer system UKT14 (filter width $\Delta\lambda = 54$ μm at 450 μm). The beamsize at 450 μm was 13 arcsec (not diffraction limited, owing to the lower dish quality and pointing precision then available on the recently commissioned JCMT). We mapped a fully sampled 7×7 grid, spaced at 6.5 arcsec and oriented along the major axis which lies at position angle PA = 65° (Nilson 1973). Using observations of the 450-μm peak at two different airmasses, the data were transformed out of the atmosphere with a zenith extinction coefficient $\tau(\mathrm{atm}) = 1.5$, taking account of the variation in airmass over the maps, which were then reduced and co-added using the continuum reduction software package NOD2 (Haslam 1974). The central pixel of the map was determined by peaking up on M82 at 800 μm. The 800-μm continuum peak is within 5 arcsec of the 2.2-μm peak (Hughes, Gear & Robson, in preparation), at RA(1950) = $09^h51^m43\overset{s}{.}47$, Dec(1950) = 69° 55′ 00″.3 (Pipher *et al.* 1987). Pointing was checked by means of five-point scans before and after each map. It was found that the second map was shifted by ≈ 7 arcsec with respect to the first, along the minor axis, and this offset was corrected before the maps were added. The estimated positional uncertainty of the map is therefore 7 arcsec. 450-μm photometry of M82 from the night of 1988 February 6 (Hughes *et al.*, in preparation) was used to scale up the integrated flux from our map. In a 13-arcsec beam centred at the peak, the 450-μm flux density is 15.5 ± 5 Jy.

3 RESULTS

The 450-μm map is presented in Fig. 1(a). Fig. 1(b), to the same scale, shows the ^{12}CO $J = 1$–0 interferometer map of Lo *et al.* (1987). It is immediately obvious that there is a morphological difference between the two maps. While in the CO map there are two distinct peaks, 25 arcsec apart, situated asymmetrically either side of the dynamical (and 2-μm) nucleus, our 450-μm map has only one peak, to the south-west of the nucleus, with an extension to the north-east. Before considering physical reasons for the apparent absence of a second peak, it is necessary to check the telescope pointing – could we have missed the second peak entirely? This seems unlikely. Both 100- and 40-μm profiles along the major axis (Joy, Lester & Harvey 1987) and a new 800-μm map (Hughes *et al.*, in preparation) show similar, single-peak structures. The centre of our map – the 800-μm peak – is within 5 arcsec of the 2.2-μm nucleus. The positional accuracy of our map is ≈ 7 arcsec, and there was good agreement between our two maps in the position on the major axis of the 450-μm peak. From physical considerations – at 450- and 800-μm we are approaching the Rayleigh–Jeans tail for characteristic dust temperatures

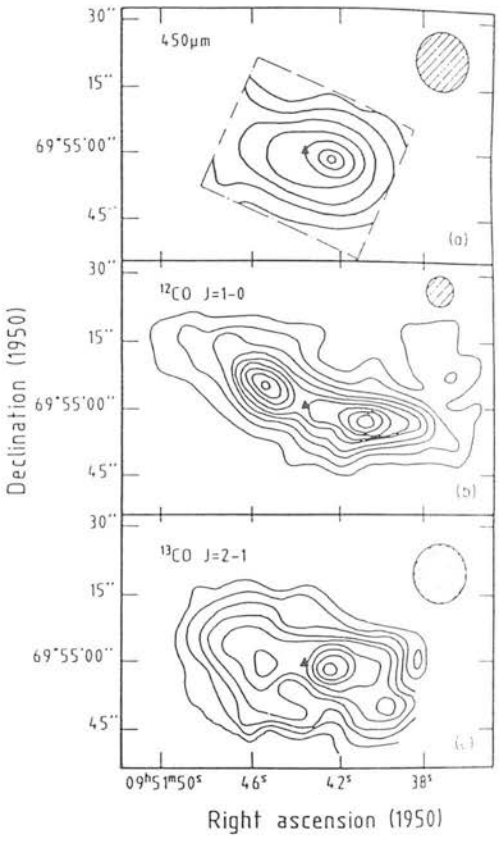


Figure 1. (a) A 450-μm continuum map of the central 40×40 arcsec of M82, with 13-arcsec resolution. The map is contoured with a spacing of 1 σ, from a 2 σ base level of 3.7 Jy beam^{-1}. The distribution of the dust continuum emission at 450 μm has one peak, with an extension to the north-east. The continuum peak is located within 7 arcsec of the 2.2-μm maximum at RA $09^h51^m43.47^s$, Dec. 69° 55′ 00.3″ (Pipher *et al.* 1987), which is also the dynamical nucleus and is marked by a triangle. The position angle of the dust major axis is orientated at ≈ 12° with respect to the optical major axis, which lies at PA = 65°. (b) A ^{12}CO $J = 1$–0 interferometer map of M82, beamsize 7 arcsec, taken from Lo *et al.* (1987). The CO map has two distinct peaks, with a weaker extension of the western lobe. The 2.2-μm nucleus (marked by a triangle) is located between the peaks, at a minimum of the CO emission. The galactic plane appears warped, in that the position angle of the CO ridge is not constant along the major axis (PA = 65°). (c) A ^{13}CO $J = 2$–1 map of M82 with 13 arcsec resolution, from Loiseau *et al.* (1988), shown to the same scale as Fig. 1(a and b) for comparison. The 2.2-μm nucleus is marked by a triangle. Note the large central peak, which dominates over the two peaks of the 'ring', indicating significant optical depth variations across the central regions of the galaxy.

(≈ 30–45 K, see below) – it seems reasonable to assume that the emission at the two wavelengths comes from similar material. For these reasons, we place the peak of our map within ≈ 7 arcsec of the 2.2-μm peak – i.e. interior to the CO ring. It is also unlikely that we are seeing two unresolved peaks. Our resolution is comparable to that of the IRAM ^{12}CO $J = 2$–1 map (13 arcsec at 230 GHz – Loiseau *et al.* 1990), and better than that of the ^{12}CO $J = 1$–0 Nobeyama map (16 arcsec – Nakai *et al.* 1987), in which two maxima are clearly seen. If there are two continuum peaks, we should expect to resolve them.

The position angle of the dust major axis is orientated at ≈ 12° to the optical major axis, as has also been observed in

the ^{12}CO $J=1-0$ maps of Olofsson & Rydbeck (1984) and Nakai *et al.* (1987).

Integrating up the emission from our calibrated map gives a total flux density of $F(450\ \mu m)=49\pm21$ Jy over our 40×40 arcsec grid, where the quoted error is due to uncertainty in the 450-μm photometry and in the relative locations of the 450-μm photometric aperture and map peak. Neither synchrotron (*cf.* Klein, Wielebinski & Morsi 1988) nor bremsstrahlung [$S(450\ \mu m)\sim0.5$ Jy, Puxley *et al.* 1989] processes contribute significantly to the integrated emission at this wavelength.

3.1 Gas mass

The physics of emission from dust grains has been discussed in detail by Hildebrand (1983), and the advantages of using sub-millimetre observations to derive gas masses convincingly argued by Thronson (1987). Gear (1988) points out that, rather than assuming a uniform source brightness distribution and top hat beam profile, as Hildebrand does, it is more realistic to assume Gaussian source and beam profiles, thus introducing a correction into Hildebrand's formulae. We will use Gear's equations in this paper.

The dust continuum optical depth is given by

$$\tau(\nu)=5.42\times10^{-16}[(\theta_s^2+\theta_b^2)/\theta_s^2\theta_b^2][S(\nu)/B(\nu,T)], \qquad (1)$$

where θ_s and θ_b are the FWHM (arcsec) of the source and beam profiles, respectively, $S(\nu)$ is the flux density into a beam centred on the peak of the source (Jy per beam), and $B(\nu,T)$ is the blackbody intensity at frequency ν and dust temperature T.

The determination of a reliable dust temperature depends on being able to fit the far-IR spectrum of the source to a modified blackbody function of the form $\nu^\beta B(\nu,T)$ (assuming optically thin emission), where β is the frequency dependence of the dust emissivity. The value of β is believed to increase from $\simeq1$ at $50\ \mu m\leq\lambda\leq250\ \mu m$ to $\beta\geq2$ by $\lambda\sim1000\ \mu m$ (Hildebrand 1983), but its exact variation with frequency is not well known. Clearly, the derived dust temperature depends on the value of β adopted. In general, β varying between 1 and 2 leads to an uncertainty of ~30 per cent in the dust temperature (Thronson *et al.* 1989). In addition, it is not normally possible to fit the whole far-IR spectrum to a single temperature – usually only the long wavelength end can be fitted, while there is an excess of emission at the shorter wavelengths indicative of hotter regions within the source (Gear 1988). It must be borne in mind that, for a composite object such as a galaxy, any derived temperature will only be an average over the system (Thronson *et al.* 1987).

Telesco & Harper (1980) and Klein *et al.* (1988) both fit a $\nu^{1.5}B(\nu,T)$ function to the far-IR spectrum of M82, for $\lambda\geq30$–$40\ \mu m$, deducing a dust temperature of 45 K. Similarly, Hughes *et al.* (1989) obtain a best-fit to the 40–1300 μm spectrum, including new data at 800 and 1100 μm, of $\beta=1.5$ and $T=47$ K, while Thronson *et al.* (1987) adopt a temperature of 30 ± 10 K from a $\beta=2$ fit to the far-IR data. We will use here a $\beta=1.5$ emissivity law, and a corresponding dust temperature of 47 K, from the work of Hughes *et al.*, although a fit to $\nu^{1.5}B(\nu,T)$ could be mimicked by an emissivity law varying smoothly from $\beta=1$–2 towards longer wavelengths (Gear *et al.* 1986). Our data are not of sufficient signal-to-noise to attempt an accurate temperature profile, although the similarity of our map to the 100-μm slice of Joy *et al.* (1987), and direct ratioing of the 100- and 40-μm profiles lead us to conclude that the temperature does not vary by more than about ±5 K over the central regions of the galaxy.

The hydrogen column density can be calculated from

$$N(H+H_2)=1.2\times10^{25}[750/\nu]^\beta\tau(\nu)\ \text{atoms cm}^{-2} \qquad (2)$$

(Hildebrand 1983), where ν is in GHz. The total mass of gas in the source is then

$$M_g=M(H+H_2)=2.7\times10^{-19}N(H+H_2)(\theta_sD)^2\ M_\odot, \qquad (3)$$

where D is the distance to the source in Mpc. Equation (3) includes a factor $\simeq1.36$ to allow for helium and other heavy elements (Hildebrand 1983). In the case of M82, molecular hydrogen dominates over atomic hydrogen in the central 1 kpc (e.g. Lugten *et al.* 1986) and so M_g closely approximates the mass of molecular hydrogen in the inner regions of the Galaxy.

The main sources of uncertainty in deriving the mass can now be clearly seen, namely, the assumed values of dust temperature and $N(H+H_2)/\tau(\nu)$. We have adopted the value of $N(H+H_2)/\tau(\nu)$ given by Hildebrand (1983), of 1.2×10^{25} atoms cm^{-2} at 400 μm, which is derived from the work of Whitcomb *et al.* (1981) on the far-IR properties of dust grains. The reliability of this number reflects the accuracy to which gas and dust parameters are known, which Thronson *et al.* (1987) estimate to be a factor of ~5 in the galactic disc. Thronson (1987) plots a number of theoretically and experimentally derived values of the ratio and finds it to be constant, within a factor of $\simeq3$, for a variety of galactic objects, although it may scale inversely with metallicity (Thronson *et al.* 1988).

The dust optical depth, hydrogen column density and total gas mass can now be calculated. Fitting a two-dimensional Gaussian profile to our data gives a map FWHM (beam convolved with source) of 28×16 arcsec. Then, using the 450-μm flux density of 15.5 ± 5 Jy in a 13-arcsec beam and taking $T=47$ K, equation (1) gives a value for the dust optical depth of 0.02 at 450 μm, consistent with our initial assumption that the sub-millimetre emission is optically thin. Adopting $\beta=1.5$ gives a hydrogen column density $N(H+H_2)=2.6\times10^{23}$ atoms cm^{-2} from equation (2). We take the distance to M82 to be 3.25 Mpc (Tammann & Sandage 1968), leading finally to a total gas mass $M_g=(2.0\pm0.8)\times10^8\ M_\odot$, where the quoted error reflects the uncertainties in the 450-μm flux density and fitted source size. Adoption of a ν^2 emissivity law and corresponding temperature of 30 K would increase the derived mass by a factor of 2.1 to $(4.2\pm1.8)\times10^8\ M_\odot$. Of this total gas mass, atomic hydrogen contributes only about $0.1\times10^8\ M_\odot$ in the central 1 kpc (Weliachew *et al.* 1984).

Our derived mass is consistent, within the accuracy of the technique, with previous estimates from sub-millimetre data. Thronson *et al.* (1988) calculate a mass of $3\times10^8\ M_\odot$ from a 1.3 mm flux of 0.9 Jy and an adopted dust temperature of 30 K. They propose that this mass is perhaps a factor of 3 too high, suggesting that the analysis of O'Connell & Mangano (1978) indicates a metallicity in M82 two to three times that in the solar neighbourhood. However, other work has suggested values of heavy element abundances which are close

to solar (Willner *et al.* 1977, Puxley *et al.* 1989). Jaffe, Becklin & Hildebrand (1984a) calculate a mass of 8×10^7 $M_\odot$ from a large-beam 400-μm flux of 30 Jy and dust temperature of 45 K. Olofsson & Rydbeck (1984) and Lo *et al.* (1987) both obtain an H_2 mass of 6×10^7 $M_\odot$ from ^{12}CO $J = 1$–0 data, assuming optically thin emission. Under the same assumption, Nakai *et al.* (1987) derive a mass of 1.1×10^8 $M_\odot$ in the central 1.5 arcmin from their ^{12}CO $J = 1$–0 map. Assumption of optically thick CO and application of a 'constant conversion factor' leads to a mass an order of magnitude higher than in the optically thin case (Nakai *et al.* 1987).

4 DISCUSSION

The difference in structure between the Co and sub-millimetre maps is surprising if, as is commonly assumed, the CO and dust continuum are both reliable tracers of the molecular hydrogen. As molecular hydrogen forms on dust grains, we can expect to find dust mixed with the H_2. On the other hand, the association of H_2 with CO is one of the cornerstones of the CO technique of mass determination. Our 450-μm map is similar in its morphology to both the 100- and 800-μm data, and intermediate-resolution radio observations, while the double-lobed structure of CO is repeated in H I, HCN and HCO$^+$ maps (Seaquist, Bell & Bignell 1985; Weliachew *et al.* 1984; Carlstrom 1988).

The peak of the 100-μm emission coincides with a region of particularly vigorous star formation ~ 10 arcsec south-west of the nucleus, displaying recombination lines from H II regions, OH masers, and a bright, non-thermal radio source associated with the supernova remnant 41.9 + 58 (Weliachew *et al.* 1984; Seaquist *et al.* 1985; Kronberg, Biermann & Schwab 1981, 1985; Joy *et al.* 1987). A weak ^{12}CO 1–0 feature is seen in the interferometer map between the nucleus and the western lobe, and an extension of the western peak of H I, as well as peaks of the HCN and ^{13}CO $J = 2$–1 emission are coincident with this area, leading to speculation on the presence of a 'giant molecular clump' in the vicinity of the star formation region (Weliachew *et al.* 1984; Lo *et al.* 1987; Carlstrom 1988; Loiseau *et al.* 1988).

A possible explanation of the differing structures then, is that the two lobes are present in dust but are 'swamped' by sub-millimetre emission from the giant star-forming region. The blackbody flux from a population of N spherical dust grains, radius a, at distance D, is given by

$$F(\nu) = N\pi(a/D)^2 Q(\nu) B(\nu, T) \quad (4)$$

(Hildebrand 1983), where $Q(\nu)$ is the emissivity of the grains at frequency ν. Thus flux density is proportional to the number of grains, which is equivalent to the amount of H_2 assuming a constant gas–dust ratio, and varies with temperature; for $\lambda \sim 450$ μm, $B(\nu, T)$ can be approximated by $B(\nu, T) \propto T^{1.6}$, to within an accuracy of 10 per cent, for $T = 20$–60 K (e.g. Jaffe *et al.* 1984b). An increase in temperature by approximately a factor of 1.5 from 47 K, or a two-fold increase in dust column density would be needed to 'fill in' the central depression suggested by the CO map of Lo *et al.* (1987). As stated before, we infer no temperature variations of this magnitude across M82, and since the emission is optically thin, we are seeing all the dust. We therefore believe it likely that the sub-millimetre emission from M82 is tracing predominantly column density variations across the central regions of the galaxy.

If the sub-millimetre emission can be assumed to be accurately following the column density of H_2 across M82, then it is apparent that some process must be enhancing the ^{12}CO emission in the lobes or depressing it in the interior regions. The CO emission from M82 is highly complex. Although the double-lobed structure apparent in ^{12}CO $J = 1$–0 and $J = 2$–1 maps (Lo *et al.* 1987; Nakai *et al.* 1987; Loiseau *et al.* 1989) has been widely interpreted as representing a rotating molecular ring, the velocity structure of the CO is disturbed, and, it has been argued, not consistent with a simply rotating disc or torus. The structural minor axis is not aligned with the axis of rotation, and the galactic plane appears warped (Lo *et al.* 1987).

The integrated intensity of CO emission, I_{CO}, is highly sensitive to changes in the physical and excitation conditions of molecular clouds, such as gas number density, metallicity, optical depth and excitation temperature, and such variations must be taken into account when deriving H_2 masses by this method.

(i) Gas number density, n: for optically thick ^{12}CO emission, Maloney & Black (1988) find that $I_{CO} \propto n^{1/2}$, thus application of a 'constant' conversion factor to extragalactic molecular clouds will not give a reliable indication of the H_2 mass if those clouds differ from the ($n = 200 cm^{-3}$) norm.

(ii) Heavy element abundance: I_{CO} increases with increasing metallicity, although, as stated earlier, the heavy element abundance in M82 may not differ significantly from solar neighbourhood values.

(iii) Optical depth: ^{12}CO $J = 1$–0 emission is almost always optically thick in galactic molecular clouds (e.g. Lada 1985). In the case of M82, however, large-beam observations give ^{12}CO $J = 2$–1/$J = 1$–0 antenna temperature ratios of ≈ 2–3.5, and ^{12}CO $J = 1$–0/^{13}CO $J = 1$–0 integrated intensity ratios of ≈ 20–30 suggest that the ^{12}CO in the central regions is at least partially optically thin (Knapp *et al.* 1980; Sofue 1988; Young & Scoville 1984; Stark & Carlson 1982). If the ^{12}CO emission was optically thin throughout the central region of M82, then we would be seeing all the gas, the ^{12}CO emission, like the dust, would be tracing molecular hydrogen column density, and the sub-millimetre and ^{12}CO maps should show similar structure. Clearly this is not the case. The recently published ^{13}CO $J = 2$–1 map (Fig. 1c) of Loiseau *et al.* (1988) shows a large central peak, coincident in position and velocity with the active south-western star-forming region, which dominates over the two peaks of the 'ring'. As ^{13}CO $J = 2$–1 emission is nearly always optically thin, the ^{12}CO $J = 2$–1/^{13}CO $J = 2$–1 antenna temperature ratio is a good tracer of changes in ^{12}CO optical depth. The ratio at the central peak of the ^{13}CO $J = 2$–1 map is 6.8, compared to 9–25 over the rest of the map, implying that the ^{12}CO optical depth varies across the central regions of M82 (Loiseau *et al.* 1988). We must therefore consider ^{12}CO optical depth variations as a likely cause of the differing CO and sub-millimetre structures, although other effects may also contribute.

(iv) Excitation temperature, T_{ex}: for optically thick CO emission, $I_{CO} \propto T_{ex}$ (Maloney & Black 1988), while in the optically thin case, for $h\nu \ll kT_{ex}$, $I_{CO} \propto 1/T_{ex}$ (Jaffe *et al.* 1984b). In addition, antenna temperature ratios in regions

with an excitation temperature gradient can be very different from those resulting from a region at uniform temperature (Cantó, Rodríguez & Anglada 1987). Knowledge of the gas excitation temperature and optical depth is therefore vital for correct interpretation of CO data and line ratios. One factor, which could have a marked effect on the gas temperature and hence the CO emission, is the presence of intense UV radiation fields. Recent H53α measurements of the central region of M82 imply an ionization rate of 1.1×10^{54} Lyman continuum photons per second (Puxley *et al.* 1989). Assuming that these photons are produced by B0 stars with an effective temperature = 31 000 K, this translates to $\simeq 5 \times 10^6$ B0 stars in the central region of M82 (Panagia 1973), with a resulting UV field $\geq 10^3$ times that in the solar neighbourhood (Maloney & Black 1988). In such intense UV fields, the CO emission may originate from the warm (100 K), dense (10^3 cm^{-3}) photodissociation regions at the boundary between H II regions and molecular clouds, which have been extensively modelled by Tielens & Hollenbach (1985). In these regions, illuminated either from within by embedded stars or from outside by galactic UV fields, grain photoelectric ejection is the dominant heating mechanism, and the gas temperature far exceeds the dust temperature. The dust emission, on the other hand, is largely independent of the UV field strength (Thronson 1988). The existence of photodissociation regions in M82 is implied by the detection of the 158-μm [C II] line which arises in such interface zones, and has the same spatial distribution and velocity structure as the ^{12}CO $J = 1$–0 line (Crawford *et al.* 1985). Maloney & Black (1988) suggest further that the anomalously high ^{12}CO $J = 2$–1/$J = 1$–0 antenna temperature ratios can be explained in terms of emission from photodissociation regions.

5 CONCLUSIONS

We have presented new high-resolution observations of M82 at 450 μm in the dust continuum. Our map of the central 40 × 40 arcsec shows distinct structural differences from ^{12}CO maps of the same region. Whereas the CO maps show a double-peaked structure believed to represent a 200 pc nuclear molecular ring, our 450-μm observations show one peak only. We believe that the difference is real, and not a result of insufficient resolution or poor pointing. We place the 450-μm peak within 7 arcsec of the 2.2-μm nucleus – interior to the CO 'ring'. The 450-μm emission is optically thin, and the integrated flux density over the map is 49 ± 21 Jy. Adoption of a dust temperature of 47 K results in a derived H_2 mass of 2.0×10^8 $M_\odot$, which increases by a factor of 2.1 if a lower dust temperature (30 K) is used. This is similar to previous estimates from sub-millimetre and CO observations, although assumption of optically thick ^{12}CO and corresponding 'constant conversion factor' leads to a somewhat higher derived mass.

The sub-millimetre and CO data cannot both be considered to be reliable tracers of molecular hydrogen. It is possible that a double-peaked dust structure is being 'swamped' by enhanced emission from a large star formation region south-west of the nucleus. However, due to the similarity of our 450-μm map to the 100-μm slice, and by direct comparison of the 100- and 40-μm profiles, we conclude that there are no large dust temperature gradients across M82; hence we expect the dust emission to trace column density variations only. We therefore consider it more likely that some effect is enhancing the CO emission in the lobes or depressing it in the inner regions. The CO emission in M82 is very complex and the integrated intensity is sensitive to changes in gas number density, metallicity, optical depth and excitation temperature. Although data on heavy element abundances suggest that the metallicity of M82 is close to solar, gas number density variations may render unreliable the standard $N(H_2)/I_{CO}$ conversion factor as applied to extragalactic molecular clouds. In addition, we consider it likely that a combination of optical depth variations and excitation temperature gradients are a major cause of the differing structures. Observations of the ^{13}CO $J = 2$–1 transition, a reliable tracer of the ^{12}CO optical depth, indicate significant variations of $\tau(^{12}CO)$ across the Galaxy. The intense UV flux in the central starburst region of M82 ($\geq 10^3 \times$ solar neighbourhood value) and the detection of [C II] 158-μm emission lead us to believe that the CO may originate in warm, dense photodissociation regions at the interface between H II regions and molecular clouds, where the gas temperature greatly exceeds the dust temperature.

We therefore believe that CO data must be treated with caution when applied to the extreme environments of vigorous extragalactic star formation regions, and that the effects on the CO emission of such parameters as optical depth and excitation temperature must be thoroughly investigated if CO lines are to be reliably used to trace the molecular gas content of extragalactic systems.

ACKNOWLEDGMENTS

We thank the staff of the Joint Astronomy Centre in Hilo for their technical support both during and after our observing run, and PATT for the allocation of JCMT telescope time. David Hughes kindly supplied us with the 800-μm peak position and 450-μm flux, and let us see the unpublished 800-μm map. We are grateful to Bill Dent and Walter Gear for helpful discussions and comments on the manuscript. Marjorie Fretwell drew the diagrams. PAS and PJP acknowledge SERC studentships.

The James Clerk Maxwell Telescope is operated by the Royal Observatory Edinburgh, on behalf of the Science and Engineering Research Council of the United Kingdom, the Netherlands Organization for Scientific Research and the National Research Council of Canada.

REFERENCES

Cantó, J., Rodríguez, L. F. & Anglada, G., 1987. *Astrophys. J.*, **321**, 877.

Carlstrom, J. E., 1988. *Galactic and Extragalactic Star Formation*, p. 571, eds Pudritz, R. E. & Fich, M., Kluwer, Dordrecht.

Crawford, M. K., Genzel, R., Townes, C. H. & Watson, D. M., 1985. *Astrophys. J.*, **291**, 755.

Gear, W. K., 1988. *Millimetre and Submillimetre Astronomy*, p. 307, eds Wolstencroft, R. D. & Burton, W. B., Kluwer, Dordrecht.

Gear, W. K., Gee, G., Robson, E. I., Ade, P. A. R. & Duncan, W. D., 1986. *Mon. Not. R. astr. Soc.*, **219**, 835.

Haslam, C. G. T., 1974. *Astr. Astrophys. Suppl.*, **15**, 333.

Hildebrand, R. H., 1983. *Q. Jl R. astr. Soc.*, **24**, 267.

Jaffe, D. T., Becklin, E. E. & Hildebrand, R. H., 1984a. *Astrophys. J.*, **285**, L31.

Jaffe, D. T., Hildebrand, R. H., Keene, J., Harper, D. A., Loewenstein, R. F. & Moran, J. M., 1984b. *Astrophys. J.*, **281**, 225.
Joy, M., Lester, D. F. & Harvey, P. M., 1987. *Astrophys. J.*, **319**, 314.
Klein, U., Wielebinski, R. & Morsi, H. W., 1988. *Astr. Astrophys.*, **190**, 41.
Knapp, G. R., Phillips, T. G., Huggins, P. J., Leighton, R. B. & Wannier, P. G., 1980. *Astrophys. J.*, **240**, 60.
Kronberg, P. P., Biermann, P. & Schwab, F. R., 1981. *Astrophys. J.*, **246**, 751.
Kronberg, P. P., Biermann, P. & Schwab, F. R., 1985. *Astrophys. J.*, **291**, 693.
Lada, C. J., 1985. *Ann. Rev. Astr. Astrophys.*, **23**, 267.
Lo, K. Y., Cheung, K. W., Masson, C. R., Phillips, T. G., Scott, S. L. & Woody, D. P., 1987. *Astrophys. J.*, **312**, 574.
Loiseau, N., Reuter, H.-P., Wielebinski, R. & Klein, U., 1988. *Astr. Astrophys.*, **200**, L1.
Loiseau, N., Nakai, N., Sofue, Y., Wielebinski, R., Reuter, H.-P. & Klein, U., 1990. *Astr. Astrophys.*, submitted.
Lugten, J. B., Watson, D. M., Crawford, M. K. & Genzel, R., 1986. *Astrophys. J.*, **311**, L51.
Lynds, C. R. & Sandage, A. R., 1963. *Astrophys. J.*, **137**, 1005.
Maloney, P. & Black, J. H., 1988. *Astrophys. J.*, **325**, 38.
Nakai, N., Hayashi, M., Handa, T., Sofue, Y., Hasegawa, T. & Sasaki, M., 1987. *Publs astr. Soc. Japan*, **39**, 685.
Nilson, P., 1973. *Uppsala astr. Obs. Ann.*, **6**.
O'Connell, R. W. & Mangano, J. J., 1978. *Astrophys. J.*, **221**, 62.
Olofsson, H. & Rydbeck, G., 1984. *Astr. Astrophys.*, **136**, 17.
Panagia, N., 1973. *Astr. J.*, **78**, 929.
Pipher, J. L., Moneti, A., Forrest, W. J., Woodward, C. E. & Shure, M. A., 1987. In: *Infrared Astronomy with Arrays*, p. 326, eds Wynn-Williams, C. G. & Becklin, E. E., Univ. Hawaii, Honolulu.
Puxley, P. J., Brand, P. W. J. L., Moore, T. J. T., Mountain, C. M., Nakai, N. & Yamashita, T., 1989. *Astrophys. J.*, **345**, 163.
Rieke, G. H., Lebofsky, M. J., Thompson, R. I., Low, F. J. & Tokunaga, A. T., 1980. *Astrophys. J.*, **238**, 24.
Seaquist, E. R., Bell, M. B. & Bignell, R. C., 1985. *Astrophys. J.*, **294**, 546.
Sofue, Y., 1988. *Galactic and Extragalactic Star Formation*, p. 409, Whistler, eds Pudritz, R. E. & Fich, M., Kluwer, Dordrecht.
Stark, A. A. & Carlson, E. R., 1982. *Astrophys. J.*, **279**, 122.
Tammann, G. A. & Sandage, A. R., 1968. *Astrophys. J.*, **151**, 825.
Telesco, C. M. & Harper, D. A., 1980. *Astrophys. J.*, **235**, 392.
Thronson, H. A., 1988. *Galactic and Extragalactic Star Formation*, p. 621, eds Pudritz, R. E. & Fich, M., Kluwer, Dordrecht.
Thronson, H. A., Walker, C. K., Walker, C. E. & Maloney, P., 1987. *Astrophys. J.*, **318**, 645.
Thronson, H. A., Walker, C. K., Walker, C. E. & Maloney, P., 1989. *Astr. Astrophys.*, **214**, 29.
Tielens, A. G. G. M. & Hollenbach, D., 1985. *Astrophys. J.*, 291, 722.
Watson, M. G., Stanger, V. & Griffiths, R. E., 1984. *Astrophys. J.*, **286**, 144.
Weliachew, L., Fomalont, E. B. & Greisen, E. W., 1984. *Astr. Astrophys.*, **137**, 335.
Whitcomb, S. E., Gatley, I., Hildebrand, R. H., Keene, J., Sellgren, K. & Werner, M. W., 1981. *Astrophys. J.*, **246**, 416.
Willner, S. P., Soifer, B. T., Russell, R. W., Joyce, R. R. & Gillett, C., 1977. *Astrophys. J.*, **217**, L121.
Young, J. S. & Scoville, N. Z., 1982. *Astrophys. J.*, **258**, 467.
Young, J. S. & Scoville, N. Z., 1984. *Astrophys. J.*, **287**, 153.

Observations of H53α Emission from Bright Galaxies

P.J. Puxley[1], P.W.J.L. Brand[2], T.J.T. Moore[3], C.M. Mountain[4], and N. Nakai[5]

[1]Institute for Astronomy, 2680 Woodlawn Drive, Honolulu, Hawaii 96822.

[2]Dept. of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ.

[3]Dept. of Physics, Queen Mary College, University of London, London E1 4NS.

[4]Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ.

[5]Nobeyama Radio Observatory, Nobeyama, Minamisaku, Nagano 384-13, Japan.

Summary

We have searched for the hydrogen recombination lines H40α and H53α in three bright spiral galaxies, and detected H53α in NGC 2146 and, possibly, in NGC 1068. The strength of the detected lines, when compared with radio continuum and far infrared data, implies that they cannot be produced solely by spontaneous emission but that about three quarters, or more, of the line photons are stimulated by continuum radiation. Model HII regions are presented which satisfy constraints imposed by the large stimulated H53α fraction, and by the radio continuum and line fluxes. These HII regions must be small, dense ($n_e \sim 10^5$ cm^{-3}) and numerous. These results contrast with our suggestion that the H53α radiation from M82 is dominated by spontaneous emission: a conclusion which is still supported by the observed H40α/H53α line ratio.

1 Introduction

Recently, the millimetre–wavelength hydrogen recombination lines H53α and H40α were detected in the nearby 'starburst' galaxy M82 (Puxley *et al.* 1989: hereinafter Paper I). Analysis of longer–wavelength radio recombination lines (RRLs), such as H102α, demonstrates that they are produced primarily by stimulated emission and that their intensity decreases with rising frequency (lower principal quantum number) due to the falling spectrum of the stimulating non–thermal background radiation and decreasing continuum optical depth (*e.g.* Bell & Seaquist 1978; Shaver, Churchwell & Walmsley 1978). The millimetre–wavelength lines in M82 go against this trend, being more than an order of magnitude brighter than would be expected based on the RRL strengths, leading to the conclusion that they are caused by spontaneous emission. Comparison with the near–infrared Brα line implies an extinction of about 1 mag at 4 μm, equivalent to $A_V \simeq 27$ mag, which is considerably larger than that inferred solely from near–IR or optical lines. This behaviour of increasing extinction with wavelength of the diagnostic lines is expected if the sources of radiation, HII regions in this case, and dust are distributed along the line of sight. This is because each line of wavelength λ contains contributions only from those sources lying within a few τ_λ of the exterior of the region, and so different lines sample different numbers of sources along the line of sight. Since the mm–wavelength lines should be unattenuated even by the largest amounts of extinction expected to be encountered, they appear to provide a most reliable measure of the ionization rate.

In this paper we describe observations of H53α and H40α in three other bright galaxies : NGC 1068, NGC 2146 and M83 (§2). The results are compared with other data in §3 where we find that there are good reasons to question a spontaneous origin for much of the H53α line emission in NGC 1068 and NGC 2146, and discuss the implications and experimental tests in §4.

2 Observations and Calibration

Observations of H53α (rest frequency, $\nu = 42.95197$ GHz) and H40α ($\nu = 99.02297$ GHz) were made in January 1989 using the 45m telescope at the Nobeyama Radio Observatory. Both lines were observed simultaneously with SIS (H53α) and Schottky diode (H40α) front ends tuned to near the galaxian systemic velocity. The 2048–channel wide–band acousto–optical spectrometer back end gave a (2 channel) frequency resolution of 250 kHz and an instantaneous bandwith of 256 MHz. During the observations, the single sideband system noise temperatures were approximately 300 and 600 K, at 43 and 99 GHz respectively. Telescope pointing was monitored approximately every 90 minutes by means of nearby SiO maser sources, yielding a pointing and tracking accuracy of $\pm 5''$ over this period. This is small compared with the beam sizes of $41''$ (43 GHz) and $17''$ (99 GHz). The spectra were calibrated internally using a warm chopper in front of the receiver and the resulting antenna temperatures were converted into flux

densities using telescope aperture efficiencies of $\eta_A = 0.48$ and 0.26, at 43 and 99 GHz respectively.

To assist calibration of the galaxy spectra, the planetary nebula NGC 6210 (PK 43+37°1) was observed in the H53α line using the same instrumental set-up. A portion of this spectrum near zero LSR velocity and re-binned into channels of width 4.4 km s^{-1} is shown in Fig. 1a. The strongest feature in this spectrum, at $V_{LSR} \simeq -20$ km s^{-1}, is identified as H53α (*c.f.* the weighted mean of independent velocity measurements of $V_{LSR} = -17.7 \pm 1.1$ km s^{-1}: Schneider *et al.* 1983). We also note, in passing, that the width of this line (FWHM=21 km s^{-1}) is in excellent agreement with that derived from optical data (FWHM= 21 ± 2 km s^{-1}; Wilson 1950). Adopting a diameter for NGC 6210 of $15\widehat{''}$ from radio continuum (Milne & Aller 1975), 10 μm (Williams *et al.* 1988), and Hα data (Balick 1987) implies a conversion factor from the T_A^* (K) scale to flux density (Jy) of :

$$\frac{2k}{A_e}\frac{\theta_r^2}{\theta_b^2} = 4.10 \qquad (1)$$

where A_e is the effective area of the telescope (equal to $\eta_A \times$ geometric area), θ_b is the beamwidth ($41\widehat{''}$ at 43 GHz) and θ_r is the response width, given by $\theta_r^2 = \theta_s^2 + \theta_b^2$ for a source of width θ_s (*e.g.* Downes 1989). Integration over the velocity range -60 to 15 km s^{-1} then gives an H53α flux of $(4.6 \pm 0.7) \times 10^{-21}$ Wm^{-2}, where the error derives from the rms noise in the spectrum.

For comparison purposes, observations of NGC 6210 were also made in the near-infrared hydrogen recombination lines Brγ ($\lambda = 2.166\,\mu$m) and Paβ ($\lambda = 1.282\,\mu$m) using the UK Infrared Telescope (UKIRT). The apertures employed were $20\widehat{''}$ (Brγ) and $5\widehat{''}$ (Paβ). These data, calibrated using the standard stars BS 4695 and BS 5384, are shown in Figs. 1b and 1c. The integrated fluxes in these lines are compared with that in H53α, and with the 5 GHz continuum measurement by Milne & Aller (1975), in Table 1. The low emission measure in NGC 6210 ($EM = 2 \times 10^6$ cm^{-6} pc; Walmsley *et al.* 1981) means that any H53α contribution by stimulated emission will be negligible. The final column of this Table shows the ionization rate derived from each measurement using Case B recombination data from Hummer & Storey (1987) assuming an electron density of 10^4 cm^{-3} and temperature of 10^4 K (Kaler 1986; Stanghellini & Kaler 1989). Since the ionization rate derived from H53α is consistent with that calculated from the other observations, we conclude that no correction need be applied to the chopper-calibrated galaxy spectra.

3 Results

In Table 2, the observed H53α and H40α fluxes or upper limits are presented for the three galaxies observed. The only positive detection was of H53α in NGC 2146, with a marginal detection in NGC 1068. The columns in Table 2 are : (1) galaxy name; (2) distance in Mpc (Tully 1988); (3) assumed source extent; (4) integrated flux in H53α line; (5) integrated flux in H40α line; and (6) ionization rate. The entries in this Table for each galaxy are discussed below, but first we make some general comments.

The source size (column 3), which is used in calculation of the factor $((\theta_b^2+\theta_s^2)/\theta_b^2)$ to convert antenna temperature to flux density (see §2), is only approximate. At 43 GHz, the beam ($41\hat{''}$) is expected to be considerably larger than the sources and therefore the conversion factors are small. Hence our uncertainty in the source size should not introduce significant errors into the derived H53α fluxes. However, at 99 GHz the beam ($17\hat{''}$) and sources are of comparable size requiring large conversion factors and thus the H40α fluxes are likely to be less reliable.

For the detected galaxies, the error in the H53α flux (column 4) is dominated by the rms noise in the continuum. The formal significance of the detection was calculated by dividing the total signal in the line by the rms noise per channel multiplied by the square root of the number of channels containing line flux. The assigned error comprises this source of uncertainty added in quadrature to an estimated 10% calibration error deriving primarily from the aperture efficiency. For the non–detections we quote 3σ upper limits determined in similar fashion with linewidths assumed equal to that of H53α (NGC 1068 and 2146) or to CO (M83; Combes *et al.* 1978).

The ionization rate in column 6 was calculated by assuming that all of the H53α flux is generated via spontaneous emission following recombination in gas with an electron temperature of $T_e = 5 \times 10^3$ K and density of $n_e = 10^4$ cm^{-3}. Under such conditions, one H53α photon is produced for every 1.13×10^4 recombinations (Paper I). The derived ionization rates are modestly sensitive to the electron temperature and less so to the electron density : if $T_e = 10^4$ K then N_{Lyc} must be increased by a factor of 1.55, whereas if $T_e = 5 \times 10^3$ K and $n_e = 100$ cm^{-3} then the multiplier is only 1.14. No correction for absorption of ionizing photons by dust has been made.

3.1 NGC 1068

The millimetre–wavelength recombination lines were observed in NGC 1068 for a total (on plus off source) integration time of 6.3 hr, with receivers tuned so that the central channel in the AOS corresponded to $V_{LSR} = 1131$ km s^{-1}. After removal of linear baselines, the data were re–binned into channels of width 8 MHz and the resulting H53α spectrum is shown in Fig. 2a. Integrating under the feature between velocities of 950 and 1250 km s^{-1}, we find a marginal detection at the 2.9σ level assuming a Gaussian noise distribution. This velocity range is comparable to that of optical line–emitting gas seen in Fabry–Perot studies of this region (*e.g.* Atherton *et al.* 1985; Cecil, Bland & Tully 1990). We would not expect to see broad line emission from high–velocity gas near the active nucleus, despite the potential of this line to penetrate large quantities of obscuring material, because of the relatively small receiver bandwidth and narrow intrinsic resolution ($\Delta v \sim 0.87$ km s^{-1}) which results in poor sensitivity to broad lines. We conclude that the emission probably arises from gas located in the central star–forming disk and ionized by a combination of stellar UV and power–law radiation from the Seyfert nucleus (*e.g.* Baldwin, Wilson & Whittle 1987). For this reason we set the source size to be $25\hat{''}$ (Telesco & Decher 1988; Cecil *et al.*

1990), and thereby derive a peak H53α line flux density of (36.7 ± 12.7) mJy and an integrated flux of $(9.5 \pm 3.3) \times 10^{-21}$ Wm^{-2}.

The H40α transition was not detected in NGC 1068. For a linewidth of 300 km s^{-1} and source size of $25''$, the rms noise implies a 3σ upper limit to the integrated flux of 1.0×10^{-19} Wm^{-2}. For spontaneous emission by gas having $n_e = 100 - 10^4$ cm^{-3} and $T_e = (5 - 10) \times 10^3$ K we would expect the line ratio $F_{H40\alpha}/F_{H53\alpha}$ to lie in the range $4.4 - 5.7$ (Paper I). On the other hand if the lines were generated by stimulated emission, the ratio would be a factor of 2 to 3 lower (see §4). The observed upper limit $F_{H40\alpha}/F_{H53\alpha} < 10.5$ is thus too large to shed light on the H53α production mechanism.

The H53α measurement may be compared with the Hα flux in a $40''$ beam centred on the nucleus of 1.73×10^{-14} Wm^{-2} (J. Bland, personal communication, data in Cecil *et al.* 1990). The latter value also excludes the AGN because of the difficulty in separating broad Hα and [NII] lines from the underlying continuum within the central $10''$. If both Hα and H53α are the result of spontaneous emission then their flux ratio indicates a beam-averaged extinction of $A_V = 4.6 - 5.2$ mag, depending on n_e and T_e and using the interstellar reddening law of Rieke & Lebofsky (1985), with an estimated 1σ uncertainty due to measurement errors of $\Delta A_V = 0.4$ mag. This is somewhat larger than the moderate reddening found in most visible studies ($A_V = 1.5 - 3$ mag), but could be interpreted either as evidence for a stimulated contribution to the H53α flux or for optically obscured HII regions.

The thermal radio continuum emission can be predicted from H53α data using (Paper I) :

$$\left(\frac{S_\nu^{th}}{\mathrm{Jy}}\right) = 5.6 \times 10^{15} \left(\frac{F_{H53\alpha}}{\mathrm{Wm^{-2}}}\right) \left(\frac{N_{Lyc}}{N_{H53\alpha}}\right) \left(\frac{\nu}{\mathrm{GHz}}\right)^{-0.1} \left(\frac{T_e}{10^4\mathrm{K}}\right)^{0.45} \qquad (2)$$

where S_ν^{th} is the bremsstrahlung flux density at a frequency ν, $N_{Lyc}/N_{H53\alpha}$ is the number of recombinations per H53α photon (values given in Paper I), and T_e is the electron temperature. Comparison with observations is complicated in this case by the presence of several components. Wilson & Ulvestad (1982) separate the 1.4 GHz emission into contributions from a central source, a linear structure, and larger scale emission extending over about $2'$, estimating the latter to be 1.0 ± 0.3 Jy. Wynn-Williams, Becklin & Scoville (1985) determine the spectral index of the emission to be in the range $\alpha = -0.8$ to -1.0. For an electron temperature of $T_e = 5 \times 10^3$ K, we predict a bremsstrahlung flux density at 1.4 GHz of 0.45 Jy, with a weak dependence on the electron density. This value rises to 0.95 Jy if $T_e = 10^4$ K. Large thermal flux densities such as these are clearly inconsistent with the observed non-thermal spectrum, assuming that the HII regions are optically thin at this frequency, and imply that the observed H53α emission cannot be entirely spontaneous. If the HII regions had sufficiently large emission measure that they were optically thick, then their thermal continuum would indeed be reduced but stimulated line emission would be enhanced.

3.2 NGC 2146

In NGC 2146 the H53α and H40α lines were observed for a total integration time of 5.3 hr, with receivers tuned to V_{LSR} = 883 km s^{-1}. The calibrated H53α spectrum is shown in Fig. 2b after re–binning into 8 Mhz–wide channels. The profile of this line resembles the CO J=1–0 spectrum in a 45$\widehat{''}$ beam by Young *et al.* (1988). At higher resolution, the CO emission delineates a narrow structure several kpc long (Jackson & Ho 1988; Young *et al.* 1988a), morphologically similar to the Hα and radio continuum emission (Kronberg & Biermann 1981; Young, Kleinmann & Allen 1988), suggesting that the H53α emission may also be extended.

Integrating the spectra in Fig. 2b between 650 and 1050 km s^{-1} indicates that the H53α line has been detected at the 6.6σ level. The H40α line was not detected. Taking a characteristic source extent of 14$\widehat{''}$ from CO, Hα and radio continuum data, we calculate a peak H53α line flux density of (28.2 $\pm$ 7.9) mJy and integrated flux of (9.5 $\pm$ 1.7) $\times$ 10^{-21} Wm^{-2}. The 3σ upper limit to the H40α strength of 7.1 $\times$ 10^{-20} Wm^{-2} implies a line flux ratio $F_{H40\alpha}/F_{H53\alpha} < 7.5$. As in the case of NGC 1068, this upper limit is too large to determine whether stimulated emission is important at these high frequencies.

In NGC 2146 the central complex emits about 0.3 Jy at 5 GHz in the continuum with a spectral index of $\alpha = -0.5$ (Kronberg & Biermann 1981). Once again, using Eqn. 2 we find that the bremsstrahlung flux density predicted from the observed H53α flux is larger than the observed value, approximately 0.4–0.8 Jy for $T_e = (5-10) \times 10^3$ K. In this instance, there is additional evidence which supports the conclusion that not all of the H53α emission can be spontaneous. The ionization rate given in column 6 of Table 2, and derived assuming purely spontaneous emission, is extremely large compared with the total luminosity. We define the latter as being the sum of IR (1–500 μm) and optical luminosities. The IR luminosity was determined from mid– and far–IR photometry through a 50$\widehat{''}$ aperture by Rickard & Harvey (1984) to be $L_{IR} = 8 \times 10^{10}$ L$_\odot$. To account for the emission at visible wavelengths we define the optical flux as λf_λ, where f_λ is the flux density through a standard B filter. (In this form the sun has an optical luminosity of 0.62 L$_\odot$; see discussion in Thronson & Greenhouse 1988). Using B photometry in a 40$\widehat{''}$ aperture by de Vaucouleurs *et al.* (1978) we derive the optical contribution to the luminosity of $L_{OPT} = 4 \times 10^9$ L$_\odot$. Thus the ratio of ionization rate to total luminosity is $N_{Lyc}/L_{tot} = (1.75 \pm 0.54) \times 10^{44}$ s^{-1} L$_\odot{}^{-1}$, larger even than that of a 100 M$_\odot$ star for which $N_{Lyc}/L_{tot} \simeq 6 \times 10^{43}$ s^{-1} L$_\odot{}^{-1}$ (*e.g.* Puxley 1988, using stellar atmospheres from Kurucz 1979). We defer a detailed discussion of the enhanced H53α line emission, over and above that expected from the observed luminosity for a star–forming complex, until §4.

3.3 M83

The observations of M83 were obtained using receivers tuned to V_{LSR} = 520 km s^{-1} with a total integration time of 2.7 hr. Neither recombination line was detected. Adopting a source size of 15$\widehat{''}$ from the extent of near–IR recombination line emission (Puxley & Brand 1990) and a linewidth of 100 km s^{-1}

(Combes *et al.* 1978), yields 3σ upper limits to the line fluxes of $F_{H53\alpha} < 3.7 \times 10^{-21}$ Wm^{-2} and $F_{H40\alpha} < 7.1 \times 10^{-20}$ Wm^{-2}. The former value is somewhat uncertain because of significant baseline curvature. The H53α upper limit can be compared with the integrated Brγ flux mapped over a $20'' \times 20''$ region using a $5''$ beam of 1.9×10^{-16} Wm^{-2} (Puxley & Brand 1990), giving $F_{H53\alpha}/F_{Br\gamma} < 1.9 \times 10^{-5}$. For Case B recombination we expect $F_{H53\alpha}/F_{Br\gamma} = (0.11 - 0.16) \times 10^{-5}$ depending on the electron density and temperature (Paper I), and thus derive an upper limit to the $2\,\mu$m extinction of 2.7 mag (equivalently, $A_V = 23$ mag). This is comparable with the extinction found by Turner, Ho & Beck (1987) from Brγ and Brα lines.

4 Discussion : H53α line emission from NGC 2146

Comparison of the observed radio continuum emission with the thermal flux density predicted from H53α data has indicated that *not all of the H53α flux from NGC 1068 and NGC 2146 can be due to spontaneous emission* (§3). Due to the marginal quality of the H53α detection in NGC 1068 and complications arising from the presence of an AGN we shall restrict our further examination of mm-wavelength line emission to the case of NGC 2146. There are at least three potential sources of H53α photons, in addition to spontaneous emission from HII regions around massive stars, which we now consider. These are (i) fast shocks, (ii) an active nucleus, and (iii) stimulated emission.

The possibility of a major contribution to the H53α line emission from fast shocks may be rejected because the expected intensity is low and thus an unreasonably large shock surface area would be required. Extrapolating the models in Shull & McKee (1979) to a shock of velocity 130 km s^{-1} ploughing into gas of density 100 cm^{-3}, we might expect an intensity in the Hβ line of 6×10^{-8} J m^{-2} s^{-1} sr^{-1} and thus $I_{H53\alpha} \sim 3 \times 10^{-15}$ J m^{-2} s^{-1} sr^{-1}, using line ratios from Paper I. Therefore, matching the observed flux of 9.5×10^{-21} Wm^{-2} would require an emitting area of about 1000 kpc^2.

Line emission from an active nucleus may be discounted on several grounds. Firstly, empirically AGN have smaller ratios of ionization rate to total luminosity than do the star-forming regions in normal galaxies (*e.g.* DePoy 1987). This is in the opposite sense to the observed ratio (§3.2). Secondly, the similarity of the H53α and CO line profiles has been invoked to argue that, like CO, the H53α emitting gas must be extended and not restricted to a small central condensation (§3.2). In fact, the observed H53α line width is consistent with the Hα and CO rotation curves presented by Jackson & Ho (1988) if the extent of the H53α emission approximates that delineated by radio continuum and Hα emission. Finally, the presence of an AGN does not receive support from optical spectroscopy which indicates gas excitation consistent with a heavily reddened HII region (Benvenuti *et al.* 1975; Keel 1984).

Thus, despite initially being excluded by analogy with M82, we are forced to reconsider stimulated emission as potentially a major source of H53α photons in NGC 2146. As the first step we attempt to quantify the discrepancy by estimating the bremsstrahlung contribution to the observed radio continuum

emission and thereby derive the spontaneous and stimulated H53α fluxes. Let the total radio continuum emission be the sum of two components : optically thin HII regions with spectral index $\alpha = -0.1$ and synchrotron emission with $\alpha = -0.75$. Given that the observed spectral index between 1.4 and 5 GHz is $\alpha = -0.5$ and that $S_{5\rm GHz} = 0.3$ Jy, we solve for the thermal emission to find $S^{th}_{5\rm GHz} = 0.14$ Jy. This result is not too sensitive to the choice of non-thermal spectral index ($\alpha = -0.9$ gives $S^{th}_{5\rm GHz} = 0.19$ Jy). From Eqn. 2 we estimate a spontaneous H53α flux corresponding to the derived thermal flux density for *optically thin* HII regions of $(1.7 - 3.1) \times 10^{-21}$ Wm^{-2}, that is, roughly *one quarter* of the observed line emission.

A spontaneous H53α fraction of 25% is not inconsistent with other available data. Converting the estimated spontaneous flux into an ionization rate gives a ratio $N_{Lyc}/L_{tot} = (4.4 \pm 1.7) \times 10^{43}$ s^{-1} L$_\odot{}^{-1}$, where the assigned 1σ error derives from uncertainties in the spontaneous fraction (30%) and in the infrared luminosity (25%). This ratio is still large, comparable in fact to the value obtained for M82 assuming that all of its H53α emission is spontaneous, but not unreasonably so if the lower stellar mass function is truncated at a relatively high value (see Paper I). A smaller spontaneous fraction, about 10%, results if the HII regions are optically thick. In this case, the ratio N_{Lyc}/L_{tot} would be lower and more consistent with a normal stellar population.

The spontaneous contribution to the H53α emission can also be compared with an Hα flux of 3.8×10^{-15} Wm^{-2}, estimated within a $40''$ aperture to be 80% of the total map flux of Young, Kleinmann & Allen (1988). Using the reddening law of Rieke & Lebofsky (1985), we calculate a visual extinction of $A_V = 4.8 - 5.4$ mag, depending on n_e and T_e, with an additional uncertainty due to H53α and Hα calibrations of $\Delta A_V = 0.3$ mag. Although this value for the extinction is consistent with indirect estimates from several sources (see discussion in Young *et al.* 1988), there being no direct measurements of reddening from recombination line ratios, for example, we caution that the visual extinction is not a strong constraint on the spontaneous H53α flux. Even if all of the observed flux were spontaneous this would imply an extinction in the range $A_V = 6.6 - 7.2$ mag, which would still be consistent with other estimates.

We now investigate whether plausible physical conditions might result in the apparent large stimulated fraction of H53α in NGC 2146. Following Bell & Seaquist (1978) we write the peak total line flux density from an HII region as :

$$S_L^{tot} = \Omega B(\nu_L, T_e) \left[\left(\frac{\tau_C + b_n \tau_L^*}{\tau_C + \gamma \tau_L^* b_n} \right) \left(1 - e^{-(\tau_L + \tau_C)} \right) - (1 - e^{-\tau_C}) \right] + S_\nu e^{-\tau_C} \left(e^{-\tau_L} - 1 \right) \qquad (3)$$

where the first term represents the spontaneous and stimulated emission from the cloud itself, and the second term accounts for stimulation by background continuum radiation. Here, Ω is the solid angle subtended by the HII regions to earth, $B(\nu_L, T_e)$ is the Planck function evaluated at the line frequency ν_L, τ_C and τ_L are the continuum and line optical depths, respectively, τ_L^* is the LTE line optical depth,

b_n and γ describe the departure of the level population from LTE, and S_ν is the flux density of the background continuum. The spontaneous line emission, alone, is :

$$S_L^{spont} = \Omega B(\nu_L, T_e) e^{-\tau_C} (1 - e^{-\tau_L}) \tag{4}$$

and the thermal continuum emission :

$$S_c^{th} = \Omega B(\nu_L, T_e) (1 - e^{-\tau_C}) \tag{5}$$

The line and continuum optical depths are defined :

$$\tau_L = b_n \gamma \tau_L^* \tag{6}$$

$$\tau_L^* = 1.7 \times 10^3 \left(\frac{\Delta\nu_L}{\mathrm{kHz}}\right)^{-1} \left(\frac{T_e}{\mathrm{K}}\right)^{-2.5} \left(\frac{EM}{\mathrm{cm^{-6}pc}}\right) \tag{7}$$

$$\tau_C = 8.2 \times 10^{-2} \left(\frac{T_e}{\mathrm{K}}\right)^{-1.35} \left(\frac{\nu_L}{\mathrm{GHz}}\right)^{-2.1} \left(\frac{EM}{\mathrm{cm^{-6}pc}}\right) \tag{8}$$

where EM is the emission measure and $\Delta\nu_L$ is the linewidth. The parameter γ is defined as :

$$\gamma = 1 - \frac{kT_e}{h\nu_L} \frac{\mathrm{d} \ln b_n}{\mathrm{dn}}. \tag{9}$$

Values of γ and b_n are tabulated by Brocklehurst (1970) for various values of n_e and T_e.

Using this formulism, we have analysed the emission from model HII regions. The constraints on these models are as follows : (*a*) linewidth — as an absolute maximum $\Delta \mathrm{v} \simeq 300$ km s^{-1}, from Fig. 2b, but considerably less for small sources ($\Delta \mathrm{v} \lesssim 25$ km s^{-1}). (*b*) source size — the maximum is $20\widehat{''} \times 10\widehat{''}$, comparable to the extent of radio continuum and Hα emission. For reference, at a distance $D = 17.2$ Mpc, a source subtending an angle $\theta = 0.1\widehat{''}$ has a linear size of about 8 pc. An O5 star having $T_{eff} = 45{,}000$ K and $N_{Lyc} = 10^{49.46}$ s^{-1} (*e.g.* Puxley, Hawarden & Mountain 1990) will generate an HII region whose Strömgren diameter is 9 pc if $n_e = 100$ cm^{-3}. (*c*) Electron density — ranging from 100 cm^{-3} for evolved HII regions up to 10^5 cm^{-3} for young ultracompact sources. Note that the choice of n_e, together with the angular source extent, effectively fixes the emission measure because $EM \simeq n_e^2 D \Omega^{1/2}$. (*d*) Radio continuum emission — the observed 5 GHz continuum flux density integrated over the central $40\widehat{''}$ is about 300 mJy with a spectral index of $\alpha = -0.5$. The peak flux density in a $1.2\widehat{''}$ beam at this frequency is 6 mJy (Kronberg & Biermann 1981). (*e*) Line emission — the integrated H53α flux is approximately 10^{-20} Wm^{-2}, but more than about three quarters of this must be stimulated so as not to conflict with the radio continuum emission and total luminosity (§4). The free parameters in each model are the electron density and temperature, source size and linewidth. Departure coefficients for the n=54 level are set by the choice of n_e and T_e and were derived from the work of Brocklehurst (1970) by Lagrangian interpolation. Two distinct series of models were examined.

4.0.1 Models of an extended HII region complex

The first class of models represents an HII region complex of maximal extent ($20^{\widehat{m}} \times 10^{\widehat{m}}$), similar to those considered by Bell & Seaquist (1978) to account for the radio recombination lines in M82. The results of varying the other free parameters are summarized in Table 3. The columns in this Table are : (1) model code; (2) electron density; (3) electron temperature; (4) linewidth; (5) peak H53α line flux density; (6) integrated H53α flux; (7) stimulated fraction of total H53α line emission; and (8) thermal continuum flux density. It is apparent that no combination of parameters can reproduce the required values cited in the final row of the Table. Models (Id – Ig) with an electron density $n_e \geq 10^3$ cm^{-3} produce too much line and continuum emission simply because of the large number of emitting particles. Whilst reducing the emitted intensity, the electron density of 100 cm^{-3} in models (Ia – Ic) implies only a small value of the parameter γ (eqn. 9) and thereby results in negligible stimulated emission. The only model shown which includes non–thermal background radiation is Ic for which we adopt $S_\nu = 300$ mJy at 5 GHz. It can be seen that this causes no significant enhancement of the stimulated emission because of the weak amplification. It is clear from these crude models that only with an electron density greater than about 10^3 cm^{-3} are the departure coefficients appropriate for producing the neccessary amount of stimulated emission, but that such HII regions must be small so as not to exceed the constraints set by the observed line flux and continuum flux density.

4.0.2 Models with compact HII regions

To reproduce the observed radio spectral index, our second grid of models comprises two components : compact HII regions, satisfying the additional conditions identified in §4.1.1, and non–thermal continuum emission. For simplicity, the compact HII regions are assumed to be entirely responsible for the recombination line emission even though one expects a contribution from less dense regions. Letting the continuum emission from the compact HII regions vary as $S_\nu \propto \nu^2(1 - e^{-\tau_\nu})$ and the non–thermal emission as $S_\nu \propto \nu^{-0.75}$, and given the observed flux density and spectral index, we solve the appropriate simultaneous equations as in §4 to find :

$$N S_5^{th} \left[1 - 0.03 \left(\frac{1 - e^{-14.5\tau_5}}{1 - e^{-\tau_5}} \right) \right] = 78 \qquad (10)$$

where N is the number of (identical) compact HII regions, S_5^{th} is their 5 GHz flux density in mJy and τ_5 is their optical depth at this frequency. For the H53α line emission, the equivalent constraint is simply : $N F_{H53\alpha}^{tot} \sim 10^{-20}$Wm^{-2}. So in the limiting cases of very optically thin ($\tau_5 << 1$) and optically thick ($\tau_5 >> 1$) HII regions we arrive at the following approximate limits on their line and continuum radiation : $N S_5^{th} = 80 - 140$ mJy and $S_5^{th}/F_{H53\alpha}^{tot} = 8 - 14$ mJy/10^{-21} Wm^{-2}. From comparison of the former constraint with the observed peak 5 GHz surface brightness of 6 mJy/beam we see immediately that many individual sources are required ($N \gtrsim 15$). Table 4 displays the model line and continuum

emission per HII region obtained by varying the free parameters in order to try and match these and the previously imposed restrictions which are summarized in the last row of the Table. The columns are : (1) model code; (2) electron density; (3) electron temperature; (4) source diameter; (5) linewidth; (6) peak H53α line flux density; (7) integrated H53α flux; (8) stimulated H53α fraction; (9) 5 Ghz thermal emission from compact HII regions; (10) ratio of 5 GHz thermal emission to integrated line flux; and (11) number of HII regions ($N = 10^{-20}/F^{tot}_{H53\alpha}$). Since the HII regions subtend only a small solid angle, stimulation by background radiation has not been incorporated into any of the models shown; all of the stimulated H53α flux is caused by the continuum of the HII region itself. Examination of Table 4 shows that models in which the emission measure is small because of relatively low density or small source size (IIa – IIe) do not provide a sufficient optical depth for substantial stimulation by the continuum even though the parameter γ may be large. Increasing the source diameter for a given electron density enhances the stimulated fraction, *e.g.* compare IIa – IIc, but for $n_e \lesssim 10^4$ cm^{-3} the restrictions imposed by the total line and continuum flux densities intervene before the required fraction is achieved. The failure of models with high gas density such as IIk and III to match the predicted ratio of $S^{th}_5/F^{tot}_{H53\alpha}$ occurs because of their large optical depth at lower frequencies. However, as cautioned at the beginning of this section, we cannot necessarily exclude them on these grounds because a more sophisticated model would incorporate lower density HII regions (*e.g.* IIa) which redress the overall balance by exhibiting large $S^{th}_5/F^{tot}_{H53\alpha}$ ratios. Nonetheless, several of the models with $N_e \geq 10^4$ cm^{-3} satisfy all of the imposed conditions on their own (see IIf, IIg and IIj). If we now consider the ionizing photon fluxes which would be required to produce such an HII region :

$$\left(\frac{N_{Lyc}}{\mathrm{s}^{-1}}\right) = 1.6 \times 10^{48} \left(\frac{n_e}{\mathrm{cm}^{-3}}\right)^2 \left(\frac{\theta}{\mathrm{arcsec}}\right)^3 \qquad (11)$$

we find that models IIf and IIg would need several hundred O5 stars ($N_{Lyc} = 10^{49.46}$ s^{-1}), whereas an HII region like that described by model IIj requires only a single late O star. Thus we conclude that in order to produce a significant fraction of stimulated H53α radiation without exceeding the absolute line flux or continuum flux density observed in NGC 2146, a population of dense ($n_e \sim 10^5$ cm^{-3}) compact ($\theta \sim 0.001$ $\widehat{''}$) HII regions is required.

4.1 Comparison with M82

In Paper I we deduced that the radio recombination line data for M82 indicates that the mm–wavelength lines H53α and H40α are almost entirely due to spontaneous emission. This conclusion is illustrated in Fig. 3 where we show the peak H$n\alpha$ line flux density as a function of the line frequency for several models : (i) the best–fit to M82 from Bell and Seaquist (1978, $n_e = 150$ cm^{-3}); HII regions with (ii) $n_e = 10^5$ cm^{-3} (model IIk from Table 4) and (iii) $n_e = 10^4$ cm^{-3} (model IIf from Table 4); and (iv) pure spontaneous emission. All of the models have been normalised relative to their *spontaneous* H53α flux density. The dependence of the γ coefficients for a given level on the electron density is readily apparent : as n_e

increases, the excess (stimulated) flux over the spontaneous emission dominates at a higher frequency (lower principal quantum number). At low frequencies, the line strength in dense HII regions deviates below that expected for spontaneous emission because of the high optical depth. Inset along the abscissa are marked several Hnα lines of interest.

It can be see from Fig. 3 that the model fit of Bell and Seaquist (1978) to radio recombination lines between H76α and H166α is unconstrained by lines at frequencies higher than about 20 GHz and therefore this data cannot exclude the presence of dense HII regions. However, we note that the ratio of H40α/H53α line fluxes differs significantly between the models. If compact HII regions were the dominant source of H53α photons then the line flux ratio would be about 2.2 (loci (ii) and (iii)), compared with $\sim$5.0 for spontaneous emission. The observed value of $F_{H40\alpha}/F_{H53\alpha} = 4.9^{+1.0}_{-1.2}$ suggests that dense HII regions do not dominate the H53α line emission, although the data is insufficient to preclude a moderate contribution.

5 Conclusions

The hydrogen recombination lines H40α and H53α have been searched for in 3 bright spiral galaxies. The H53α transitions was detected in NGC 2146, possibly in NGC 1068, but not in M83. Comparison of the line strength with radio continuum data indicates that a substantial fraction of the H53α flux is produced by stimulated emission. Models show that the emitting ionized regions in NGC 2146 must be dense, in order that their ratio of stimulated to spontaneous line emission be large, small, so as not to exceed the inferred thermal continuum emission, and numerous.

This picture differs greatly from that deduced for M82, the only other galaxy in which H53α emission has been detected, where the relative strength of H40α and H53α recombination lines argues for spontaneous emission. The reason for this disparity is uncertain but we speculate that it may reflect different stages in the evolution of the starburst, with NGC 2146 being significantly less evolved than M82. Confirmation of this hypothesis requires measurements of other high frequency recombination lines, such as H66α at 22.4 GHz and H30α at 231.9 GHz, to discriminate between the two emission processes and further constrain physical conditions in the ionized gas.

Acknowledgments

We thank Joss Bland for kindly deriving the integrated Hα flux from NGC 1068, Tim Hawarden for assistance at UKIRT, and Nobeyama Radio Observatory for their continued support of this project. PJP is grateful to SERC Fellowships for financial support.

References

Acker, A., Stenholm, B. & Tylenda, R., 1989. *Astr. Astrophys. Suppl.*, **77**, 487.

Atherton, P.D., Reay, N.K. & Taylor, R., 1985. *Mon. Not. R. astr. Soc.*, **216**, 17P.

Baldwin, J.A., Wilson, A.S. & Whittle, M., 1987. *Astrophys. J.*, **319**, 84.

Balick, B., 1987. *Astron. J.*, **94**, 671.

Bell, M.B. & Seaquist, E.R., 1978. *Astrophys. J.*, **223**, 378.

Benvenuti, P., Capaccioli, M. & D'Odorico, S., 1975. *Astr. Astrophys.*, **41**, 91.

Brocklehurst, M., 1970. *Mon. Not. R. astr. Soc.*, **148**, 417.

Cecil, G., Bland, J. & Tully, R.B., 1990. Submitted to *Astrophys. J.*

Combes, F., Encrenaz, P.J., Lucas, R. & Weliachew, L., 1978. *Astr. Astrophys.*, **67**, L13.

DePoy, D.L., 1987. *Ph.D. Thesis*, University of Hawaii.

de Vaucouleurs, G., de Vaucouleurs, A. & Corwin, H.G., 1978. *Astron. J.*, **83**, 1331.

Downes, D., 1989. In : *Introductory Course in Galaxies' Evolution and Observational Astronomy*, eds. Appenzeller, I., Habing, H., & Lena, P., (Springer–Verlag, Heidelberg), in press.

Hummer, D.G. & Storey, P.J., 1987. *Mon. Not. R. astr. Soc.*, **224**, 801.

Jackson, J.M. & Ho, P.T.P., 1988. *Astrophys. J.*, **324**, L5.

Kaler, J.B., 1986. *Astrophys. J.*, **308**, 322.

Keel, W.C., 1984. *Astrophys. J.*, **282**, 75.

Kohoutek, L. & Martin, W., 1981. *Astr. Astrophys. Suppl.*, **44**, 325.

Kronberg, P.P. & Biermann, P., 1981. *Astrophys. J.*, **243**, 89.

Kurucz, R.L., 1979. *Astrophys. J. Suppl.*, **40**, 1.

Milne, D.K. & Aller, L.H., 1975. *Astr. Astrophys.*, **38**, 183.

Puxley, P.J., 1988. *Ph.D. Thesis*, University of Edinburgh.

Puxley, P.J., Brand, P.W.J.L., Moore, T.J.T., Mountain, C.M., Nakai, N. & Yamashita, T., 1989. *Astrophys. J.*, **345**, 163. (Paper I)

Puxley, P.J. & Brand, P.W.J.L., 1990. In preparation.

Puxley, P.J., Hawarden, T.G. & Mountain, C.M., 1990. Submitted to *Astrophys. J.*

Rieke, G.H. & Lebofsky, M.J., 1985 *Astrophys. J.*, **288**, 618.

Schneider, S.E., Terzian, Y., Purgathofer, A. & Perinotto, M., 1983. *Astrophys. J. Suppl.*, **52**, 399.

Shaver, P.A., Churchwell, E. & Walmsley, C.M., 1978. *Astr. Astrophys.*, **64**, 1.

Shull, J.M. & McKee, C.F., 1979. *Astrophys. J.*, **227**, 131.

Stanghellini, L. & Kaler, J.B., 1989. *Astrophys. J.*, **343**, 811.

Telesco, C.M. & Decher, R., 1988. *Astrophys. J.*, **334**, 573.

Thronson, H.A. & Greenhouse, M.A., 1988. *Astrophys. J.*, **327**, 671.

Tully, R.B., 1988. *Nearby Galaxies Catalog*, Cambridge University Press.

Turner, J.L., Ho, P.T.P. & Beck, S.C., 1987. *Astrophys. J.*, **313**, 644.

Walmsley, C.M., Churchwell, E. & Terzian, Y., 1981. *Astr. Astrophys.*, **96**, 278.

Williams, T.C., Hackwell, J., Gehrz, R.D. & Grasdalen, G.L., 1988. *Publ. astr. Soc. Pac.*, **100**, 124.

Wilson, A.S. & Ulvestad, J.S., 1982. *Astrophys. J.*, **263**, 576.

Wilson, O.C., 1950. *Astrophys. J.*, **111**, 279.

Wynn–Williams, C.G., Becklin, E.E. & Scoville, N.Z., 1985. *Astrophys. J.*, **297**, 607.

Young, J.S., Claussen, M.J., Kleinmann, S.G., Rubin, V.C. & Scoville, N., 1988. *Astrophys. J.*, **331**, L81.

Young, J.S., Kleinmann, S.G. & Allen, L.E., 1988. *Astrophys. J.*, **334**, L63.

Table 1: Observations and derived ionization rates for the calibration source, NGC 6210.

Line/Freq.	Integrated Flux (Wm^{-2})	Ionization Rate (10^{47} s^{-1}) [a]
H53α	$(4.56 \pm 0.70) \times 10^{-21}$	1.51 ± 0.23
Brγ	$(3.38 \pm 0.34) \times 10^{-15}$	1.33 ± 0.13
Paβ	$(4.08 \pm 0.41) \times 10^{-15}$	1.25 ± 0.13 [b]
5 GHz	(0.311 ± 0.015) Jy	1.19 ± 0.06

[a] adopted distance of 2.08 kpc (Milne & Aller 1975). [b] scaled using small aperture ($4''$) and total Hβ fluxes from Acker *et al.* (1989) and Kohoutek & Martin (1981).

Table 2: H53α and H40α results for three bright spiral galaxies.

Galaxy Name	Distance (Mpc)	Source Size ($''$)	$F_{H53\alpha}$ (10^{-21} Wm^{-2})	$F_{H40\alpha}$ (10^{-21} Wm^{-2})	N_{Lyc} (s^{-1})
NGC 1068	14.4	25	9.5 ± 3.3	<100	$(9.5 \pm 3.3) \times 10^{54}$
NGC 2146	17.2	14	9.5 ± 1.7	<71	$(1.4 \pm 0.3) \times 10^{55}$
M83	4.7	15	<3.7	<71	$< 3.9 \times 10^{53}$

Table 3: Model results for an HII region with extent $20\widehat{''} \times 10\widehat{''}$. The observed and inferred values for NGC 2146 are given in the final row.

Model Code	n_e (cm^{-3})	T_e (K)	Δv (km s^{-1})	$S^{tot}_{H53\alpha}$ (mJy)	$F^{tot}_{H53\alpha}$ (10^{-21} Wm^{-2})	$F^{stim}_{H53\alpha}/F^{tot}_{H53\alpha}$ (%)	S^{th}_{5GHz} (mJy)
Ia	100	10^4	100	180	26	1.0	2.9×10^3
Ib	100	10^4	200	90	26	0.8	2.9×10^3
Ic	100	10^4	200	90	26	0.9	2.9×10^3
Id	10^3	5×10^3	200	1.4×10^5	4.0×10^4	88	1.3×10^4
Ie	10^3	10^4	200	2.9×10^4	8.2×10^3	71	2.6×10^4
If	10^3	10^4	300	1.9×10^4	8.0×10^3	70	2.6×10^4
Ig	10^4	10^4	300	8.0×10^7	3.5×10^7	100	2.6×10^4
Target Values				~ 30	~ 10	$\gtrsim 75$	~ 140

Table 4: Model results for compact HII regions. Constraints imposed by observations of NGC 2146 and by the assumptions described in the text (§4.1) are given in the last line.

Model Code	n_e (cm^{-3})	T_e (K)	θ ($''$)	Δv (km s^{-1})	$S^{tot}_{H53\alpha}$ (mJy)	$F^{tot}_{H53\alpha}$ (10^{-21} Wm^{-2})	$F^{stim}_{H53\alpha}/F^{tot}_{H53\alpha}$ (%)	S^{th}_{5} (mJy)	$S^{th}_{5}/F^{tot}_{H53\alpha}$ (mJy/10^{-21} Wm^{-2})	N
IIa	10^3	10^4	0.1	25	0.03	1.0×10^{-3}	2	0.11	110	10^4
IIb	10^3	10^4	0.5	25	3.9	0.14	9	12	84	70
IIc	10^3	10^4	1.0	25	35	1.2	17	78	65	8
IId	10^3	5×10^3	0.5	25	12	0.42	22	11	28	25
IIe	10^4	10^4	0.01	25	4×10^{-3}	1.5×10^{-4}	21	7.8×10^{-3}	27	7×10^4
IIf	10^4	10^4	0.05	25	1.0	0.04	61	0.30	8	250
IIg	10^4	10^4	0.05	10	4.0	0.06	76	0.30	5	170
IIh	10^4	5×10^3	0.05	25	8.1	0.3	88	0.17	0.6	33
IIi	10^4	10^4	0.1	25	15	0.5	80	1.4	3	20
IIj	10^5	10^4	0.001	25	7×10^{-4}	2.5×10^{-5}	53	1.4×10^{-4}	6	4×10^5
IIk	10^5	10^4	0.005	25	0.4	1.4×10^{-2}	93	3.4×10^{-3}	0.2	700
III	10^5	10^4	0.005	10	10	0.14	99	3.4×10^{-3}	0.02	71
Target Values					< 30	< 10	$\gtrsim 75$	< 6	$\sim 8 - 18$	$\gtrsim 15$

Figure Captions

Figure 1

Hydrogen recombination line observations of the calibration source NGC 6210 : (a) H53α, (b) Brγ, and (c) Paβ.

Figure 2

H53α recombination line spectra of the galaxies (a) NGC 1068 and (b) NGC 2146.

Figure 3

Peak Hnα line flux density as a function of line frequency for (i) the best fit to radio recombination line data in M82; models of dense HII regions with (ii) $n_e = 10^5$ cm^{-3} and (iii) $n_e = 10^4$ cm^{-3}; and (iv) pure spontaneous emission. All of the models have been normalised to the same spontaneous H53α flux density.

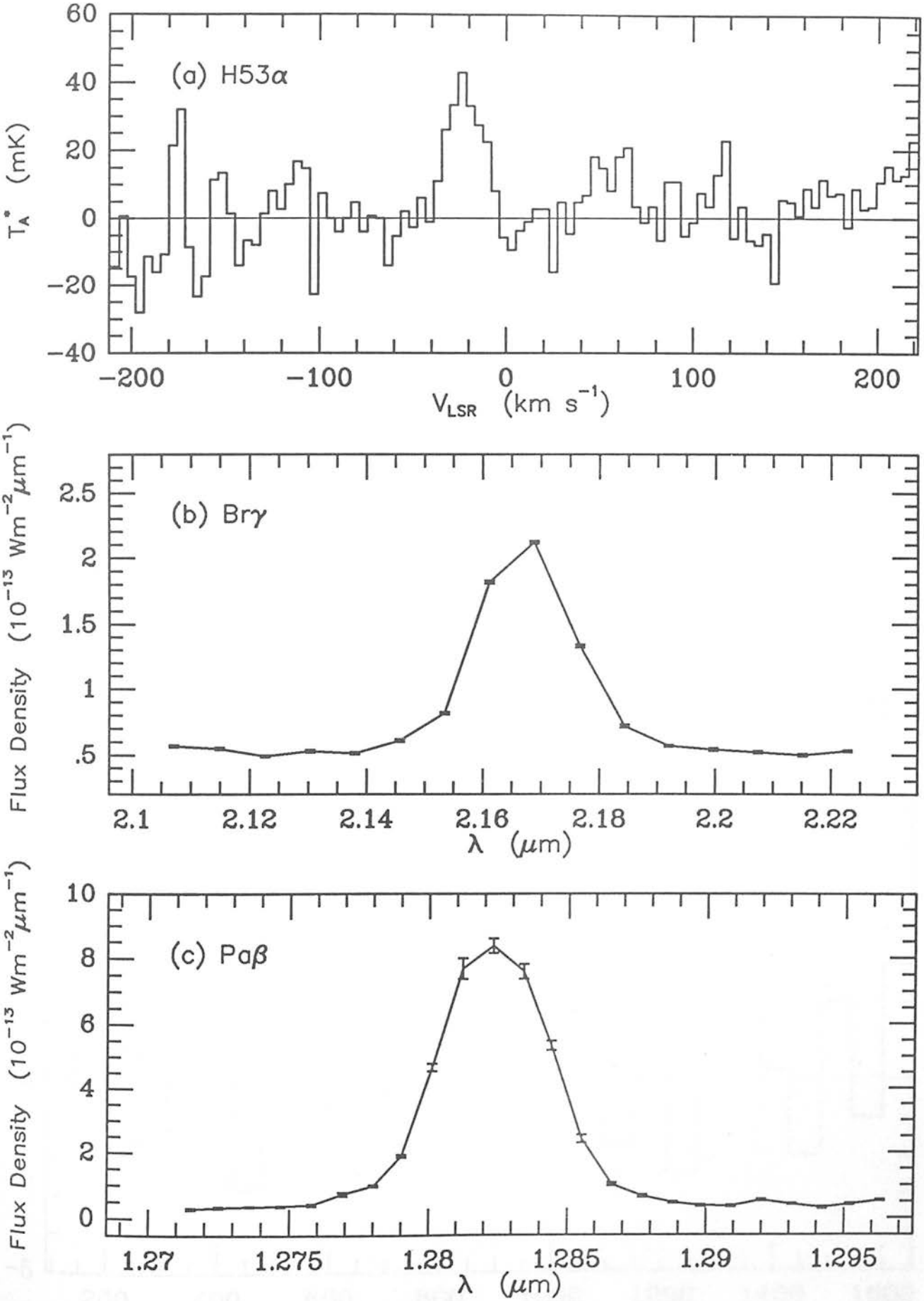
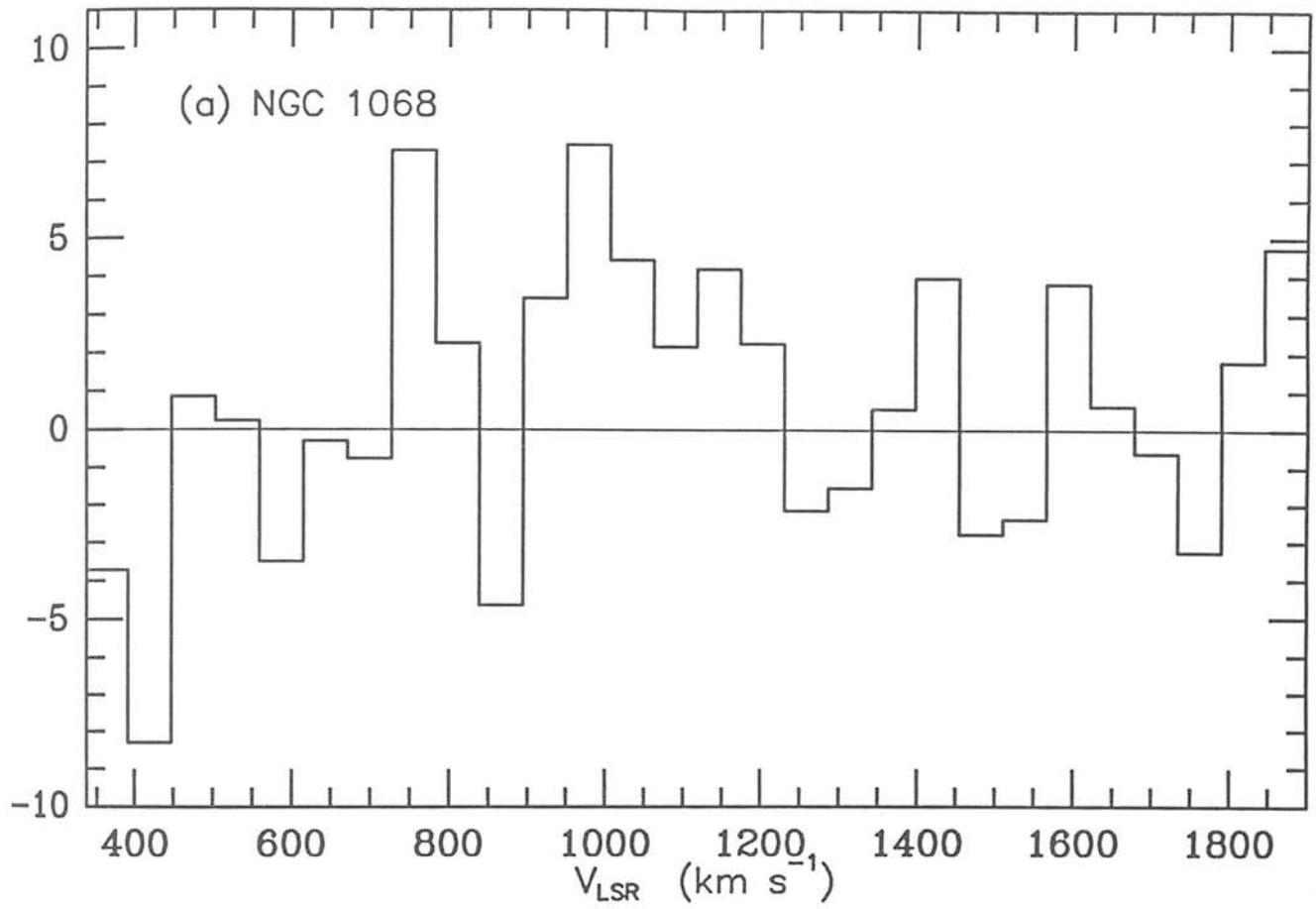
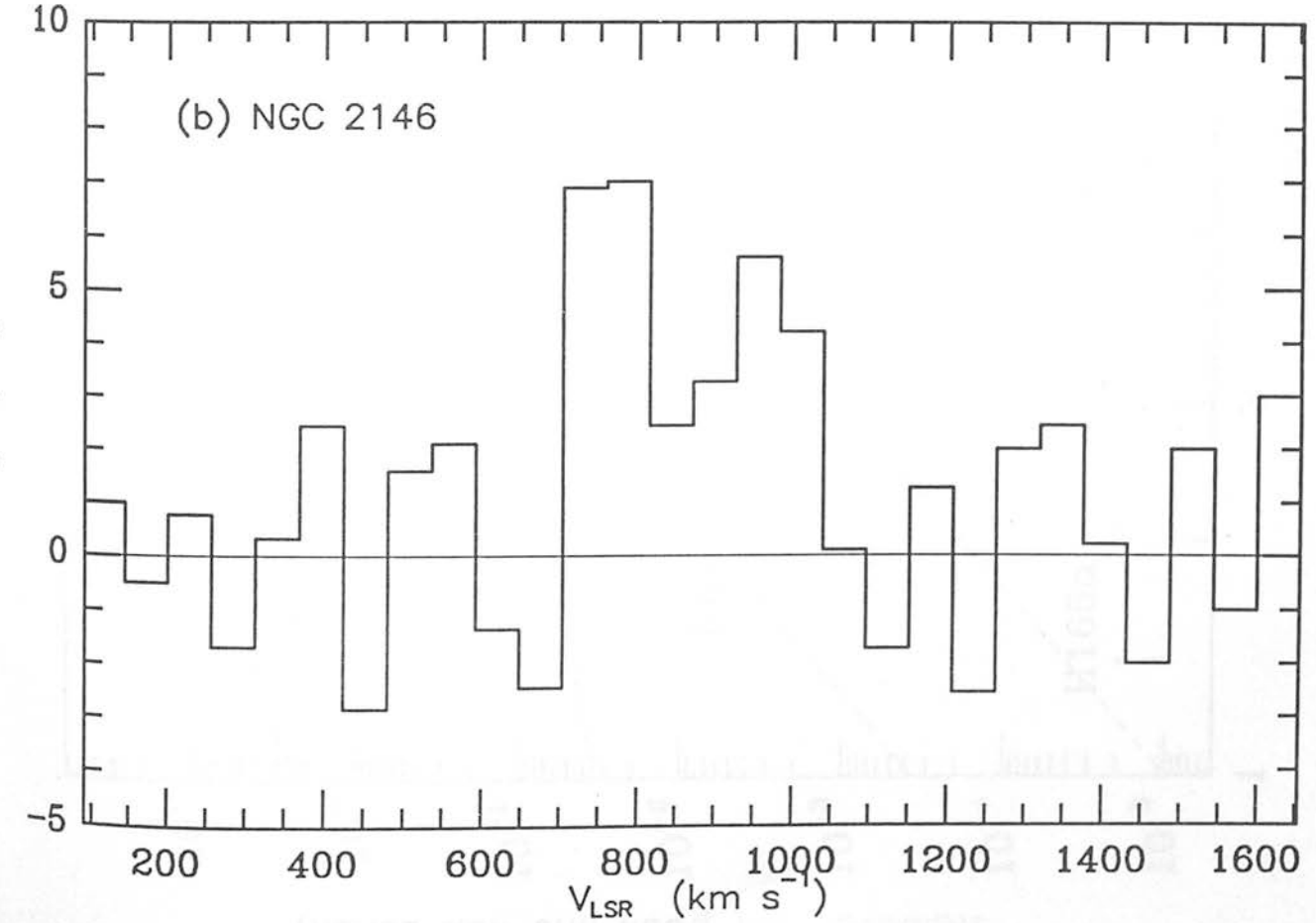
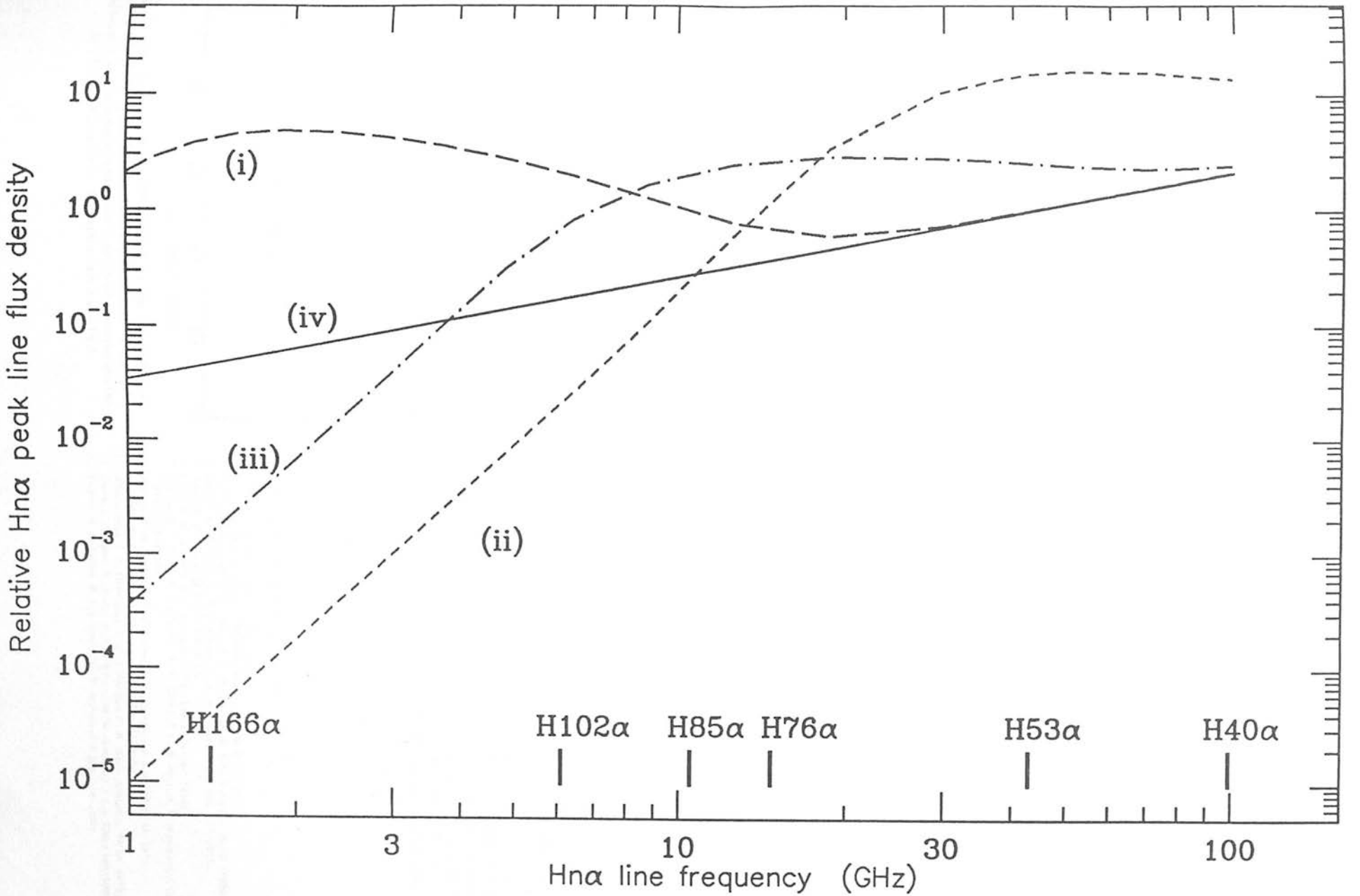

(a) H53α
T_A^* (mK)
V_{LSR} (km s^{-1})
(b) Brγ
Flux Density (10^{-13} Wm$^{-2}\mu$m^{-1})
λ (μm)
(c) Paβ
Flux Density (10^{-13} Wm$^{-2}\mu$m^{-1})
λ (μm)

Fig 2


(a) NGC 1068
10
5
0
−5
−10
400
600
800
1000
1200
1400
1600
1800
V_{LSR} (km s^{-1})


(b) NGC 2146
10
5
0
−5
200
400
600
800
1000
1200
1400
1600
V_{LSR} (km s^{-1})


Relative Hnα peak line flux density
10^{1}
10^{0}
10^{-1}
10^{-2}
10^{-3}
10^{-4}
10^{-5}
(i)
(ii)
(iii)
(iv)
H166α
H102α
H85α
H76α
H53α
H40α
1
3
10
30
100
Hnα line frequency (GHz)

Reprinted from Nature, Vol. 295, No. 5846, pp. 214-215, January 21 1982

Occurrence of the 3.3-μm feature in galaxies

T. J. Lee*, D. H. Beattie*, Ian Gatley*, P. W. J. L. Brand†, T. Jones‡ & A. R. Hyland‡

* UK Infrared Telescope, 900 Leilani Street, Hilo, Hawaii 96720, USA
† Department of Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK
‡ Australian National University, Canberra ACT, Australia

Many emission features have been discovered, in dusty ionized regions such as planetary nebulae or H II regions, with wavelengths longer than 2 μm (ref. 1). Some are gaseous emission lines but many that are too broad to be due to single atoms are attributed to emission from dust grains in or near the ionized gas. The feature at 3.28 μm ('the 3.3-μm feature') has been observed in H II regions[2,3] in planetary nebulae[4] and in particular, in the following extragalactic objects: 3C273 (ref. 5) (quasar), NGC4151 (ref. 6), M82 (ref. 7), NGC253 (ref. 8), NGC2903 (ref. 9), NGC5195 (ref. 9) and NGC6946 (ref. 9). We report here new measurements which contradict some of these results. In particular, we find the feature in the late-type galaxy M83 but do not observe it in either the Seyfert galaxy NGC4151 or the quasar 3C273.

The mechanism that produces the line is thought to be due either to fluorescence of the grains in a UV radiation field[10] (but this requires very high efficiencies) or to the heating of very small grains to high temperatures (several hundred degrees)[11]. The presence of the feature in extragalactic objects such as late-type galaxies is not surprising because these objects contain many H II regions. Measures of equivalent width of the feature will indicate the chemical composition, once studies of galactic objects have established the exact nature of the mechanism. In objects such as Seyfert galaxies and quasars, however, the presence or absence of the 3.3μm feature is important as it is one of the few unambiguous dust indicators available to settle the question of the presence of dust grains in the nuclei of active galaxies.

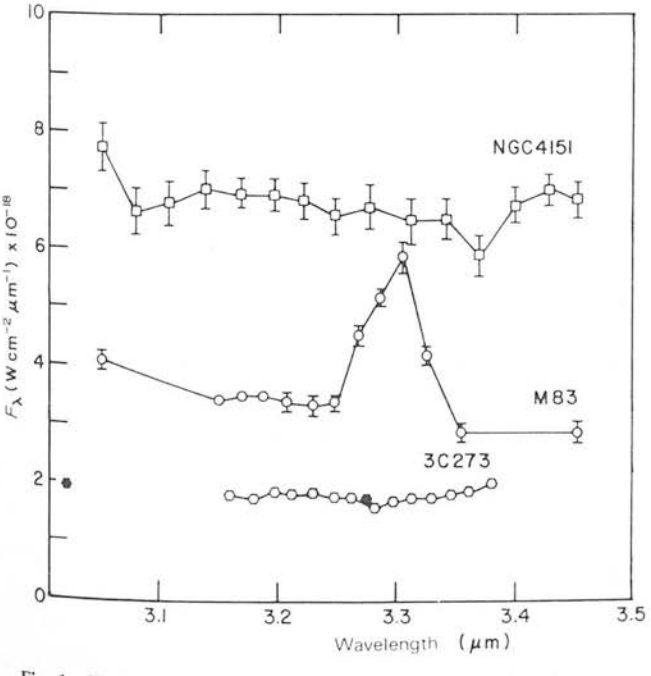


Fig. 1 The circular variable filter spectra of the three objects: NGC4151, M83 and 3C273. The wavelengths are rest wavelengths and broad-band filter measurements are also plotted for the quasar 3C273 as solid symbols. The error bars are ±1 s.d.

As part of a programme to observe a larger number of objects, we have measured the spectra of 3C273, M83 and NGC4151 in the region of the local rest wavelength of the 3.3-μm feature. This was carried out on the nights of 21, 22 and 23 May (UT) 1981, one object being measured each night. In addition, NGC4151 had been observed on 12 April 1981. During these nights the sky was clear above Mauna Kea summit and the visibility good. The telescope and instrumentation used were the UK Infrared Telescope (3.8 m) with an InSb CVF spectrophotometer (1% resolution). Apertures and chopper throws respectively were 7.3 and 13 arc s for 3C273, 10.8 and 18 arc s for M83, and 14.4 and 18 arc s for NGC4151. The wavelengths at which measurements were taken were spaced at 0.02-μm intervals near to the position of the feature and a few more widely spaced points were also measured. In addition broad-band data were obtained. At each observed point several 10-s integrations were made and total integration times were 5,600 s for 3C273, 4,820 s for M83 and 5,640 s for NGC4151. The spectra are shown in Fig. 1, where the flux density F_λ is plotted against the rest wavelength. For both 3C273 and M83 two spectra were averaged, while for NGC4151 the April data are shown.

The April data on NGC4151 are the more extensive in wavelength and of better signal-to-noise ratio. These data set a 3σ upper limit of 8% of continuum on line height above the continuum (while the May data set a poorer upper limit of 23%). For 3C273, the corresponding limit is 9% of continuum. These limits are conservative, being based on the assumption of a flat baseline, and the spectral profile of the CVF. If the continuum outside the region of the line were to be slightly curved this would reduce the upper limits by a factor of ~2. These limits should be compared with previously[5,6] suggested values of 30% and 40% respectively. Although both 3C273 and NGC4151 are variable by a few tens of per cent, our continuum levels are sufficiently close to, and also fainter than, the previously measured continuum levels, and this cannot cause such a discrepancy.

In M83, a late-type galaxy with H II regions in its nuclear region[12], the 3.3-μm feature is clearly seen in Fig. 1. Equally clearly, at the levels discussed above, the feature is not present in either the Seyfert galaxy NGC4151 or the quasar 3C273. The excellent transparency in the 3-μm region at the summit of Mauna Kea and high instrumental instantaneous signal-to-noise ratios mean that previous detection of the 3.3-μm feature must be considered suspect. This experiment shows no evidence for dust in the Seyfert galaxy NGC4151 or the quasar 3C273, but clearly demonstrates that the dust in M83 and in H II regions in our Galaxy share similar properties.

Received 7 October; accepted 30 November 1981.

1. Aitken, D. K. *IAU Symp. No.* **96**, (1980).
2. Grasdalen, G. L. & Joyce, R. R. *Astrophys. J. Lett.* **205**, L11 (1976).
3. Aitken, D. K., Roche, P. F., Spencer, P. M. & Jones, B. *Astr. Astrophys.* **76**, 60, (1979).
4. Merrill, K. M., Soifer, B. T. & Russell, R. W. *Astrophys. J. Lett.* **200**, L37 (1975).
5. Allen, D. A. *Nature* **284**, 323-324 (1980).
6. Cutri, R. M. & Rudy, R. J. *Astrophys. J. Lett.* **241**, L141-144 (1980).
7. Willner, S. P., Soifer, B. T., Russell, R. W. Gillett, F. C. & Joyce, R. R. *Astrophys. J. Lett.* **217**, L121 (1977).
8. Russell, R. W., Soifer, B. T. & Willner, S. P. *Astrophys. J. Lett.* **217**, L148 (1977).
9. Lebofsky, M. J. & Rieke, G. H. *Astrophys. J.* **229**, 111 (1979).
10. Allamandola, L. J., Greenberg, J. M. & Norman, C. A. *Astro. Astrophys.* **77**, 66 (1979).
11. Dwek, E., Sellgren, K., Soifer, B. T. & Werner, M. W. *Astrophys. J. Lett.* **238**, L140 (1980).
12. Dufour, R. J., Talbot, R. J. & Jensen, E. B. *Astrophys. J.* **236**, 119 (1980).

Introduction

Many types of object have in their infrared spectra several emission features which have not been satisfactorily explained. One of the more prominent of the near-infrared features is at 3.28 μm with a width of 0.05 μm.

This line has been observed in several HII regions including the Orion nebula (Sellgren 1981) and Sharpless 106 (Williams and Zealey 1983). Sellgren concluded that the 3.28 μm emission arises in the HII region - molecular cloud boundary with little or no 3.28 μm emission coming from inside the HII region itself. Williams and Zealey however conclude that there is no evidence for preferential 3.28 μm emission where the HII region interacts with the molecular cloud, but the observed variations in S-106 can be explained in terms of variations in its gas density.

We conducted the observations described in this paper to elucidate this contradiction. NGC 2024 (W12, Orion B) is the second most prominent source in the Orion complex. Its optical appearance shows a dust lane dividing the nebula, which lies in the foreground of the HII region. The identity of the exciting star for this HII region has been in question for a long time. Two possible stars are star #1 (Grasdalen 1974; also IRS 1, Johnson and Mendoza 1964), lying on the western edge of the dust lane and a bright infrared point source lying in the middle of the dust lane (star #2 Grasdalen 1974). Figure 4 in Grasdalen (1974) shows the optical appearance of NGC 2024 with the position of star #2.

An interpretation of NGC 2024 can be made using a 'blister' model (Zuckerman 1973) with star #2 as the exciting star. Figure 1 shows the

geometry of NGC 2024 using this model. However, Thompson et al (1981) have argued that neither star #1 or star #2 is the exciting star. They concluded that the exciting star is obscured by the dust and gas to the south of star #2 and has a temperature of 34,000K. In any case, the geometry indicated by both optical and infrared appearance is that of an edgeways-on blister-type HII region obscured by neutral material.

Observations

The observations were carried out at the United Kingdom Infrared Telescope (UKIRT) during February 1983. The UKT6 InSb photometer used incorporates a 1% circular variable filter (CVF) in the 2.3 - 4.6 μm wavelength range. A 5mm aperture was used to give an 8 arcsec aperture on the sky. Sky chopping and beam switching were carried out in the north-south direction with a beam separation of 3 arc minutes.

The standard star used is HD 40335 (spectral type A). During each night this star was observed several times. After the individual spectra had been ratioed with the appropriate spectrum of the standard star, and wavelength calibrated, the flux calibration was carried out using infrared photometry of HD 40335 by Elias et al (1982).

The observations were made at 13 spatial position in two strips in the north-south direction across the ionization front and into the neutral cloud. One strip passes through the peak of the 8.4 μm map of Grasdalen (1974). The second N-S strip is 16 arcsec to the east and each spatial position in the N-S direction is separated by 8 arcsec. Figure 2 shows the spectra arranged according to spatial position. The peak position in Grasdalen's (1974) map is used as our zero position.

At most spatial positions two spectra were taken; (i) 3.2 μm to 3.5 μm covering both the 3.28 μm and 3.4 μm features and (ii) 4.02 μm to 4.14 μm covering the hydrogen Brackett α(Bα) line at 4.05 μm. The remainder of the positions contained just the 3.2 μm to 3.5 μm spectra. For each spectrum an average of 6 scans were made and then coadded.

Results

The spatial map of 3.28 μm and Bα spectra is shown in Figure 2. The Bα map shows the location of the ionization front. However the 3.28 μm feature is seen to be present at all spatial positions across the ionization front demonstrating that it is not located only in the HII region but also in the neutral cloud. This means that for these emission features (both 3.28 μm and 3.4 μm) the dust is excited by non-ionizing UV photons. The strength of the 3.28 μm feature varies along the North-South strips peaking around the edge of the HII region, and being below detectable limits at 32 arcsec South. Also the equivalent width of the 3.28μm feature increases from north to south up to the ionization front. The 3.4 μm feature is seen at all spatial positions with a relative strength to the 3.28 μm feature of 0.3 ± 0.1.

In order to produce the above results, blackbody fitting was carried out to the continuum points in the 3 μm and 4 μm spectra in order to evaluate the strengths of the two features. The colour temperature so deduced is 550-600K at the position of the peak emission, and drops from this value by approximately 20 percent at either end of each scan. Allowing for an emissivity varying as λ^{-2} reduces the necessary colour temperature by thirty percent throughout. However, fitting between our continua and the 8.4 μm points of Grasdalen (1974) produces colour temperatures in the

range of 300K, whether allowing for emissivity or not. It is clear that (not unexpectedly) there must exist a range of temperatures of dust grains. Figure 3 shows the variation of equivalent width for the 16 arcsec. E strip with a maximum value of 0.63μm at the position of the ionization front. This variation is the same for the Peak strip. This is consistent with a bimodal distribution of dust in the region.

The fact that there is no sudden drop in 3.28 μm emissivity at the edge of the HII region demonstrates that the photons heating the grains are dominated by those with wavelengths greater than 912Å, and this confirms the suggestion (Thompson et al 1981) that NGC 2024 is a low excitation HII region. At their suggested temperature of 34000K, the fraction of the star's energy which is non-ionising is 0.70. If we adopt the temperature of the exciting star suggested by Grasdalen (1974), this fraction increases to 0.92. The distance over which the intensity of the feature decays to zero from the highest value 8 arcsec South of the peak is 24 arcsec, with a 1/e distance of 10 arcsec. This latter corresponds (at a distance of 450 pc, Grasdalen 1974) to 0.02 pc, which represents $\tau(\text{dust}) \sim 1$. Very approximately, this will imply a gas density of 5×10^4 atoms cm^{-3}, based on a normal gas to dust ratio, and a standard extinction law (Allen 1973). This density corresponds to the compressed shell of (presumably) molecular gas between the ionisation front and the shock propagating ahead of it. This shock, which will be an isothermal shock, can compress the gas by up to 100 times its initial density, although at this stage in the evolution of NGC 2024 (a well-developed blister) the compression will be less. Thus we predict that the density of the gas ahead of the shock will be of the order of 5000 cm^{-3}, typical of the cores of molecular clouds.

Conclusions

1. The 3.28 μm dust feature is seen to be present on either side of the ionization front in NGC 2024. This implies that the feature is excited by soft UV photons copiously produced by a low excitation source, such as the star predicted to ionise the nebula.

2. The equivalent width increases from the north to the south with a maximum value at the ionization front. This is consistent with a bimodal distribution of dust in the region.

3. The 3.4 μm feature is seen accompanying the 3.28 μm feature at all spatial positions with a relative strength of 0.3 ± 0.1, and the constancy of this ratio supports previous suggestions that the features have a common origin.

References

Allen, C.W., 1973. Astrophysical Quantities, Athlone, London.

Elias, J.H., Frogel, J.A., Matthews, K. and Neugebauer, G., 1982. Astron. J., 87, 1029.

Grasdalen, G.L., 1974. Astrophys. J., 193, 373.

Johnson, H.L. and Mendoza, E.E., 1964. Bol. Obs. Tonantzintla y Tacubaya, 3, 331

Sellgren, K., 1981. Astrophys. J., 245, 138.

Thompson, R.I., Thronson, Jr. H.A., Campbell, B.G., 1981. Astrophys. J. 249, 622

Williams, P.M. and Zealey, W.J., 1983. Mon. Not. R. astr. Soc. 203, 433

Zuckerman, B., 1973. Astrophys. J., 183, 863.

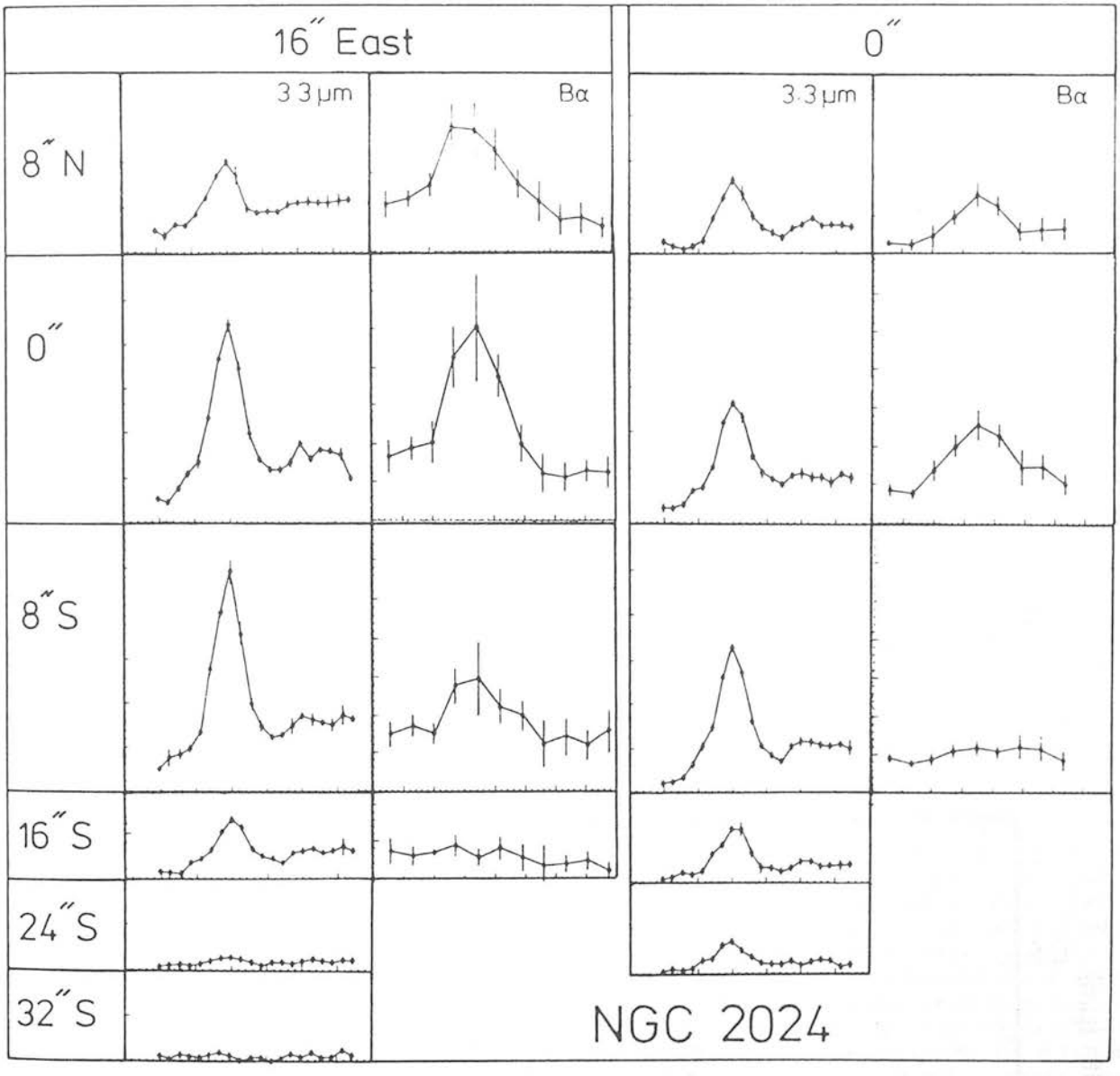


Figure 2. The 3.3 μm and 4 μm spectra in their correct relative positions in NGC 2024. The zero point on this grid is the peak position of the 8.4 μm map of Grasdalen (1974).

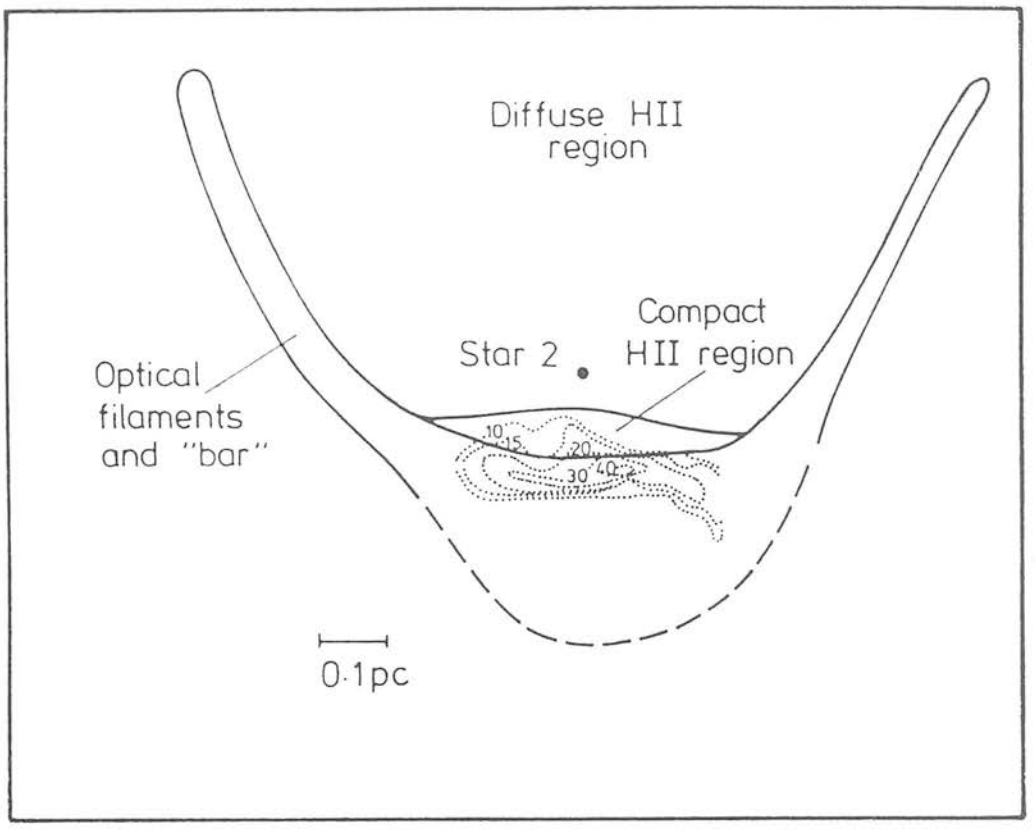


Fig. 1. A schematic of an edge-on blister model for an HII region, superposed on the 8.4 μm map of Grasdalen (1974).

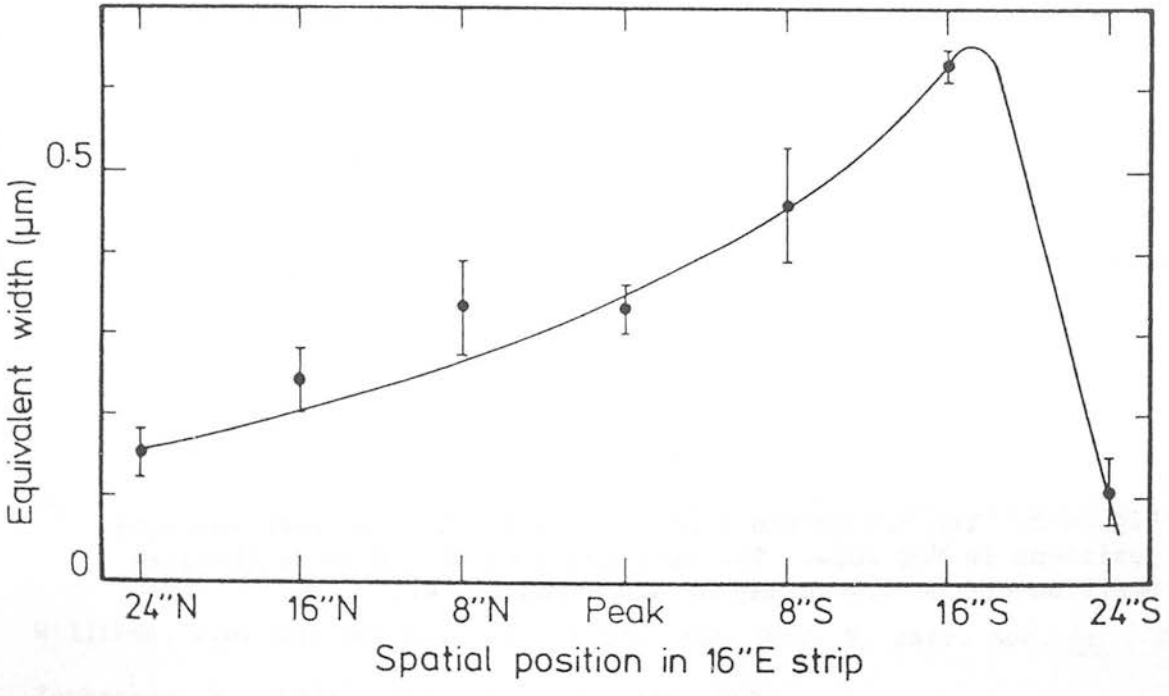


Fig. 3. The equivalent width of the 16 arcsec. E strip.

THE ASTROPHYSICAL JOURNAL, 341:278–287, 1989 June 1

42

SPATIAL VARIATIONS OF THE 3 MICRON EMISSION FEATURES WITHIN UV-EXCITED NEBULAE: PHOTOCHEMICAL EVOLUTION OF INTERSTELLAR POLYCYCLIC AROMATIC HYDROCARBONS

T. R. GEBALLE,[1,2] A. G. G. M. TIELENS,[3,4] L. J. ALLAMANDOLA,[3] A. MOORHOUSE,[5] AND P. W. J. L. BRAND[5]

Received 1988 June 17; accepted 1988 October 28

ABSTRACT

We have obtained 3 μm spectra at several positions in the Orion Bar region and in the "Red Rectangle," the nebula surrounding HD 44179. The recently discovered weak emission features at 3.40, 3.46, 3.51, and 3.57 μm (2940, 2890, 2850, and 2800 cm^{-1}) are prominent in the Orion Bar region. The 3.40 μm and 3.51 μm features increase in intensity relative to the dominant 3.29 μm (3040 cm^{-1}) feature when going from the ionized to the neutral zone across the Orion Bar. The spectrum obtained in the Red Rectangle, 5″ north of HD 44179, is somewhat similar to those in the Orion Bar. However, only a weak and rather broad 3.40 μm feature is present at the position of HD 44179. These spectra demonstrate that some of the 3 μm emission components vary independently of each other and in a systematic way within UV-excited nebulae. This spatial variation is discussed in terms of the UV excitation and photochemical evolution of polycyclic aromatic hydrocarbons and related molecular structures. The spatial behavior of the weak emission features can be understood qualitatively in terms of hot bands of the CH stretch and overtones and combination bands of other fundamental vibrations in simple PAHs. An explanation in terms of emission by molecular sidegroups attached to the PAHs is less straightforward, particularly in the case of the Red Rectangle and other evolved, mass-losing objects. We estimate PAH sizes of 20–50 carbon atoms based on the susceptibility of PAHs to destruction by the far ultraviolet fields present in the Orion Bar and the Red Rectangle; the size range is similar to independent estimates made previously.

Subject headings: infrared: spectra — interstellar: molecules — nebulae: individual (Red Rectangle) — nebulae: Orion Nebula

I. INTRODUCTION

Recent observations of the 3 μm unidentified emission features found in many UV-excited nebulae have revealed that their spectra are considerably more complex than had earlier been appreciated. Although the dominant feature is the well-known 3.29 μm (3040 cm^{-1}) band, other features at 3.40, 3.46, 3.51, and 3.57 μm (2940, 2890, 2850, and 2800 cm^{-1}) are present in some objects and, in addition, the above features appear to be perched on a rounded "plateau" of emission, above the flat continuum, which extends approximately from 3.2 to 3.6 μm (3100 to 2800 cm^{-1}) (Geballe *et al.* 1985; de Muizon *et al.* 1986; Nagata *et al.* 1988). These features have been attributed to the vibrations of polycyclic aromatic hydrocarbons (PAHs) and to related species (e.g., Duley and Williams 1981; Leger and Puget 1984; Allamandola, Tielens, and Barker 1985; Sakata *et al.* 1984; de Muizon *et al.* 1986; Borghesi, Busoletti, and Colangeli 1987). However, precise molecular identifications of the emitting species are unknown.

Clues to the origin of the unidentified 3 μm emission features might be found, not only by accurate determination of their wavelengths and profiles or by studies of their intensity variations from source to source, as have been done previously (e.g., Geballe *et al.* 1985), but also by studying their spatial variations within extended sources. The most basic of the latter approaches is to study the dependence of feature strengths with distance from the star which pumps the emission. We have taken this last approach by obtaining spectra from 3.1 to 3.7 μm (3225 to 2700 cm^{-1}) at three locations in the Orion Bar which are located at different distances from the Trapezium stars, and by observing the Red Rectangle at two locations, one on HD 44179 itself and the other 5″ north of that star. At most of these locations fairly accurate UV fluxes can be estimated. Knowledge of the spectrum and flux of the ultraviolet radiation field is crucial to proper understanding of the physical conditions experienced by the emitting material and can be used to place constraints on the structure of the emitters. This information, in turn, sheds light on the spatial, chemical, and physical evolution of the band carriers and their environment.

II. OBSERVATIONS

All spectra were obtained at the United Kingdom 3.8 m infrared telescope on Mauna Kea, using the facility liquid and solid nitrogen–cooled grating spectrometer with a 5″ aperture. A spectrum of the Red Rectangle, which consists of 12 individual scans, was obtained in 1984 December, 5″ north of the central star HD 44179. A spectrum obtained with the aperture centered on the star already has been published (Geballe *et al.* 1985); however, a second such spectrum, obtained in 1987 January with the aperture centered on the star, is presented here. Total integration times were 18 and 1.1 minutes, respectively. During both of these measurements a chopper throw of ~40″ (EW) was used. Most of the Orion Bar measurements were obtained in 1986 December. The locations observed were position 4 (Aitken *et al.* 1979) and 10″ and 20″ south of this position. The number of independent spectra and the total integration times at these positions were 4, 6, and 7 and 15, 22, and 26 minutes, respectively. An additional 16 spectra at the 20″ south position were obtained in 1987 August and brought the total amount of observing time at this position to about

[1] Joint Astronomy Centre, Hilo, Hawaii.
[2] Foundation for Astronomical Research in the Netherlands (ASTRON).
[3] NASA/Ames Research Center.
[4] Space Sciences Laboratory, University of California (Berkeley).
[5] Royal Observatory Edinburgh.

TABLE 1

PEAK INTENSITIES OF EMISSION FEATURES

SOURCE	PEAK SPECTRAL INTENSITY OF FEATURE (10^{-17} W cm^{-2} μm^{-1})[a]				
	λ (μm) = 3.29	3.40	3.46	3.51	3.57
	ν (cm^{-1}) = 3040	2940	2890	2850	2800
Orion Position 4	3.9	0.50	0.20	0.17	0.10
Orion Position 4 (10″ S)	2.0	0.35	0.06	0.11	0.05
Orion Position 4 (20″ S)	0.50	0.13	<0.02	0.06	<0.02
Red Rectangle (HD 44179) ...	160.	10.	<5.	<5.	...
Red Rectangle (5″ N)	4.5	0.6	<0.2	0.2	<0.15

[a] Above continuum plus plateau for Orion Bar; above adjacent flux density levels for Red Rectangle.

2 hr. During all Orion Bar observations the chopper throw was 120″ (EW), so that the reference beam, although well within the ionized gas of the Orion Nebula, was off of the extended infrared continuum emission (see Sellgren 1981).

The spectra obtained 5″ north of HD 44179 in 1984 December were sampled every one-half resolution element. All other spectra were sampled every one-third resolution element. The spectral resolution was $\sim$0.0075 μm ($\sim$6.5 cm^{-1}); however, the spectra as shown here are Hanning-smoothed and have a resolution of $\sim$0.009 μm ($\sim$8 cm^{-1}).

The co-added spectra were divided by spectra of early-type stars observed during the same nights and were then multiplied by Planck functions corresponding to assumed fluxes and temperatures of these stars. For the Orion Bar measurements the calibration star was BS 1713 (L = 0.13, T = 10,000 K); for the Red Rectangle the stars were BS 2294 (L = 2.66, T = 18,000 K) and BS 1899 (L = 3.53, T = 25,000 K). The final spectra of the Orion Bar and of the Red Rectangle are displayed in Figures 1 and 2. The poor signal-to-noise ratios near 3.31 μm in these spectra are caused by strong absorption by telluric CH_4. Peak intensities of most of the emission features are given in Table 1.

III. DESCRIPTION OF SPECTRA

a) Orion Bar

The three spectra of the Orion Bar (Fig. 1) show the unidentified emission features superposed on weak continua. The spectrum at position 4 shows the well-known emission peaks at 3.29 and 3.40 μm and, in addition, peaks at 3.46, 3.51, and (marginally above the noise level) 3.57 μm. The latter three features were only recently discovered in IRAS 21282+5050 and in GL 437 by de Muizon *et al.* (1986) and have now also been identified in NGC 7027 (Nagata *et al.* 1988) as well as in the ρ Oph source, WL 16 (T. R. Geballe, unpublished). The "plateau" of emission extending from 3.2 to 3.6 μm is also evident in all three Orion Bar Spectra; such emission has also been observed previously in several sources (Geballe *et al.* 1985; de Muizon, d'Hendecourt, and Geballe 1987). The emission features in the Orion Bar decrease in intensity to the south of position 4. However, it is clear that the 3.40 and 3.51 μm features do not decrease as rapidly as the others. In particular the ratio of peak intensities of the 3.40 and 3.29 μm features is 0.13 at position 4, 0.18 10″ to the south, and 0.26 20″ to the south. The peak intensities of the 3.51 μm feature are 0.04, 0.06, and 0.12 relative to those of the 3.29 μm feature at these positions. Note that the ratios of the integrated intensity of each of the weaker features relative to the 3.29 μm feature is roughly a

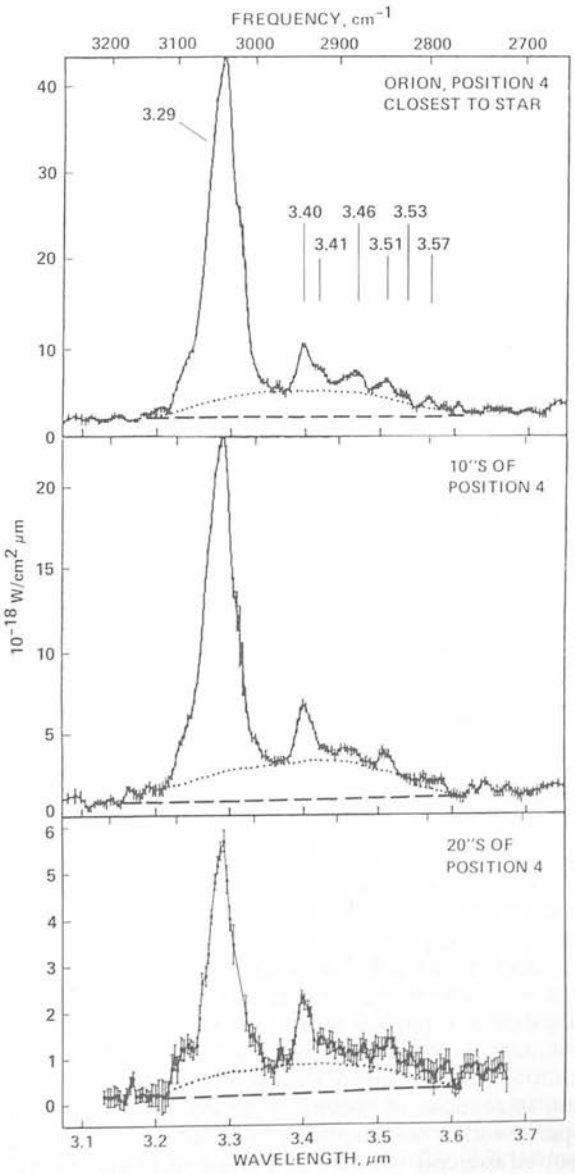


FIG. 1.—Spectra at 3 μm obtained in a 5″ beam at three locations within the Orion Bar. The spectral resolution is 0.009 μm. Dotted lines indicate the "plateau," dashed lines, the underlying continuum. Wavelengths of emission features are indicated in the top panel. Error bars are ± 1 σ.

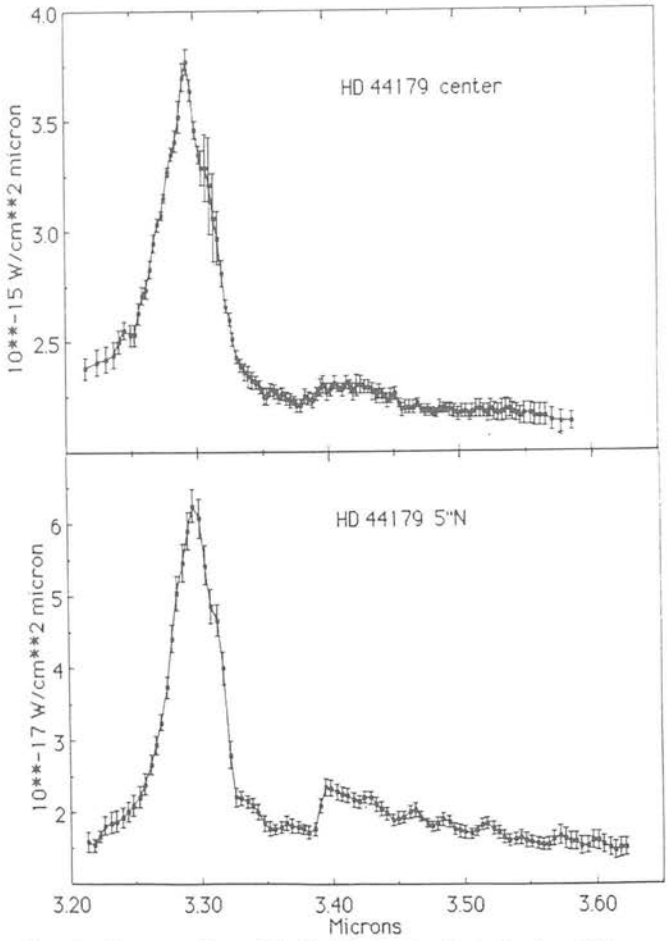


FIG. 2.—Spectra at 3 μm of the Red Rectangle obtained using a 5″ beam at the position of the exciting star and 5″ north of it.

factor of 2 less than the ratio of the peak intensities, because the width of the 3.29 μm emission is about twice that of each weaker feature.

In contrast to the behavior of the 3.40 and 3.51 μm features, the 3.46 μm component appears to be reduced in strength relative to the 3.29 μm feature 10″ south of position 4. At the 20″ south location there is no evidence for this feature at all. It also appears that the long-wavelength shoulders, present on the 3.40 and 3.51 μm features at position 4, decrease to the south. These shoulders may be due to independent and unresolved emission features at 3.415 and 3.53 μm.

b) Red Rectangle

Somewhat similar behavior to that in the Orion Bar is seen in the Red Rectangle spectra (Fig. 2). The spectrum observed at the position of HD 44179 shows a strong 3.29 μm feature and a weak, rather broad 3.40 μm feature superposed on a strong continuum. There is little evidence of the 3.2–3.6 μm plateau, although the 3.40 μm feature appears to be somewhat broader than in most other sources exhibiting it. At the off-star position the continuum, as judged by the flux densities at 3.2 and 3.6 μm, is reduced in spectral intensity by a factor of 150 (compared with a reduction of a factor of $\sim$450 measured for unresolved sources), but the 3.29 μm feature has decreased by a factor of only 36. The ratio of peak intensity (above the continuum) of the 3.40 μm feature to that of the 3.29 μm features has markedly increased, from 0.06 at HD 44179 to 0.13 at the 5″ north position. The short-wavelength cut-on of the 3.40 μm feature appears to occur slightly shortward of the cut-on in Orion, and the peak of the feature shows a similar shift, although these differences are only marginally significant.

In addition to the 3.29 μm and 3.40 μm features, some weak spectral structures appear in the 5″ north spectrum. The weaker ones of these (3.43, 3.48, 3.57 μm) are probably the result of smoothing noise fluctuations in the spectrum. The most prominent of the weak features is at 3.51 μm; we believe that this feature is real. Note that this is one of the features seen to become relatively more prominent going across the Orion Bar, and that it is not seen at HD 44179. The reality of the next most prominent feature, at 3.46 μm, is questionable. Thus regarding the 3.29 μm, 3.40 μm, and 3.51 μm features, the trend is similar to that in the Orion Bar; the ratios of the latter two to the former increase with increasing distance from the stellar (UV) source. We also note that the 3 μm spectrum of HD 44179 obtained by Tokunaga *et al.* (1988) in a 2.″7 aperture is consistent with the above pattern; it shows no 3.40 or 3.51 μm features at all.

The present spectrum obtained at the location of HD 44179 (Fig. 2, *top panel*) shows some slight differences from a similar spectrum published by Geballe *et al.* (1985). Two faint emission features at 3.46 and 3.52 μm were marginally present (although not commented upon) in the earlier spectrum. Because of their proximity to two of the new features found by de Muizon *et al.* (1986), it was suggested by the latter authors that they might be the same ones. However, in the present spectrum of HD 44179 there is no evidence for these emission features. Moreover, it has been determined that spurious features at these wavelengths can be generated by dividing the spectrum of a bright continuum source by that of a late-type star which contains strong OH absorption lines (Tokunaga *et al.* 1988). The present spectrum of HD 44179 was divided by that of a hot star, and should contain no such spurious emission features.

c) Width of the 3.29 Micron Feature

Tokunaga *et al.* (1988) found that within their 2.″7 aperture the FWHM of the 3.29 μm feature in HD 44179 is only 0.023 μm ($\sim$20 cm^{-1}), roughly half that observed in other sources (Geballe 1984; Geballe *et al.* 1985; de Muizon *et al.* 1986; Nagata *et al.* 1988). Our spectra of the Orion Bar and of HD 44179 obtained in a 5″ beam, yield deconvolved FWHMS of $\sim$0.040 μm ($\sim$37 cm^{-1}) and $\sim$0.035 μm ($\sim$32 cm^{-1}), respectively. They were made at a lower spectral resolution than used by Tokunaga *et al.*, but one which easily resolved the feature. A recent (1988 October) spectrum of HD 44179, obtained by us in a 2.″5 aperture, confirms the narrow width found by Tokunaga *et al.*

There are some difficulties in determining the widths of 3.29 μm features, largely due to the deep and only partially resolved telluric CH_4 absorption on the long-wavelength side of the feature and the difficulty in determining the continuum level. These cause uncertainties of perhaps 0.005 μm (5 cm^{-1}) in the FWHMs. In determining the FWHMs given above, we have used a straight line to interpolate across the 3.31–3.32 μm interval. However, the Tokunaga *et al.* (1988) spectrum shows that the major difference between the profiles of the 3.29 μm feature in HD 44179 and in other objects occurs on the short-wavelength side of the 3.29 μm feature, where the atmospheric transmission is fairly good. The results of Tokunaga *et al.*, combined with ours, imply that the width of the 3.29 μm

feature can vary significantly, not only from source to source, but also with position in a source.

IV. DISCUSSION

The 3.29 μm feature has long been ascribed to emission from the C—H stretch in aromatic systems. Because of its high intensity relative to the lower frequency emission features, even far from the powering source, the carriers have to be small, containing 20–50 carbon atoms. Such molecular-sized aromatic entities are generically called polycyclic aromatic hydrocarbons (see Allamandola, Tielens, and Barker 1987 for a review).

Detailed spectra of the emission features in the 3 μm region should provide strong constraints on the carriers once laboratory spectra of realistic analogs become available. Within the framework of the PAH hypothesis, two general classes of interpretations for the weaker features (at 3.40, 3.46, 3.51, and 3.57 μm) already have been suggested. One is that some of these features are vibrational "hot bands" (e.g., $v = 2$–1 and 3–2) of the C—H stretch, blending with combination bands involving the low frequency C—C stretching and bending modes (Barker, Allamandola, and Tielens 1987). The other is that they are fundamental vibrations of molecular subgroups attached to the basic PAH structure (Duley and Williams 1981; de Muizon *et al.* 1986). Below we discuss the present spectra, which show positional variations in the relative strengths of these emission features, in the context of the above interpretations. At present each explanation can accommodate at least some of these observations, mainly because of the lack of laboratory data regarding these transitions in free PAHs. Of course, both explanations may actually contribute. Each explanation will be discussed in turn, within the context of what is currently known.

a) Vibrational Overtones and Hot Bands

In PAHs the fundamental C—H stretching vibration always occurs close to 3.29 μm. Anharmonicitiy of the vibrational potential as a function of atomic separation causes the transitions between higher adjacent vibrational levels (i.e., $v = 2$–1, $v = 3$–2, ...; often called "hot bands") to be shifted to successively longer wavelengths (lower frequencies). Two of the weak emission features observed in these and other objects, those at 3.40 and 3.51 μm, are quite close to, although not exactly at, the wavelengths of the 2–1 and 3–2 vibrational transitions of the C—H stretch in the small aromatic hydrocarbons, benzene, napthalene, and anthracene (Barker, Allamandola, and Tielens 1987). The anharmonicities of ionized and larger PAHs are not yet determined but are likely to be similar. It is possible that some of them might provide an exact match.

Absorption of an ultraviolet (UV) photon by a molecule is followed by a rapid redistribution of the photon energy among the various vibrational modes of the molecule. In the absence of any other deactivation route, this is followed by spontaneous deexcitation via infrared fluorescence. In the case of PAHs the latter process takes about 0.1 s. The PAHs which are believed to dominate the near-infrared and mid-infrared emission spectrum are thought to contain 20–50 carbon atoms (Leger and Puget 1984; Allamandola, Tielens, and Barker 1985). The larger molecules in this range contain many times more vibrational modes than can be simultaneously excited by a single UV ($h\nu < 13.6$ eV) photon. Thus population of high-energy modes such as C—H stretches in $v = 2$ or higher states will be unusual occurrences. However, because smaller PAHs have fewer modes, more frequent excitation above the $v = 1$ levels is expected for them. By the same token the smaller PAHs will also be more susceptible to photodissociation processes by the UV radiation field (Tielens *et al.* 1987).

For a highly vibrationally excited PAH, H loss is much more likely than fragmentation of its carbon skeleton. This is because the aromatic C—H bond strength ($\sim$4.5 eV) is significantly lower than the aromatic C—C bond strength (5.5 eV; cf. Table 2) and because at least two C—C bonds have to be broken simultaneously in order to lose a C atom. This requires 11 eV and a simultaneous bond excitation, a very improbable pair of events. Indeed, photodestruction is thought to play only a minor role in the breaking up of the carbon skeletons of interstellar PAHs (Crawford, Tielens, and Allamandola 1985). Emission of IR photons (IR fluorescence) and C—H bond rupture are competing relaxation mechanisms for highly vibrationally excited PAHs, but C—H bond rupture will dominate only when the amount of energy which is deposited in the molecule exceeds a critical threshold, E_c, whose value depends strongly on the size of the PAH and the internal excitation energy (Tielens *et al.* 1987). Small decreases in the size of the molecule or small increases in the internal excitation energy (e.g., the energy of the absorbed UV photon) will increase the probability of H loss by large factors. If reattachment of hydrogen is unlikely before the following UV photon with $E > E_c$ is absorbed, a small PAH eventually will be stripped of all of its H atoms.

Thus in regions with strong UV fields (e.g., at the surface of a molecular cloud or near the exciting star), interstellar PAHs with sizes below a critical value will be quickly dehydrogenated. Larger PAHs will not be dehydrogenated at all, but because these have many vibrational modes, little emission from the $v = 2$ or higher levels will occur, even in the intense UV field near the star. In a shielded environment, such as a molecular cloud (e.g., 20″ S of position 4 in the Orion Bar), or sufficiently far from the exciting star, rehydrogenation dominates over UV destruction, and all PAHs are completely hydrogenated. The smaller ones will be highly excited upon absorption of a UV photon and will emit strongly in the hot bands (i.e., near 3.40 and 3.51 μm). The resultant spectra of the region will be strongly dependent on the details of the PAH size distribution, but clearly, the observed spatial variations in the relative strengths of the 3.29, 3.40, and 3.51 μm bands shown in Figures 1 and 2 could be a natural consequence of this model.

To determine if this explanation is tenable, one must estimate how susceptible small PAHs are to dehydrogenation by UV radiation in the Orion Bar and the Red Rectangle. Within the framework of this model, the intensity ratios of the 3.29 to 3.40 μm features 10″ south of position 4 in the Orion Bar and 5″ north of HD 44179 in the Red Rectangle imply that the smallest hydrogenated PAHs contain about 15–20 carbon atoms, depending on the spectrum of absorbed UV photons (Barker, Allamandola, and Tielens 1987), In contrast, at position 4 and (especially) at HD 44179 the PAHs responsible for the 3.40 μm emission contain about 50% more C atoms. It is likely that this is because of the complete dehydrogenation of the smaller ones at the latter positions. Using quantum RRK theory (Barker 1983) one can calculate the C—H bond rupture rate as a function of PAH size and vibrational energy content. Tielens *et al.* (1987) show that the UV threshold for efficient dehydrogenation of PAHs containing about 18–20 carbon

TABLE 2

PHYSICAL PROPERTIES OF SIDEGROUPS ON POLYCYCLIC AROMATIC HYDROCARBONS

SIDEGROUP[a]	BENZENE DERIVATIVE	BOND[b]	ENERGY[b] (eV)	H STRETCHING MODE[c] (cm^{-1})	(μm)	NOTES
Hydrogen	Benzene	C_6H_5-H	4.47	3030	3.29	
	Benzene	$C-C_5H_6$	11.0	...	...	1
Methyl	Toluene	$C_6H_5CH_2-H$	3.69	2925 s	3.42	2
	Toluene	$C_6H_4CH_2-H$		2870 m	3.48	
	Toluene	$C_6H_5-CH_3$	4.0	...	...	
Ethyl	Ethylbenzene	$C_6H_5CH_2-CH_3$	3.1	2926 s	3.42	
	Ethylbenzene	$C_6H_5CH_2-CH_3$	3.1	2853 m	3.51	
Methylene	Benzyl	$C_6H_5-CH_2$	4.9	2986	3.35	3, 4
	Benzyl	$C_6H_5-CH_2$	4.9	3062	3.27	
	Benzyl	C_6H_5CH-H	4.47	...	...	
Hydroxyl	Phenol	C_6H_5O-H	3.69	3610	2.77	5
	Phenol	C_6H_5-OH	4.5	...	...	
		C_6H_5-O	3.2	...	...	
Amine	Aniline	C_6H_5NH-H	3.47	3481	2.87	
	Aniline	C_6H_5NH-H	3.47	3395	2.95	
	Aniline	$C_6H_5-NH_2$	4.0	...	...	
Aldehydic hydrogen ..	Benzaldehyde	C_6H_5CO-H	3.21	2820	3.55	6
	Benzaldehyde	C_6H_5CO-H	3.21	2730	3.66	

[a] Structural formulae shown in Fig. 3.

[b] Energy of bond denoted by — (Benson 1965)

[c] Peak frequencies and wavelengths of H stretching modes (Bellamy 1966) Variation of about 10 cm^{-1} (0.01 μm) likely within a class of compounds. The indicators s and m denote strong and moderate intensities.

NOTES.—(1) Elimination of a carbon atom from the carbon backbone of a PAH requires two bonds to be broken simultaneously (thus 11.0 eV; Crawford *et al.* 1985). (2) Frequencies are for methyl groups directly attached to a benzene ring. In pure, saturated aliphatic hydrocarbons, the CH stretching frequencies are 2960 and 2870 cm^{-1} (3.38 and 3.48 μm). (3) Phenyl is a very stable radical (see text). (4) Theoretically calculated frequencies (Lutoshkin, Kotorlenko, and Krugylak 1972). Actual values may be slightly different. (5) Frequency and wavelength are for isolated hydroxyl group. If strong intramolecular hydrogen bonding occurs, peak shifts to ~ 3200 cm^{-1} (3.13 μm). (6) Presence of two bands is due to a Fermi resonance of the aldehydic CH stretching mode with overtones of the bending vibrations.

atoms is about 75,000 cm^{-1} (1300 Å). Therefore to estimate the dehydrogenation rates in these objects one must know the UV flux in the 912–1300 Å range.

i) *The Orion Bar*

In the Orion Bar region most of the emission in the features originates in the neutral zone adjoining the H II region (Sellgren 1981; Bregman *et al.* 1988). The UV flux in the above band incident from θ^1 Ori C ($T = 40{,}000$ K) is estimated to be 3×10^{12} cm^{-2} s^{-1}, from the observed far-infrared intensity (cf. Werner *et al.* 1976; Tielens and Hollenbach 1985*b*). Multiplying this by an assumed PAH UV cross section (5×10^{-16} cm^2) yields an H loss rate of 1.5×10^{-3} s^{-1}. Assuming that rehydrogenation occurs with unit efficiency upon each H atom–PAH collision, the rehydrogenation rate is given by the expression, $R = 10^{-9} n$ s^{-1} (Tielens *et al.* 1987). At the Orion Bar, where $n = 10^5$, the rehydrogenation rate is then 10^{-4} s^{-1}. Clearly for such small PAHs, dehydrogenation dominates at position 4. For slightly larger PAHs (~ 25 C atoms), however, the critical UV photon energy for dehydrogenation falls above the Lyman limit, and they will be completely hydrogenated.

The emitting material 10″ and 20″ south of the bar is embedded in the neutral zone where the UV field is heavily attenuated. Assuming that the observed intensity of the 3.29 μm feature directly measures the flux of pumping UV photons, (i.e., that most of the 3.29 μm feature is due to large, completely hydrogenated PAHs), the UV field 20″ south of position 4 is eight times less than at position 4 (i.e., $\tau_{UV} \simeq 2$). The C—H rupture rate is then calculated to be about 2×10^{-4} s^{-1}, comparable to the rehydrogenation rate. If instead the 3.29 μm emission is dominated by smaller, partially dehydrogenated PAHs, then the UV field, and thus the C—H rupture rate, will be lower than estimated above. In either case it appears that dehydrogenation of PAHs close to the surface of a molecular cloud can quantitatively explain the observed spatial variations of the 3 μm spectra in the Orion Bar.

ii) *The Red Rectangle*

A similar analysis of the emission in the Red Rectangle leads to a slightly different model than above. Speckle observations at 2.2 μm have recently been made of this object, and show that its $1/e$ dimensions are 0.″8 × 0.″6 (Beckers *et al.* 1988). Our observations of the line-to-continuum ratio indicate that the 3.29 μm emission is slightly more extended. Therefore we evaluate the physical conditions at a distance of 1″ from HD 44179. For this star, we take the distance (330 pc) and luminosity (1×10^3 $L_\odot$) from Cohen *et al.* (1975), use $T = 10{,}000$ K to estimate that about 0.009 of the luminosity is radiated between 912 and 1300 Å (Kurucz, Peytremann, and Avrett 1974), and obtain a UV photon flux in this band of 6×10^{12} cm^{-2} s^{-1} at a radius of 1″. This flux is twice that at position 4 in Orion and implies that the H loss rate from PAHs containing 20 carbon atoms is $\sim 3 \times 10^{-3}$ s^{-1}. The hydrogen density

is not well known for this object. Hα measurements indicate a density of $\sim 100\ cm^{-3}$ at a distance of 15″ in the lobes (Warren-Smith, Scarrott, and Murdin 1981). Assuming that the nebula is due to a steady outflow from HD 44179, the density at 1″ is $2 \times 10^4\ cm^{-3}$, which translates into a rehydrogenation rate of $2 \times 10^{-5}\ s^{-1}$. Thus, in the region of brightest 3.29 μm emission, small PAHs probably are rapidly dehydrogenated; indeed, conditions appear to be considerably harsher than those in the Orion Bar.

It must be cautioned, however, that the geometry of the Red Rectangle suggests the presence of a bipolar flow; hence, the emission close to HD 44179 might actually originate within a circumstellar disk, whose density cannot easily be guessed. Whether the weakness of the 3.40 μm band close to HD 44179 can be attributed to photochemically driven dehydrogenation depends critically on the assumed density of the emitting circumstellar material.

A further problem is encountered in evaluating the conditions 5″ north of HD 44179. Due to geometric dilution, the UV field will have dropped by a factor of 25 compared to its value at $r = 1''$. In an outflow the density also will have decreased by this factor, so that the ratio of dehydrogenation and rehydrogenation rates should remain approximately constant within the nebula, until attenuation by dust becomes dominant. As the 3.29 μm feature has decreased by only a factor of 36 at the 5″ north position, significant attenuation of UV radiation by dust appears unlikely (i.e., $\tau_{UV} \sim 0.4$). Thus, unless the density is much higher 5″ north of HD 44179 than expected, the relative spatial variations of the 3 μm emission features in the Red Rectangle cannot be explained by the above simple model.

However, an important, and perhaps crucial, distinction between the conditions close to HD 44179 and those of the Orion Bar region is the intensity of the entire illuminating UV field (evaluated over all UV wavelengths longward of 912 Å). At a distance of 1″ from HD 44179 the absorption rate of all such UV photons by PAHs is about $0.1\ s^{-1}$, which is about 30 times larger than in the Orion Bar. Consequently, events governed by the absorption of two UV photons within a radiative relaxation time scale (i.e., $\Delta t_{abs} < t_{IR} = 0.1$ s) occur at least 10^3 times more frequently in the Red Rectangle than at position 4 in Orion. The increased likelihood of a high internal excitation energy, even larger than the Lyman limit, coupled with the (possible) low rehydrogenation rate, should result in the dehydrogenation of much larger PAHs close to HD 44179 than in the Orion Bar. Since the rate of two-photon events is proportional to r^{-4}, rehydrogenation will quickly dominate beyond a certain distance. Therefore two-photon events may be able to explain the weakness or absence of the 3.40 and 3.51 μm emission features close to HD 44179.

Thus within the framework of the "hot band" interpretation of the 3.40 and 3.51 μm features, the spatial dependence of these features both in Orion and in HD 44179 can be interpreted in terms of photochemically driven dehydrogenation of PAHs. The steep dependence of the dehydrogenation rate on the molecular size (Tielens *et al.* 1987) implies that the smallest PAHs, which are responsible for most of the emission in the 3.40 and 3.51 μm features, contain less than 25 C atoms. In the Red Rectangle, within 1″ of HD 44179, somewhat larger PAHs may also be dehydrogenated. The intensity ratio of the 3.29 μm and 3.40 μm features independently implies that the smallest hydrogenated PAHs present contain 15 to 20 carbon atoms (see, e.g., Barker, Allamandola, and Tielens 1987). The similarity of these size estimates is no coincidence, as both the emission in the hot bands as well as dehydrogenation depend on appreciable population of the higher vibrational levels of the CH stretching mode.

Finally, within this anharmonicity picture, not all of the weaker features in the 3 μm region are due to hot bands of the CH stretching vibrations. Some could be attributed to overtones (i.e., $v = 2$–0) and combination bands of the CC stretching vibrations, whose fundamentals occur in the 5–9 μm region. These would not be affected by dehydrogenation of the small PAHs. The weak 3.46 and 3.53 μm features in Orion may be examples of this emission since their spatial variations suggest a different origin than the 3.40 and 3.51 μm features. This idea is supported by laboratory spectra of several PAHs, some of which (i.e., coronene and pyrene) show weak, narrow absorption features at the former wavelengths as well as a broad, weak feature from 3.2 to 3.6 μm (Cyvin and Klaboe, reproduced in Allamandola, Tielens, and Barker 1987). The absence of these features in the spectrum of HD 44179 implies then that the PAH family responsible for the emission in this source is somewhat different from that in Orion. Such a compositional difference is also implied by the strikingly different 5–13 μm spectrum observed from this source (Cohen, Tielens, and Allamandola 1985; Cohen *et al.* 1986). Although the Orion Bar and the Red Rectangle are associated with objects in vastly different evolutionary stages, it is tempting to speculate that such compositional differences are created by the different intensities of the UV fields where most of the emitting PAHs are located.

b) *Molecular Sidegroups*

Figure 3 shows structural formulae of some typical sidegroups that can attach to PAHs. Table 2 lists physical properties of these sidegroups. Included in the table are the XH— stretching frequencies of each sidegroup when attached to benzene (i.e., each benzene derivative). In the sidegroup picture the best candidates for the bands in the 3.4–3.6 μm region are CH stretching vibrations of saturated aliphatic subgroups,

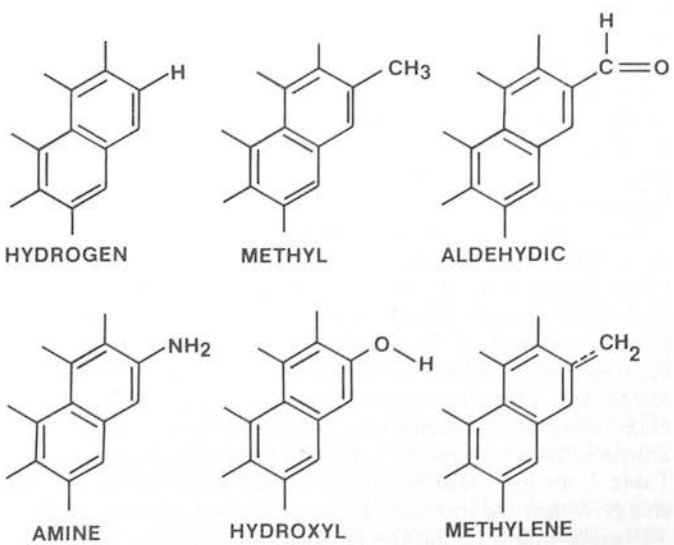


FIG. 3.—Structural formulae of some possible sidegroups on PAHs. The hexagonal networks schematically indicate the carbon backbone of the PAHs. Sidegroups, each of which would replace one H atom, are indicated only on one of the possible edge positions of a PAH. The "methylene" group actually refers to a relatively stable radical (see text).

such as $-CH_3$ and $-C_2H_5$ (Duley and Williams 1981; de Muizon *et al.* 1986; de Muizon, d'Hendecourt, and Geballe 1987). Although close, such saturated aliphatic hydrocarbon sidegroups do not provide an exact match to the observed spectral structure in the 3 μm region. However, it has been suggested that for ionized, dehydrogenated, or electronically excited PAHs these bands would be shifted and might provide a better match (de Muizon *et al.* 1986). Finally, one might perhaps expect the presence of sidegroups other than CH_2 and CH_3, notably CH and NH_2. However, due to their low binding energies, such sidegroups are very susceptible to photochemical destruction, and their expected intensities are small.

The present observations, particularly those of the Orion Bar, provide some constraints on sidegroup identifications. The observed 3.40 and 3.51 μm bands might be attributed to (shifted) methyl or methylene CH stretching vibrations. In contrast, since the observed 3.40 and 3.46 μm bands are decoupled, the latter cannot be the (shifted) methyl or methylene symmetric stretching vibration. Indeed, the difference in spatial behaviors of the 3.40 and 3.51 μm features relative to the other weak features (3.415 μm, 3.46 μm, 3.53 μm) is not readily explained within the sidegroup model (see below). However, even within the sidegroup model, $\Delta v = 1$ emission from higher vibrational levels of the CH stretching mode, and overtones and combination bands of the low-frequency CC stretching vibrations are expected at some level.

In the case that several of the weak bands are emitted by different molecular sidegroups, their intensities relative to other bands are a measure of the ratio of aromatic to aliphatic groups attached to the PAHs. The integrated band strengths for gas-phase PAHs with and without sidegroups are not known, but may be similar, as in the case of benzene derivatives (Gribov and Smirnov 1962; Wexler 1967). If so, the observed ratio (5–10) of the band intensity at 3.29 μm to the sum of those at longer wavelengths is simply the ratio of the number of aromatic to aliphatic H atoms in the emitting species. Assuming that each saturated aliphatic group contains three H atoms, the ratio of aromatic H atoms to aliphatic sidegroups is 15–30. This suggests that saturated aliphatic sidegroups are at best only a minor component of the IR-emitting PAHs. Attributing the emission to PAHs containing 20–50 carbon atoms with about 10–20 edge atoms results in an average of one sidegroup (other than H) per PAH, although one might construct models in which sidegroups dominate the peripheral groups on small PAHs and are absent on larger ones (see below).

In the sidegroup model, variations in the relative strengths of the features may be attributed to differences in the degrees to which different sidegroups survive exposure to the UV radiation field and the different rates of sidegroup reattachment after dissociation. It is therefore important to consider the binding energies of the various bonds involved with aliphatic sidegroups. These are summarized in Table 2. In view of the observed spectral structure, we restrict ourselves mainly to the aromatic hydrogen, methyl, and methylene groups. Examining Table 2, we infer that if there were just sufficient vibrational energy within the molecule to break a bond, the first ones to dissociate would be the C—H bond in the methyl side group and the C—C bond which connects methyl groups to a larger saturated aliphatic hydrocarbon sidegroup (i.e., $-CH_2-CH_3$). In both cases a very stable radical is produced because the unpaired electron on the remaining CH_2 sidegroup can participate in the aromatic π molecular orbital. In the case of methyl-benzene, the benzyl radical is formed. Due to the increased binding energy, further hydrogen loss from this stripped-down methyl group is about as likely as the loss of peripheral aromatic hydrocarbons. Thus, photochemistry will tend to convert saturated aliphatic hydrocarbon sidegroups into CH_2 radical sidegroups. Subsequently, the PAH may lose further hydrogens from this sidegroup, peripheral aromatic H atoms, or both, depending on the exact binding energies involved. Finally, loss of the carbon in the sidegroup may become important. As in the case of the anharmonic model, fragmentation of the carbon backbone of the PAH is expected to be of little importance.

i) *A Model for the Orion Bar Emission*

The above considerations suggest the following simple model for the photochemical evolution of PAHs in clouds near sources of intense far-UV radiation, such as in Orion. At a distance within the molecular cloud corresponding to $A_V \sim$ 5–10 mag, a small PAH (< 25 carbon atoms) is likely to absorb a UV photon and lose an aromatic H during the lifetime of the region ($\sim 10^5$ yr for Orion). Because most of the hydrogen is locked up in H_2, abundances of atomic C and O are comparable to that of atomic H, while atomic N is somewhat less abundant (Tielens and Hollenbach 1985*a*). A PAH which has lost an H atom there is likely to grow an aliphatic sidegroup containing O, C, H, and, to a lesser extent, N. The expected sidegroups at this cloud depth may consist of an aliphatic carbon backbone with hydroxyl (e.g., OH), aldehydic (e.g., HC=O), amine (NH_2), and methyl and methylene (e.g., $-CH_3$, $-C_2H_5$) functional groups attached (cf. Fig. 3). Closer to the cloud surface ($A_V < 5$ mag) the UV field will photodissociate H_2 and atomic H will be more abundant than C and O, so the strengths of the sidegroup features relative to the 3.29 μm feature will decrease. In terms of the spatial behavior of the emission features, only the 3.40 and 3.51 μm features are even qualitatively consistent with the above model. However, since for small PAHs the photochemical rupture and rehydrogenation time scales are much shorter than the dynamical evolution of the Orion region, the composition of the sidegroups on PAHs should reflect the composition of the gas phase, and thus within the Orion Bar region observed ($A_V < 1$) only aromatic H groups are expected. One is then forced to assume that the smallest PAHs present in Orion have a time scale for photochemical evolution which is comparable to the lifetime of the region and, thus, will not have "equilibrated" with the gas phase. Possibly, the observed sidegroups might result from the photochemical evolution of PAHs accreted onto ice grain mantles (Allamandola 1988).

In principle, in the sidegroup model, the 3.4–3.6 μm features could originate in slightly larger PAHs than in the anharmonic model, as the average energy per oscillator is not required to be as high for emission in the CH_2 and CH_3 stretching vibrations as that needed to pump $v = 2$ and 3 levels of the aromatic CH stretch. Somewhat larger PAHs are also implicated if the observed variations are interpreted, as above, as implying a time scale of $\sim 10^5$ yr for photochemical evolution. Note that a short lifetime for the emitting PAHs immediately implies emission in CH stretch "hot bands," since a radiation field which destroys PAHs populates excited vibrational levels in the process. Thus, if $v = 1$–0 vibrations in sidegroups are the only cause of the weak 3 μm features, a long photochemical time scale for these PAHs is actually a necessity.

To summarize this discussion, in the Orion Bar the observed spatial intensity pattern of the 3.40 and 3.51 μm features relative to the 3.29 μm (fundamental CH stretch) feature is qualitatively similar to that expected from sidegroups, but the patterns seen for the other combinations of weak features do not match expected behavior. In addition, if the sidegroup model is valid, it appears that the PAHs must acquire sidegroups by some means other than UV photodehydrogenation followed by reactions with gas-phase species.

ii) *The Red Rectangle*

The situation in the Red Rectangle is quite different from that in Orion, because the 3 μm emitting material has been freshly synthesized from atomic gas formerly within HD 44179. In the harsh environment very close to the star, PAHs can only be formed without aliphatic sidegroups. As the PAHs move to greater distances from the star, conditions moderate to the point where sidegroups could survive. However, unless the carbon elemental abundance is extremely high (C/H $\gtrsim$ 1), complete hydrogenation is expected to dominate over sidegroup formation. Thus, we see no simple model that can account for the observed spatial variations in the 3 μm spectrum of the Red Rectangle within the context of the sidegroup model. In fact, this argument may apply to the structure in the 3 μm spectra in all outflow sources, such as planetary nebulae (NGC 7027) and other evolved objects (e.g., IRAS 21282+5050) which produce PAHs locally.

d) *The* ^{13}C—H *Stretch*

When the ^{12}C—H fundamental stretch occurs at 3.29 μm the ^{13}C—H fundamental stretch occurs at 3.42 μm. Thus ^{13}C may be responsible for the weak, long wavelength shoulder sometimes observed on the 3.40 μm feature. In Orion the emitting material is interstellar, where ^{12}C/^{13}C is $\sim$40 (e.g., see Hawkins and Jura 1987), in reasonable agreement with the strength of the 3.415 μm shoulder, after subtraction of the underlying broad component (plateau). In evolved objects mixing of CNO-processed material from the interior can lead to observed values of ^{12}C/^{13}C as low as $\sim$4. The spectrum of HD 44179 by Tokunaga *et al.* (1988), which shows no detectable 3.40 μm or 3.415 μm feature, demonstrates that ^{12}C/^{13}C is at least 50 in that object. CO observations of NGC 7027, another object which has a clear-cut 3.40 μm band, yield ^{12}C/^{13}C > 36 (Thronson 1983). Here, the ^{13}C may be responsible for the observed shoulder on the 3.40 μm emission feature.

e) *Line Widths*

In general the width of the 3.29 μm feature (FWHM $\sim$ 0.04 μm) is about twice that of the weaker features (FWHM $\sim$ 0.025 μm). The latter are somewhat difficult to determine accurately due to their blending with other weak features as well as their placement relative to the strong, underlying plateau, or "quasi continuum." We noted earlier that the 3.29 μm feature obtained with a 2".7 aperture centered on HD 44179 (Tokunaga *et al.* 1988) has a narrower width ($\sim$0.023 μm) than it does in our spectrum obtained with a 5" aperture, or in any other object observed to date. We also note that the 3.40 μm feature observed on HD 44179 (Fig. 2) is either a blend of more than one feature, or is unusually broad. Thus it appears that the widths of some of these features can vary substantially. In this section we will discuss possible interpretations of these observations, although we realize that further observational tests of these effects are warranted.

The width of an IR fluorescent feature is determined by the combined effects of the time-scale for intramolecular vibrational energy transfer (the Heisenberg uncertainty relation, $\Delta E\,\Delta t > \hbar$), intensity enhancement and blending with weaker modes that are close in frequency (i.e., Fermi resonance), and small shifts in frequency of the emitting mode between different molecules within the family of interstellar PAHs (Allamandola, Tielens, and Barker 1985). The first is generally cited for the widths of the narrower observed emission features (e.g., at 3.29 μm and 6.2 μm), while the last may be responsible for the broad 7.7 μm feature and the plateau of emission underneath the 11.3 μm feature (Allamandola, Tielens, and Barker 1985; Cohen, Tielens, and Allamandola 1985).

In the case that intramolecular energy transfer dominates, the width will reflect the density of states at the internal excitation energy and the coupling of these states with the emitting mode. Since the density of such states is a strong function of the excitation energy, one might expect that the widths of "hot bands" are much greater than those of the fundamentals. Likewise, given their similar excitation energies, one might expect that the CH stretching modes of CH_3 and CH_2 have widths similar to that of aromatic H. However, recent elegant studies of the line widths of overtones of the CH stretching modes in benzene (Reddy, Heller, and Berry 1982), as well as of other molecules, have shown that intramolecular energy transfer from a particular mode to the general thermal bath usually takes place through a very limited number of so-called doorway states. Apparently, partly due to near-energy resonance and partly due to the geometric structure of the molecule, the couplings to these states (other fundamental modes which are close in frequency, as well as overtones and combination bands of lower frequency fundamentals) are very strong, leading to a very fast energy transfer ($\Delta t \sim 10^{-13}$ s). The "doorway" states of the particular modes under consideration themselves couple rapidly with a slightly different set of "doorway" states, so that coupling of a particular vibrational mode with all of the other available modes in the thermal bath occurs on a much slower, diffusion time scale. Nevertheless, all modes communicate on a time scale much less than the typical time for IR emission (0.1 s), so that the assumption of statistical distribution of the available energy over all modes, used in calculating IR emission spectra (e.g., Allamandola, Tielens, and Barker 1985) is fully justified. The line width, however, is determined by those "doorway" modes that dominate the energy transfer from the energy level in question, and the observed widths of the 3.29 μm and 6.2 μm features is in general agreement with that expected from such time scales (Allamandola, Tielens, and Barker 1985).

Since the actual line width of a fluorescing mode depends to a large extent on the coupling to a few states, no large increase in line width with vibrational excitation energy is expected. In fact, experimental studies on benzene show that the width of CH stretching modes increases slightly for the first few modes, but decreases again for even higher modes. In view of the absence of laboratory data on astrophysically relevant PAHs, the decrease in line width for "hot bands" or sidegroups is not presently in conflict with the PAH hypothesis.

Although the line width of the 3.29 μm emission feature generally has been interpreted only in terms of energy transfer time scales, the observations of a large variation in the linewidth by Tokunaga *et al.* (1988) suggest that other contributions should be considered. In particular, close to HD 44179 perhaps even the largest PAHs contain only a few H atoms. Since the

3.29 μm line width is probably dominated by interactions with the stretching modes of other peripheral H atoms on the PAH, such an isolation in HD 44179 might lead to a narrower emission feature. Similarly, we note that in benzene a Fermi resonance between the CH stretching vibration and an overtone or combination band of CC stretching modes gives rise to a strong band at 3.25 μm (3080 cm^{-1}), close to that of the fundamental aromatic CH stretch (Herzberg 1945). The absorption spectrum of coronene and, to a lesser extent, pyrene, also show a strong overtone or combination band in this wavelength region (Cyvin and Klaboe, reproduced in Allamandola, Tielens, and Barker 1987). Blending of these two emission features and broadening by intramolecular energy transfer will give rise to an increased width and perhaps a slight shift in frequency of the CH stretching mode. The narrow 3.29 μm feature observed in a small beam around HD 44179 may then reflect a difference in composition of the PAH family responsible for the emission in this source. That is, the PAHs which dominate the emission within 1″ of HD 44179 may lack this overtone or combination band.

Such a compositional difference in the PAH family close to HD 44179 has already been suggested on the basis of its 5–8 μm spectrum, which shows a "7.7 μm emission feature" which peaks at much longer wavelengths ($\sim$8.0 μm) than in Orion (Cohen *et al.* 1986; Bregman *et al.* 1988; Allamandola, Tielens, and Barker 1987). This might mean that an overtone or combination band of the "7.7 μm" CC stretch is involved in broadening the 3.29 μm emission feature. In that case, one would expect that other sources that show a large wavelength shift in the peak of the 7.7 μm feature (e.g., CPD $-56°8032$ and He 2–113; Cohen *et al.* 1986) would also show a narrower 3.29 μm feature. The observed variations of this feature in the Red Rectangle would suggest that this Fermi responance is associated with the smallest PAHs. Very close to HD 44179, such species may have lost all of their aromatic H atoms, and, in the absence of the fundamental CH stretch, the intrinsic strengths of the overlapping overtone or combination band would decrease considerably. The 3.29 μm band would then be dominated by the CH stretch in somewhat large PAHs, which do not possess this Fermi resonance. Further laboratory and astronomical studies are needed to test these speculations.

V. CONCLUSION

The spectra presented here provide unambiguous evidence for the presence of several components to the 3 μm emission spectrum of UV-excited nebulae which vary independently of one another within each nebulae. The observed spatial behavior of these features, in conjunction with laboratory studies, suggests that the weak 3.46 and 3.53 μm features may be due to overtones and combination bands of low-frequency CC stretching vibrations. The observed spatial variations of the intensities of the 3.40 and 3.51 μm features relative to that of the 3.29 μm feature can be attributed to the photochemical evolution of the interstellar PAH family within UV-excited nebulae. In the Orion Bar, these two features can be either "hot bands" (v = 2–1, 3–2) of the CH stretch in simple PAHs (containing only aromatic H atoms), or emission in the CH stretching modes of saturated aliphatic subgroups. However, in the Red Rectangle, whose emitting material is mass recently lost from the central star, the same type of spatial variations appear difficult to explain in the context of the sidegroup model. Therefore, if the 3.40 and 3.51 μm features in the Orion Bar region are emitted by the same molecular species as they are in the Red Rectangle, one must seriously question whether sidegroups are involved. Given the paucity of relevant laboratory data and the ingenuity of theoreticians, however, both interpretations of these features still must be considered viable. The observed photochemical evolution of interstellar PAHs implies that the emission carriers are small (between 20 and 50 C atoms, depending on the model adopted). It is significant that estimates of PAH sizes obtained from this work are similar to those previously obtained from other considerations.

Progress in understanding the identifications and emission mechanisms of the various features awaits the availability of laboratory spectra. However, it is also important to quantitatively test further the dependency of the weak emission features on the UV radiation field. Additional observational tests might be carried out on planetary nebulae and other objects that show the 3 μm features. We note, for example that the 3 μm spectra of the planetary nebulae NGC 7027 and BD $+30°3639$ are considerably different (Geballe *et al.* 1985; Nagata *et al.* 1988). In BD $+30°3639$, where the 3.29 μm emission occurs rather close to the exciting star, the emission features in the 3.4–3.6 μm region are weak or absent, similar to HD 44179. In NGC 7027, where the emission occurs at an angular distance from the star perhaps 5 times greater (and hence where the UV radiation field is likely considerably less intense), the 3.4–3.6 μm features are relatively prominent. Spectra at different locations within the above objects might prove illuminating.

We wish to thank the staff of UKIRT for its support of this research. Some of the data described herein were obtained during UKIRT Service observing time. We are indebted to Professor Cyvin of Trondheim University, Norway for pointing out some important spectroscopic characteristics of condensed aromatics and to A. Tokunaga for a number of helpful comments. This work was partially supported by NASA grant 188-41-57 at NASA/Ames.

REFERENCES

Aitken, D. K., Roche, P. F., Spenser, P. M., and Jones, B. 1979, *Astr. Ap.*, **76**, 60.
Allamandola, L. J. 1988, in *Carbon in the Galaxy*, ed. J. Tarter (NASA CP), in press.
Allamandola, L. J., Tielens, A. G. G. M., and Barker, J. R. 1985, *Ap. J. (Letters)*, **290**, L25.
———. 1987, in *Polycyclic Aromatic Hydrocarbons and Astrophysics*, ed. A. Leger *et al.* (Dordrecht: Reidel), p. 255.
Barker, J. R., 1983, *J. Chem. Phys.*, **77**, 301.
Barker, J. R., Allamandola, L. J., and Tielens, A. G. G. M. 1987, *Ap. J. (Letters)*, **315**, L61.
Beckers, J. M., Christou, R. G., Probst, R. G., Ridgway, S. T., and von der Luhe, O. 1988, preprint.
Bellamy, L. J. 1966, *The Infrared Spectra of Complex Molecules* (New York: Wiley).
Benson, S. W. 1965, *J. Chem. Ed.*, **42**, 502.
Borghesi, A., Busoletti, E., and Colangeli, L. 1987, *Ap. J.*, **314**, 422.
Bregman, J., Allamandola, L. J., Tielens, A. G. G. M., Geballe, T. R., and Witteborn, F. C. 1988, *Ap. J.*, submitted.
Cohen, M., *et al.* 1975, *Ap. J.*, **196**, 179.
Cohen, M., Allamandola, L. J., Tielens, A. G. G. M., Bregman, J. D., Simpson, J., Witteborn, F. C., Wooden, D., and Rank, D. 1986, *Ap. J.*, **302**, 737.
Cohen, M., Tielens, A. G. G. M., and Allamandola, L. J. 1985, *Ap. J. (Letters)*, **199**, L93.
Crawford, M. K., Tielens, A. G. G. M., and Allamandola, L. J. 1985, *Ap. J. (Letters)*, **293**, L45.
de Muizon, M., d'Hendecourt, L. B., and Geballe, T. R. 1987, in *Polycyclic Aromatic Hydrocarbons and Astrophysics*, ed. A. Leger *et al.* (Dordrecht: Reidel), p. 287.

de Muizon, M., Geballe, T. R., d'Hendecourt, L. B., and Baas, F. 1986, *Ap. J. (Letters)*, **306**, L105.
Duley, W. W., and Williams, D. A. 1981, *M.N.R.A.S.*, **196**, 269.
Geballe, T. R. 1984, in *Workshop on Laboratory and Observational Infrared Spectra of Interstellar Dust*, ed. R. D. Wolstencroft and J. M. Greenberg (Edinburgh: Royal Observatory), p. 93.
Geballe, T. R., Lacy, J. H., Persson, S. E., McGregor, P. J., and Soifer, B. T. 1985, *Ap. J.*, **292**, 500.
Gribov, L. A., and Smirnov, V. N. 1962, *Soviet Phys.—Uspekhi*, **4**, 919.
Hawkins, I., and Jura, M. 1987, *Ap. J.*, **317**, 374.
Herzberg, G. 1945, *Molecular Spectra and Molecular Structure. II. Infrared and Raman Spectra of Polyatomic Molecules* (Princeton: van Nostrand Co.).
Kurucz, R. L., Peytremann, E., and Avrett, E. H. 1975, *Blanketed Model Atmospheres For Early Type Stars* (Washington: Smithsonian Astrophysical Observatory).
Leger, A., and Puget, J. L. 1984, *Astr. Ap.*, **137**, L5.
Lutoshkin, V. I., Kotorlenko, L. A., and Krugylak, Y. A. 1972, *Teor. Eksp. Khim.*, **8**, 542.
Nagata, T., Tokunaga, A., Sellgren, K., Smith, R. G., Onaka, T., Nakada, Y., and Sakata, A. 1988, *Ap. J.*, **326**, 157.
Reddy, K. V., Heller, D. F., and Berry, M. J. 1982, *J. Chem. Phys.*, **76**, 2814.
Sakata, A., Wada, S., Tanabe, T., and Onaka, T. 1984, *Ap. J. (Letters)*, **287**, L51.
Sellgren, K. 1981, *Ap. J.*, **245**, 138.
Thronson, H. A. 1983, *Ap. J.*, **264**, 599.
Tielens, A. G. G. M., Allamandola, L. J., Barker, J. R., and Cohen, M. 1987, in *Polycyclic Aromatic Hydrocarbons and Astrophysics*, ed. A. Leger *et al.* (Dordrecht: Reidel), p. 273.
Tielens, A. G. G. M., and Hollenbach, D. J. 1985*a*, *Ap. J.*, **291**, 722.
———. 1985*b*, *Ap. J.*, **291**, 747.
Tokunaga, A. T., Nagata, T., Sellgren, K., Smith, R. G., Onaka, T., Nakada, Y., Sakata, A., and Wada, S. 1988, *Ap. J.*, **328**, 709.
Warren-Smith, R. F., Scarrott, S. M., and Murdin, P. 1981, *Nature*, **292**, 317.
Werner, M. W., Gatley, I., Harper, D. A., Becklin, E. E., Loewenstein, R. F., Telesco, C. M., and Thronson, H. A. 1976, *Ap. J.*, **204**, 420.
Wexler, A. S. 1967, *Appl. Spectrosc. Rev.*, **1**, 9.

L. J. ALLAMANDOLA and A. G. G. M. TIELENS: NASA/Ames Research Center, MS 245–6, Moffett Field, CA 94035

P. W. J. L. BRAND and A. MOORHOUSE: Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, Scotland

T. R. GEBALLE: Joint Astronomy Centre, 665 Komohana Street, Hilo, HI 96720

INFRARED IMAGES OF REFLECTION NEBULAE AND ORION'S BAR : FLUORESCENT MOLECULAR HYDROGEN AND THE 3.3 μm FEATURE

M.G. Burton*, A. Moorhouse**, P.W.J.L. Brand**,
P.F. Roche*** and T.R. Geballe****

* NASA Ames Research Center, Moffett Field, California 94035, USA.
** Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, Scotland.
*** Royal Observatory Edinburgh, Blackford Hill, Edinburgh, EH9 3HJ, Scotland.
**** Joint Astronomy Center, 665 Komohana Street, Hilo 96720, USA.

Abstract

Images have been obtained of the (fluorescent) molecular hydrogen 1–0 S(1) line, and of the 3.3 μm emission feature, in Orion's Bar and three reflection nebulae. The emission from these species appears to comes from the same spatial locations in all sources observed. This suggests that the 3.3 μm feature is excited by the same energetic UV–photons which cause the molecular hydrogen to fluoresce.

Observations

We have obtained infrared images of emission from molecular hydrogen and the 3.3 μm feature in the ionisation front Orion's Bar, and in portions of three reflection nebulae, NGC 1333 SVS 3, NGC 2023 and Pars 18. In all these sources ultra–violet radiation is beleived to be responsible for the excitation of the molecular hydrogen. The data was obtained at the UKIRT in January of 1988 using the infrared camera IRCAM with 0.6″ pixel scale. Narrow band (1%) filters were used to image the 3.3 μm emission feature, the molecular hydrogen v=1–0 S(1) line (2.12 μm), and the atomic hydrogen Brackett γ line (2.16 μm). The images shown here were constructed by mosaicing together several overlapping frames (note that East is to the right in the figures). Typical integration times were 5 minutes a frame. Images were also obtained through narow band filters at 2.1 μm and 3.1 μm to assess the

contribution of continuum radiation to the emission. Apart from emission from stars, there was essentially no continuum detected at any of the positions observed.

Contour maps of the H_2 line emission from the two reflection nebulae NGC 1333 SVS 3 and NGC 2023, overlapped on those of the 3.3 μm emission feature, are shown in Figures 1 & 3 (the image for Pars 18 is not shown). In all three nebulae there is a tight correlation between the location of the 3.3 μm feature emission and the fluoresced molecular hydrogen emission. Although the relative proportions of S(1) and 3.3 μm emission vary between the sources, this correlation suggests they arise from the same spatial locations in each source. Brackett γ recombination line emission from ionised gas was not detected in these sources.

In Orion's Bar the 3.3 μm emission feature is clearly defined by a sharp, linear, ridge (Fig. 2). The ridge is parallel to the ionisation front seen in optical recombination lines (Munch & Taylor 1974), but lies behind it. The emission reaches a maximum within 5" of the ionisation front, and then falls off approximately exponentially. The feature is still detected 30" away. There are two components to the H_2 line emission beyond the ionisation front, with the strongest emission located in a layer about 15" away from it. There is 3.3 μm emission in front of and behind this layer.

The Reflection Nebulae

In the reflection nebulae observed, the H_2 and 3.3 μm emission arises in photo-dissociation fronts. Beyond the front there are insufficient UV photons with enough energy to photo-dissociate the molecules. The structure of such fronts has been modelled by several authors (*e.g.* Tielens & Hollenbach 1985, Black & Van Dishoeck 1987). The UV flux can heat, through the photo-electric mechanism, a column of gas near the cloud surface to high temperatures ($\sim$ 100 to 1000 K). UV-pumping can excite the molecular hydrogen, leading to an appreciable amount of vibrationally excited H_2 in the gas. Cooling occurs primarily by the OI 63 μm and CII 158 μm lines. The near-infrared spectrum of NGC 2023 is dominated by emission from high-vibrational states of molecular hydrogen (Gatley *et al.* 1987), demonstrating that the emission is fluorescently excited.

In all of the reflection nebulae observed the 3.3 μm emission arises in almost exactly the same regions as the (fluorescent) H_2 emission. The correlation is particularly tight in NGC 2023. These emission regions form shell-like structures around the exciting stars of the nebulae, with the fluoresced H_2 located at an optical depth

$A_v \sim 1$ from the stars. The relative intensity of the H_2 and 3.3 μm emission does vary between sources, and in the northern portion of NGC 1333 SVS 3, H_2 line emission was not detected. This may just be a sensitivity effect, the S(1) line flux falling below the detection limit. In addition, the extinction to the S(1) line, at 2.1 μm is greater than to the 3.3 μm feature; the variation may result from variable extinction.

The 3.3 μm emitting gas must therefore also lie in the photo-dissociation region with the excited H_2. This strongly suggests that the same UV–photons which excite the molecular hydrogen (with $\lambda = 912 - 1100$Å) can also excite the 3.3 μm emission feature. This is in fact slightly surprising for it is likely that UV photons with wavelengths greater than 1100Å can also excite the 3.3 μm feature, whereas they cannot induce H_2 to fluoresce. This may account for the more extended 3.3 μm emission region in NGC 1333 SVS 3 than the H_2, with lower energy photons penetrating further into the surrounding molecular cloud. However in NGC 2023 and Pars 18 the emission regions appear coincident. If the optical depth is rising rapidly behind the emission shell, then all energetic photons will be effectively stopped and neither H_2 nor 3.3 μm feature will be excited beyond it.

Alternatively the 3.3 μm emitting material may be preferentially formed in the hot, dense photo-dissociation regions. If this were the case, though, we might expect to observe the feature away from the photo-dissociation front if UV photons were not responsible for its excitation. This has not been observed in the reflection nebulae. The absence of Br γ line emission in any of the sources indicates that the excitation of the 3.3 μm emission feature is not related to presence of ionised gas.

Orion's Bar

Orion's Bar appears is a classical example of an edge-on ionisation front. The edge is well defined by the 3.3 μm image of the Bar (Fig. 2), lying just behind the ionisation front. It shows clearly that the 3.3 μm emission feature arises from the neutral region behind the front, as other evidence indicates (*e.g.* Sellgren 1981, Aitken *et al.* 1979). The approximately exponential fall-off of the emission strength with distance from the front is consistent with what would be expected if UV radiation were responsible for the feature's excitation, with the optical depth of the emitting region increasing linearly with distance into the cloud from the source of UV photons.

The H_2 v=1–0 S(1) line emission is more complicated, possibly containing two components. There is in fact diffuse H_2 line emission from all over the mapped

region. The emission may arise from both shocked and fluorescent line emitting regions in the gas (Hayashi *et al.* 1985). The H_2 line emission nearer the front, around the 3.3 μm emission ridge, is probably excited by UV-fluorescence, as in the reflection nebulae. The emission is weaker, relative to the 3.3 μm emission, than the H_2 line emission from the reflection nebulae. This could result from an underabundance of H_2 molecules, which may occur if the grains become so hot that reformation of dissociated hydrogen molecules is inefficient. The carriers of the 3.3 μm feature are then presumably not dissociated to a comparable extent as the hydrogen molecules.

The strongest H_2 line emission observed originates 15" behind the front, in a layer assigned as shocked emission by Hayashi *et al.* There is 3.3 μm emission behind and in front of this layer. If this is indeed a shocked layer, it would indicate that that the shock does not significantly affect the material responsible for the feature emission.

We feel it more likely, however, that the H_2 line emission from this layer arises in a dense photo-dissociation region rather than behind a shock. It is hard to drive a shock wave by an expanding HII region into molecular gas sufficiently fast to excite the molecular hydrogen. The shock front would also be expected to lie adjacent to the ionisation front, rather than be offset 15" from it. If the molecular gas is sufficiently dense ($\geq 10^5 - 10^6$ cm^{-3}), collisions can thermalise the fluoresced molecules before they can radiate. The emission spectrum from such a region can therefore appear similar to that from hot, shocked gas, although UV-fluorescence is responsible for its excitation.

References

Aitken, D.K., Roche, P.F., Spenser, P.M. & Jones B., 1979. *Astron. Astrophys.*, **76**, 60.

Black, J. & Van Dishoeck, E., 1987. *Astrophys. J.*, **322**, 412.

Gatley, I., Hasegawa, T., Suzuki, H., Garden, R., Brand, P., Lightfoot, J., Glencross, W., Okuda, H. & Nagata, T., 1987. *Astrophys. J. (Lett)*, **318**, L73.

Hayashi, M., Hasegawa, T. Gatley, I., Garden, R. & Kaifu N., 1985 *Mon. Not. R. astr. Soc.*, **215**, 31p.

Munch, G. & Taylor, K., 1974. *Astrophys. J. (Lett)*, **192**, L93.

Sellgren, K., 1981. *Astrophys. J.*, **245**, 138.

Tielens, A.G.G.M. & Hollenbach, D.J., 1985. *Astrophys. J.*, **291**, 722.

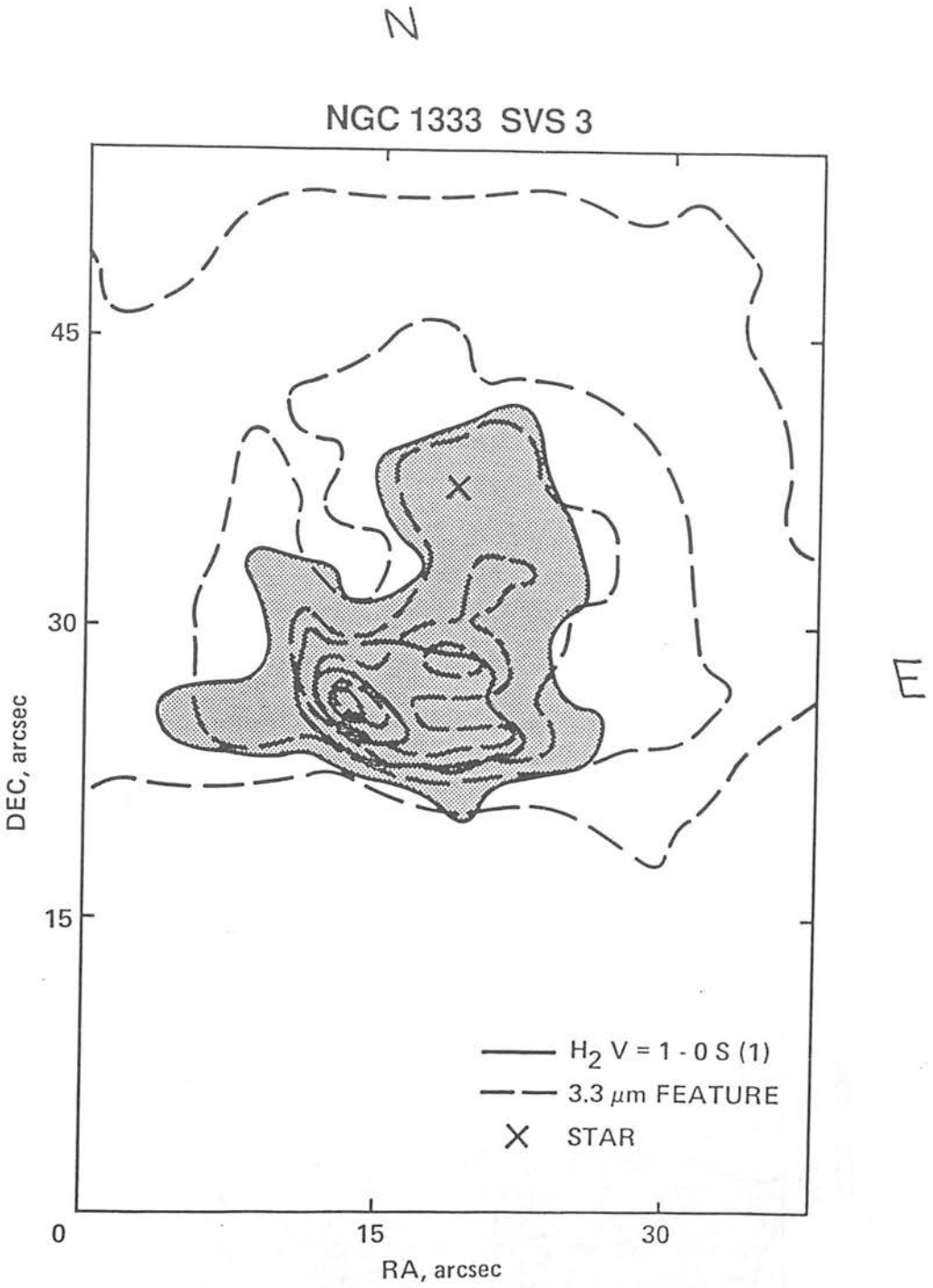


Figure 1. Contour maps of the molecular hydrogen v=1-0 S(1) line (2.12μm) (continuous line, shaded region) and the 3.3μm emission feature (dashed line) in the refelction nebula NGC 1333 SVS 3. The location of the exciting star is marked by the X. The image scale is 0.6" and the map consists of two overlapping frames.

Figure 2.

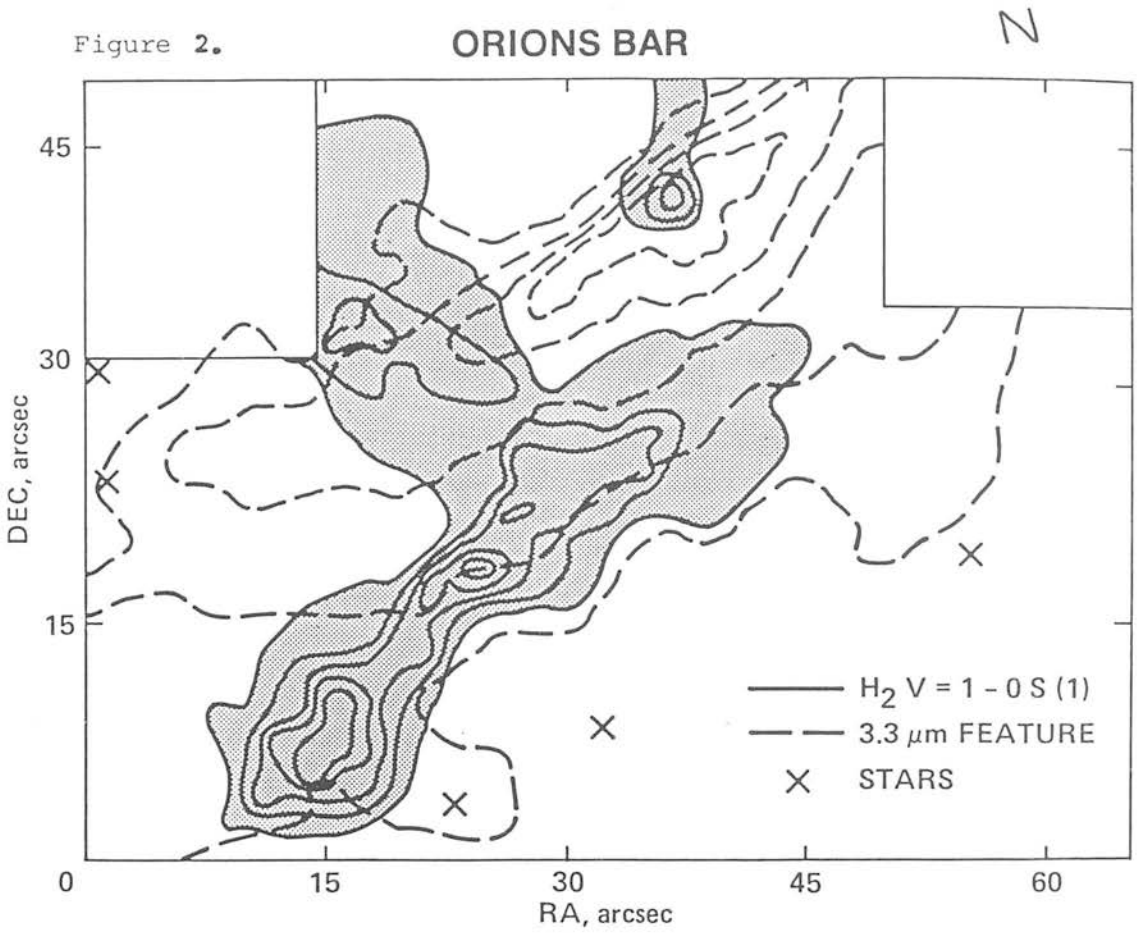
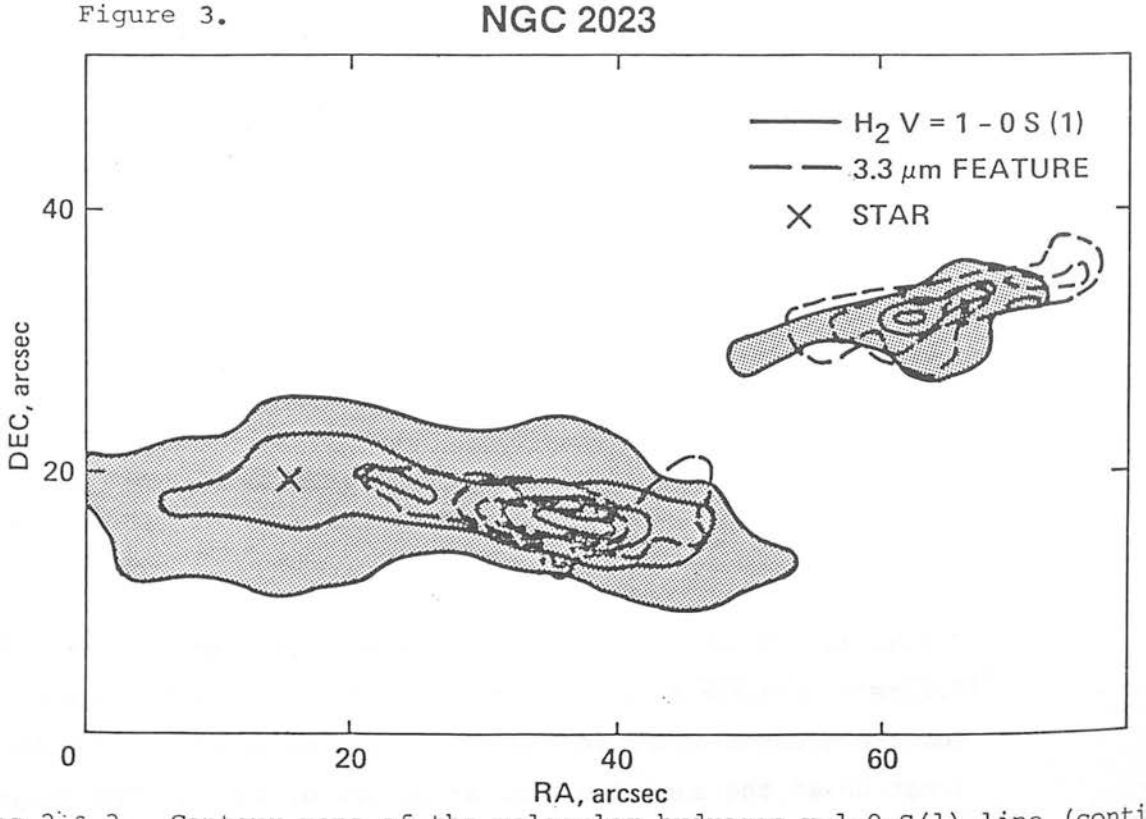


Figure 3.



Figures 2 & 3. Contour maps of the molecular hydrogen v=1-0 S(1) line (continuous line, shaded region) and the 3.3μm feature (dashed line) in Orion's Bar and the reflection nebula NGC 2023. The ionisation front of the HII region of the Bar is to the NW of the ridge of 3.3μm emission. The exciting star for NGC 2023 is located 80" N of the peak of the molecular hydrogen emission.

Mon. Not. R. astr. Soc. (1976) **175**, *Short Communication*, 19P–22P.

'COMETARY' GLOBULES AND THE STRUCTURE OF THE GUM NEBULA

T. G. Hawarden and P. W. J. L. Brand

UK 48-in. Schmidt Telescope Unit, Private Bag, Coonabarabran, NSW, Australia 2857
and
Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ

(Received 1976 February 27)

SUMMARY

Tailed, comet-like objects found in the outskirts of the Gum nebula (and elsewhere) are shown to be $> 10^5$ yr old. This implies that the H II region is much older than the Vela Pulsar 0833−45. Tailed structure found in NGC 5367 suggests the presence of ionized hydrogen falling into the galactic plane near $l = 320°$.

Several elongated, comet-like objects have been noted during inspection of hypersensitized IIIaJ plates taken for the ESO/SRC Southern Sky Survey with the UK 48-in. Schmidt telescope at Siding Spring Observatory. The first two seen were close together and accurately parallel and were for a while suspected to be a new and curious form of plate blemish.

So far 12 of these objects, hereafter termed Cometary Globules (CGs) have been found. They have compact, dusty heads which are almost or completely opaque. Long, faintly luminous tails from 3 to 55 arcmin in apparent length extend from one side of the head while the other often shows a narrow, bright rim (Plate I). Details of these 12 Globules are given in Table I.

Inspection of plates taken with a narrow-band Hα mosaic filter (Elliot & Meaburn 1976) clearly established that the tails and bright rims are emission features.

Ten of the 12 CGs in Table I lie in the outskirts of the Gum nebula in the 'classical H II region' (Bok 1973) defined by γ^2 Velorum, with their heads directed approximately inwards towards the centre of the complex (Fig. 1). Numbers 4, 5, 6, 8, 9 and 10 are associated with the bright rims marking the periphery of the H II region. The non-filamentary appearance of the leading edges of these objects argues against excitation by high energy particles or photons, which would produce a network of intersecting luminescent shocks characteristic of highly-supersonic motion and observed in supernova remnants such as that surrounding PSR 0833−45 in the Gum nebula. The leading edge has all the characteristics of a classical bright rim, with no bow-shock structure apparent. Thus a simple model is suggested based on a subsonic flow of ionized hydrogen in which the dense head acts as a source of radially-moving ionized material. A similar structure would result from the combined radiation pressure from internal sources and a uni-directional overall radiation flux, although one might expect luminescent shocks in this case.

TABLE I

Cometary globules

CG	h m (1975)	° ′	Angular size of head (arcmin)	Length of tail (arcmin)	PA of tail	Remarks
1	07 18·6	−44 32	1 × 2	25	270	V. bright rim, star involved?
2	07 15·3	−43 55	2 × 2	25	270	Thick, faint rim, narrow tail
3	07 38·5	−47 48	3 × 2	7	230	
4	07 33·5	−46 51	1·5 × 4	23	250	Complex head, 5′ across. Position and dimensions for largest blob
5	07 40·0	−43 45	0·6 × 0·8	2·7	275	
6	07 29·7	−46 40	1 × 3	12	230	Tail barely separate from BR and absorption to SW
7	09 13·3	−42 43	1 × 2	19	100	Very thin, bright rim
8	07 41·9	−41 12	0·7 × 1	4	265	Very thin, bright rim
9	07 40·0	−41 23	0·8 × 1·2	3·3	280	Tail joins BR to west
10	07 41·7	−42 01	0·7 × 0·7	6	255	Projected on BR, very faint in H
11	19 38·9	−34 50	1 × 2	12	230	N edge of tail ragged but well defined
12	13 56·2	−39 51	4 × 5	55	135	Star(s) imbedded in head NGC 5367

N.B.: ' BR ' indicates nearby bright rim structure; ' rim ' indicates leading edge of head of globule.

The length of the tail is $L = Ut$ where U is the speed of the background flow and t the time since the flow commenced at the head. A distance of 400 pc (Bok 1975) implies a tail length of > 2 pc and hence, for subsonic flow, an age of more than 10^5 yr. It is likely that in most cases where the globules are completely detached from other nebulosity the detection of the tail is limited by dilution so that this age limit is conservatively low.

Nevertheless this is much larger than the spindown age (a few times 10^4 yr) of the Vela Pulsar PSR 0833 − 45 which has often been assumed to be genetically associated with the vast outskirts of the Gum complex (Brandt 1973). Conversely the globule age estimate tallies with the order-of-magnitude age of a well-developed H II region with bright rims (Osterbrock 1974) providing tentative evidence in favour of Bok's (1973) suggestion (and our opinion) that the Gum nebula is not a single entity.

These tailed globules provide useful probes of the condition and dynamics of the interstellar medium. Density measures in the tail of CG1 on plate SRC-J1315, the known plate characteristic and an assumed sky brightness of 22·0 mag arcsec^{-2} provide an estimate of 10^4 cm^{-6} pc for the emission measure which, if the tail is 0·5 pc thick, implies an electron density of very roughly 10^2 cm^{-3}. Spectrophotometry of these globules is planned in order to test the tentative model and results derived here.

Outside the Gum nebula globules 11 and 12 are the only ones identified thus far, although the ' Bright Dark Nebula ' in Chamaeleon (FitzGerald 1974) referred to in Fig. 1 as 'MPF' may be another example. This latter nebula contains formaldehyde (Goss, private communication). Globule 12 (Plate I(b)) is noteworthy:

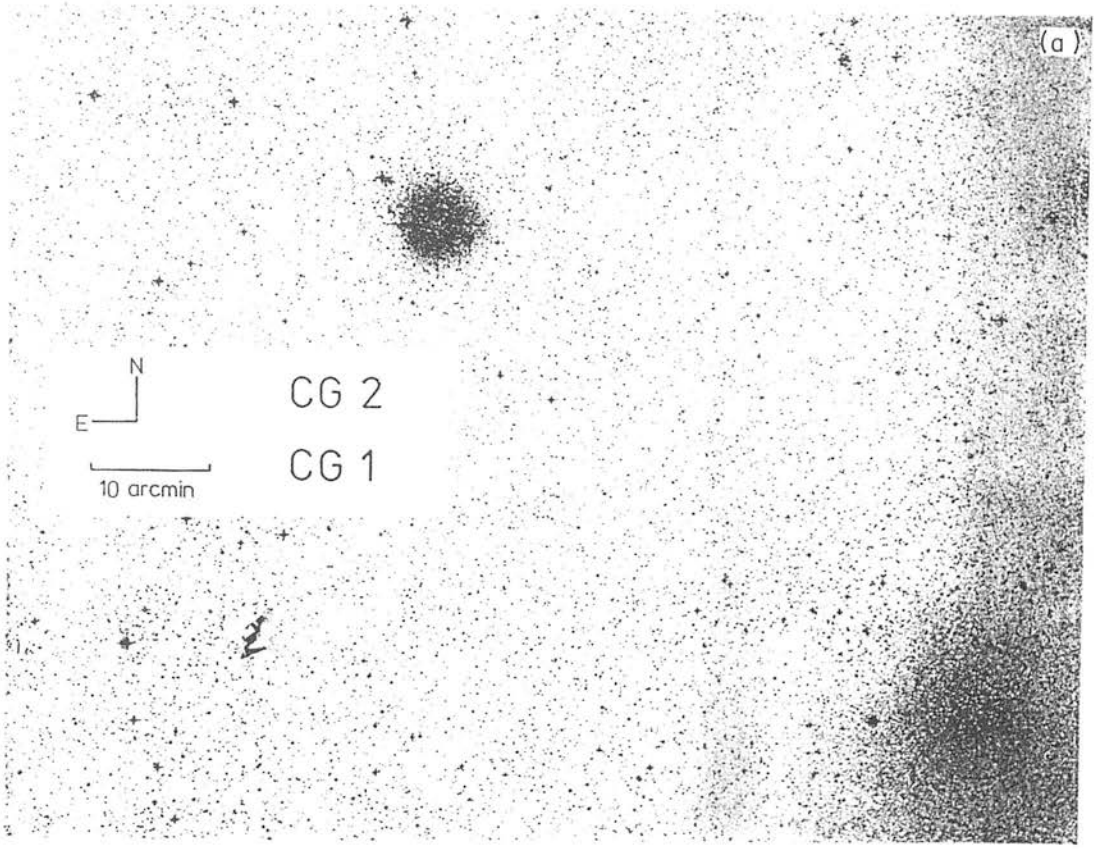


PLATE I. (*a*) *Cometary Globules* 1 *and* 2, (*b*) *Cometary Globule* 12.

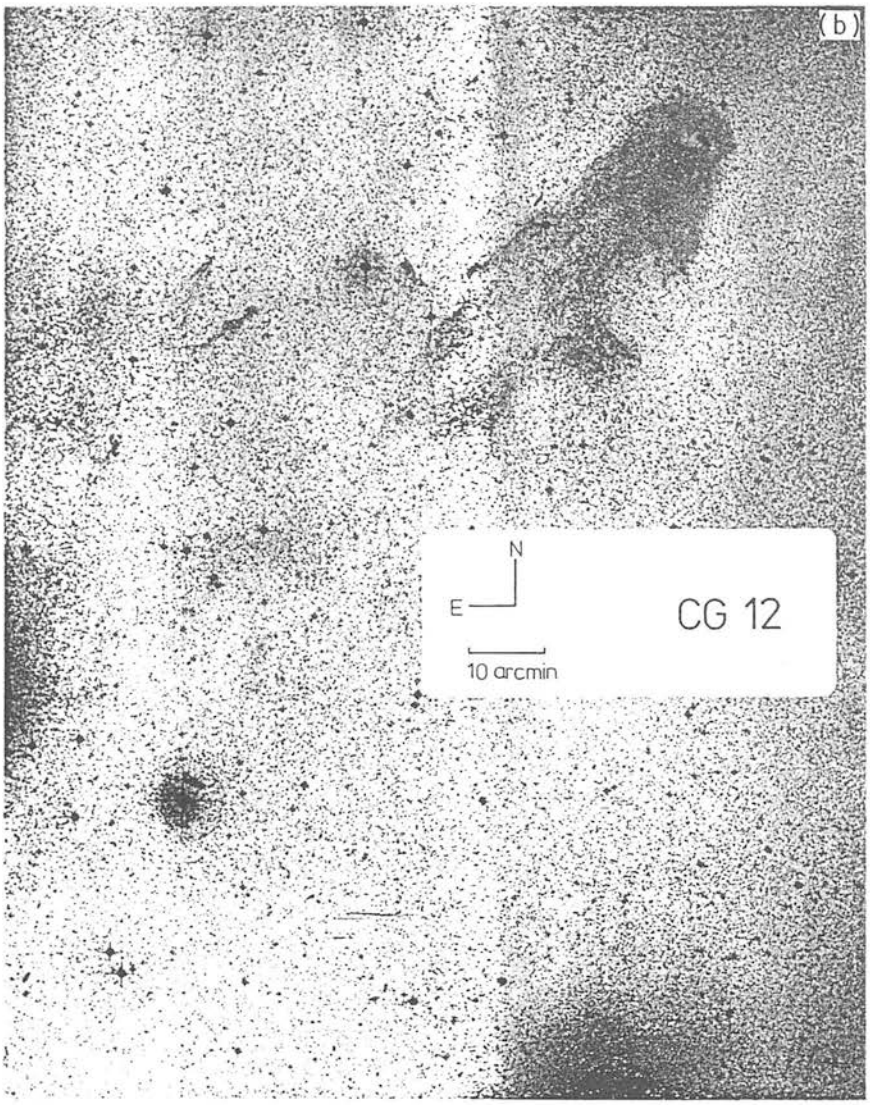


PLATE I(b)

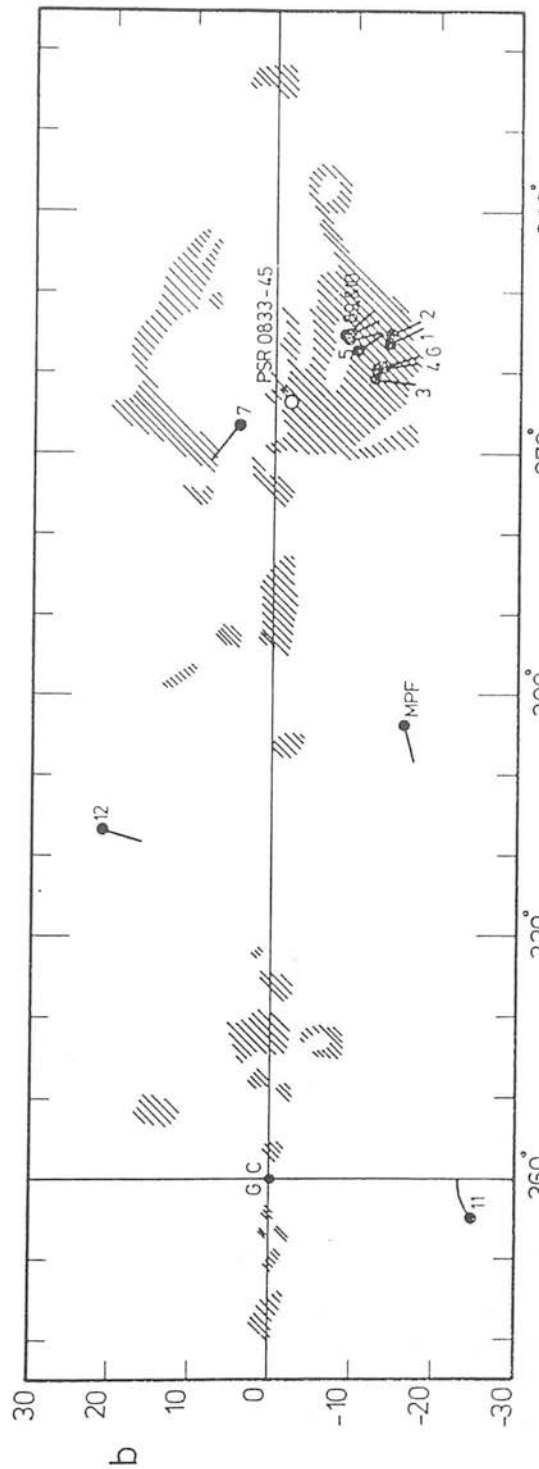


FIG. 1. *The galactic plane with major emission features marked by hatching. The Gum nebula lies between* $l = 240°$ *and* $l = 270°$. *This chart was copied from* '*An Atlas of H-Alpha emission in the Southern Milky Way*' (*Rodgers et al.* 1960).

its head is NGC 5367. The tail in this case is more than a degree long and lies roughly perpendicular to the galactic plane; neighbouring bright rims point in the same general direction. This suggests infall of material towards the galactic plane at $l = 320°$, although it is difficult to explain the fact that the tail rather than the head is pointing towards the plane.

Van Till, Loren & Davis (1975) have mapped carbon monoxide emission from NGC 5367 and from their measurements estimate the mass of the cloud to be 30 solar masses. They suggest on the basis of *UBV* photometry that the embedded stars have been recently formed and that NGC 5367 is an example of star formation in a low mass cloud.

ACKNOWLEDGMENTS

We are grateful to John Meaburn for permission to use the Manchester University Hα interference filter and to Martha Cleary, Don Mathewson and Miller Goss for helpful discussions.

REFERENCES

Bok, B. J., 1973. In *The Gum nebula and related problems*, NASA SP-332.
Brandt, J., 1973. In *The Gum nebula and related problems*, NASA SP-332.
Elliott, K. H. & Meaburn, J., 1976. *Astrophys. Space Sci.*, in press.
FitzGerald, M. P., 1974. *Astr. Astrophys.*, **32**, 465.
Osterbrock, D. E., 1974. *Astrophysics of gaseous nebulae*, Freeman, San Francisco.
Rodgers, A. W., Campbell, C., Whiteoak, J. B., Bailey, H. H. & Hunt, V. O., 1960. *An atlas of H-alpha emission in the southern Milky Way*, Mount Stromlo Observatory, Canberra.
Van Till, H., Loren, R. & Davis, J., 1975. *Astrophys. J.*, **198**, 235.

Mon. Not. R. astr. Soc. (1977) 181, 709–715

The young cluster NGC 5367 and A1353 – 40

P. M. Williams *Royal Observatory, Edinburgh*
P. W. J. L. Brand *Department of Astronomy, University of Edinburgh, Royal Observatory Edinburgh*
A. J. Longmore and T. G. Hawarden *UK 1.2 m Schmidt Telescope Unit of the Royal Observatory, Edinburgh*

Received 1977 March 10

Summary. The sparse young cluster embedded in and illuminating NGC 5367 in the head of Cometary Globule 12 is shown to suffer about 1 mag of visual extinction in the cloud and to be about 630 pc away. Infrared photometry and scanner spectroscopy of the two most luminous members, comprising the double h 4636, indicates that one component is surrounded by a circumstellar shell producing Balmer line emission and thermal re-radiation by grains. We do not favour the proposed identification of NGC 5367 with the transient X-ray source A1353 – 40. We suggest that NGC 5367 lies on an H I loop and that star formation in it was induced by a supernova explosion near $l = 320°$, $b = 30°$ about 10^7 yr ago.

1 Introduction

The intermediate latitude ($b = 21°$) reflection nebula NGC 5367 surrounding the double star h 4636 (Herschel 1847) lies in the head of Cometary Globule 12 (Hawarden & Brand 1976). Its proposed identification (Cooke 1976; Soderblom 1976) with the transient X-ray source A1353 – 40 (sometimes referred to incorrectly as A0353 – 40) has stimulated interest in NGC 5367, particularly as it has been suggested as a low mass centre of star formation by van Till, Loren & Davis (1975), who mapped it in CO. The unexpected possibility of an association between these two phenomena prompted the further observations reported here.

2 Observations

The 'cometary globule' nature of NGC 5367 was first noticed on an ESO/SRC Sky Survey plate taken with the UK Schmidt telescope, with which two further plates of the region have now been taken. One was taken in Hα ($\Delta\lambda = 100$ Å) and the other in broad-band red light on Eastman Kodak 098-04 emulsion through an RG 630 filter. The exposures were for 180

and 30 min respectively and the broad-band plate showed all the nebulosity more strongly than the Hα plate. A careful comparison of the plates showed no evidence for significant Hα emission and it appears that the tail is illuminated by reflection, presumably by light from the galactic plane (*cf.* Sandage 1976). This revises the earlier generalization from Hα observations of some of the globules around the Gum nebula by Hawarden & Brand that the tails of the cometary globules are emission features.

Using the Image Dissector Scanner (Robinson & Wampler 1972) on the Anglo-Australian 3.9-m telescope, on a night of poor transparency but good seeing, spectra covering λλ 3900–7500 were obtained of both components of h 4636 and of the nebulosity in their vicinity. The spectrum of the nebulosity is that expected from reflection of starlight and there is no indication of intrinsic emission from the cloud. The spectrum of the southern component*

Table 1. The *UBV* photometry.

Star	V	$(B-V)$	$(U-B)$	$(B-V)_0$	V_0	Remarks
N	10.70	0.51	−0.19	−0.18	8.60	h 4636 north component
S	10.26	0.20	−0.15	−0.11	9.35	h 4636 south component
1	9.70	0.31	−0.17	–	–	h 4636 = CoD −39°8581
2	10.08	0.13	−0.19	−0.11	9.35	CoD −39° 8583
3	12.45	0.59	0.09	0.41	11.87	
4	9.13	1.02	0.85	1.02	9.13	−39°8586, foreground
5	12.77	0.94	0.62	0.20	10.51	If embedded; see text
6	13.30	0.50	0.14	−0.09	11.50	Bluer case; see text
7	13.31	0.64	0.18	0.34	12.39	
8	12.91	0.68	0.48	0.16	11.32	
9	11.96	0.50	0.08	0.35	11.50	
10	12.11	0.53	0.07	0.38	11.65	

shows mainly the Balmer lines from Hα to Hε in absorption and we estimate a late B spectral type, which is compatible with the intrinsic *UBV* colours derived below (*cf.* Table 1). The spectrum of the northern component has Hα strongly in emission, Hβ almost completely filled in by emission and Hγ to Hε in absorption. The relative strengths of D_3 to Hγ suggest a slightly earlier spectral type than that of the southern component for the underlying star. The northern component is clearly the source of the Hα emission reported by Henize (1976) and Soderblom (1976) although our spectral resolution was insufficient to show the velocity structure in the Hα line reported by Soderblom.

Photometry in *UBV* for 11 stars near NGC 5367 was obtained using the St Andrews' photometer with a 14-arcsec diaphragm on the 1-m Elizabeth reflector now at the SAAO station at Sutherland, Cape. The results are presented in Table 1 and the stars are identified on Plate 1. Apart from the measures of the separate components of h 4636, the results are believed to be accurate to 0.01 mag. Comparison with the *UBV* photometry of stars 1 (presumably both components of h 4636) and 2 by van Till, Loren & Davis shows excellent agreement in $(B-V)$, an apparent difference of 0.04 mag in the $(U-B)$ scales and a discrepancy of almost 0.2 mag in the V of star 1. We note that our combined light V magnitude is compatible with those of the two components and that their use of a larger (40 arcsec) focal plane diaphragm than our own may have made the photometry of van Till *et al.* more susceptible to contamination. Alternatively, of course, one of the components may be a variable.

* We use 'north' and 'south' to distinguish the components of h 4636 instead of *A* and *B* owing to the small magnitude difference and the fact that the position angle has hardly changed since discovery (Innes 1927; Knipe 1963).

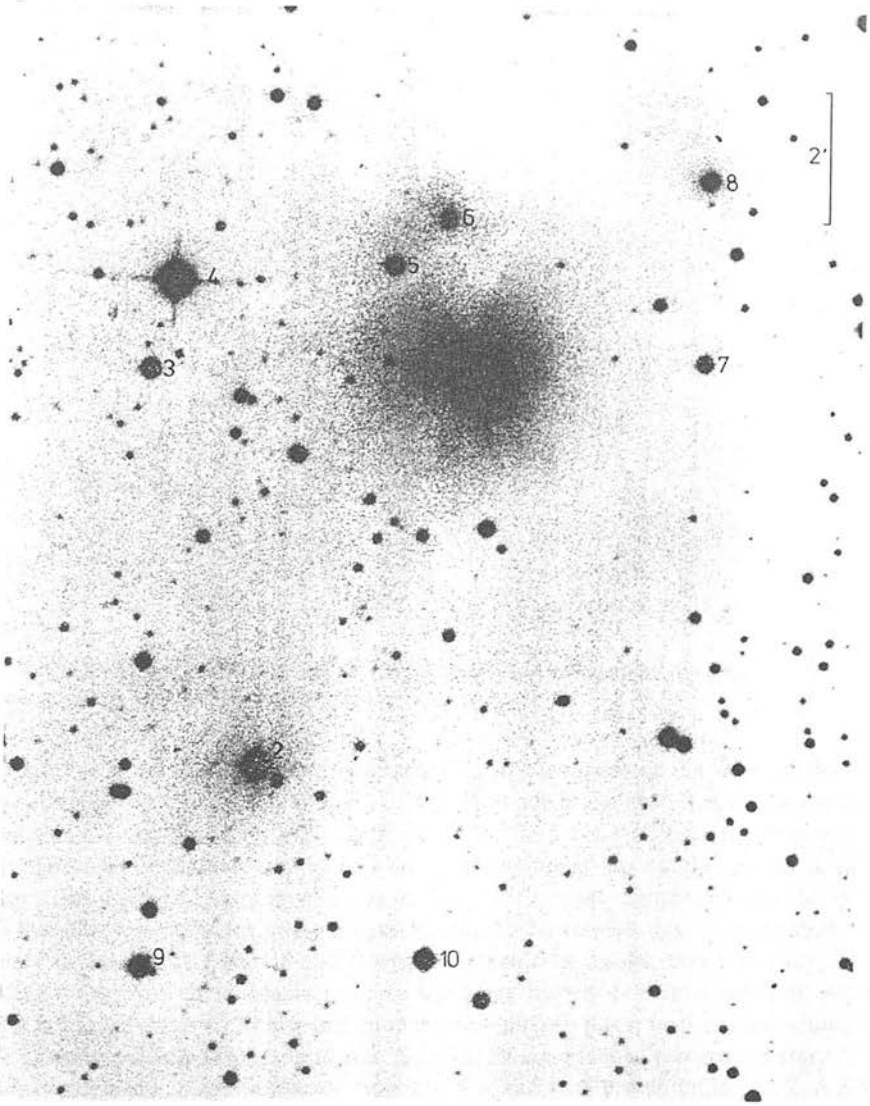


Plate 1. Finding diagram reproduced from the red plate mentioned in the text. The central double, star 1 = h 4636, is in the obscuration below star 6. North is at the top, east to the left. A plate of the whole globule in blue light was reproduced in Hawarden & Brand.

[*facing page* 710]

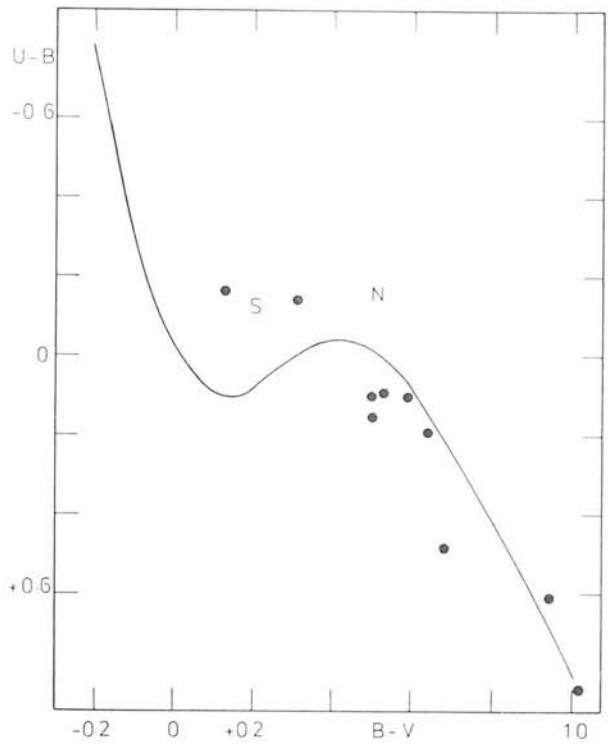


Figure 1. Two-colour diagram for NGC 5367; N and S denote the components of h 4636 and the line the standard relation between $(U-B)$ and $(B-V)$ for unreddened stars.

It is apparent from the two-colour diagram (Fig. 1) that most of the stars are reddened variously. The values of $(B-V)_0$ and V_0, found by tracing the stars back along standard reddening vectors, are also given in Table 1. Since the field as a whole is very clear – on the Schmidt plates faint galaxies are visible in large numbers up to the outskirts of the nebulosity – the reddened stars must be at or beyond the dust complex. The reddening of many of the stars is ambiguous and we have attempted to resolve this by inspection of the plates. Only in the cases of stars 5 and 6 were we unable to resolve this ambiguity and the values of $(B-V)_0$ given in the table refer to the more heavily reddened solution. Alternatively, star 6 has $(B-V)_0 = 0.29$ but still enough reddening to put it in the cloud while star 5 would be a barely-reddened foreground star. Star 4 is an unreddened foreground star.

A colour–magnitude diagram formed from $(B-V)_0$ and V_0 is presented in Fig. 2. A ZAMS with a distance modulus of 9 mag was fitted to the stars most clearly illuminating the nebulosity (Nos 2, 7 and 8). Both components of the double also fall on this ZAMS. If it is not a foreground object, star 5 lies above the main sequence and may still be contracting to it. Both solutions of the reddening of star 6 place it below the ZAMS but this does not necessarily rule our membership of the group as there are examples of very young stars lying below their group main sequence (e.g. W 90 in NGC 2264, Strom *et al.* 1972).

Since its spectrum is quite ordinary, h 4636 S was assumed to have an intrinsic $(V-R)$ colour for a B7 V star, the type corresponding to its intrinsic *UBV* colours (Johnson 1966). This, corrected for the component's reddening, and integrations under the spectra of the two stars yielded $(V-R) = 0.52$ for h 4636 N. This confirmed the result from the *UBV* photometry apparent in the two-colour diagram that h 4636 N suffers more than twice the reddening of its companion. Unless the extinction in NGC 5367 varies by a factor of 2 on a

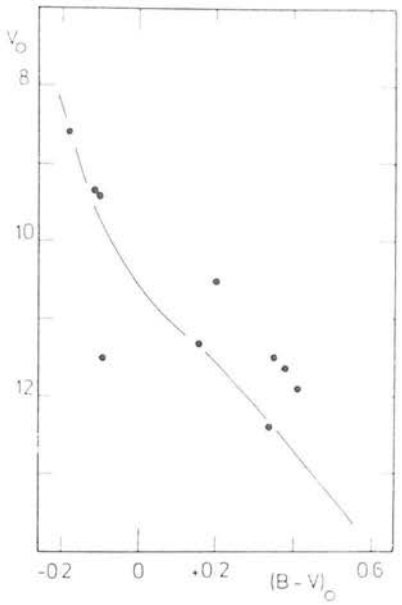


Figure 2. Colour–magnitude diagram for NGC 5367 using the de-reddened data in Table 1 and omitting the foreground star, No. 4. The ZAMS is displaced for a distance modulus of 9 mag.

scale of 4 arcsec, or 2500 AU at the distance derived, the extra reddening of h 4636 N must arise in a circumstellar shell.

Photometry in *JHKL* of the combined light from h 4636 was obtained with the 1-m telescope of the Australian National University, at Siding Spring Observatory. The magnitudes are given in Table 2. It is immediately apparent that the infrared colours are not simply those of a reddened early-type star and that there is an additional source of emission from the double. The flux distribution of this was determined as follows. First, the combined light from the double was corrected for the interstellar reddening associated with h 4636 S. The contribution of the latter was then subtracted assuming it to have normal infrared colours for a B7 V star. The flux distribution of h 4636 N and the infrared source, which we assume to be associated with the north component on account of its Hα emission and circumstellar extinction, is shown in Fig. 3. The underlying star was assumed to have normal infrared colours for a B4 V star, the type corresponding to its intrinsic *UBVR* colours. Its flux distribution, reddened in accordance with the extra A_v of 1.2 mag, was then subtracted to give that of the excess. These two flux distributions are also shown in Fig. 3 and superimposed on the latter is that of a 1600 K blackbody.

Table 2. Infrared magnitudes of both components of h 4636.

J	*H*	*K*	*L*
8.33 (± 0.1)	7.40 (± 0.02)	6.55 (± 0.02)	5.62 (± 0.1)

3 Discussion

The spectrum of the infrared excess from h 4636 appears to be that of thermal re-radiation by the dust grains presumably responsible for the extra visual extinction around the north

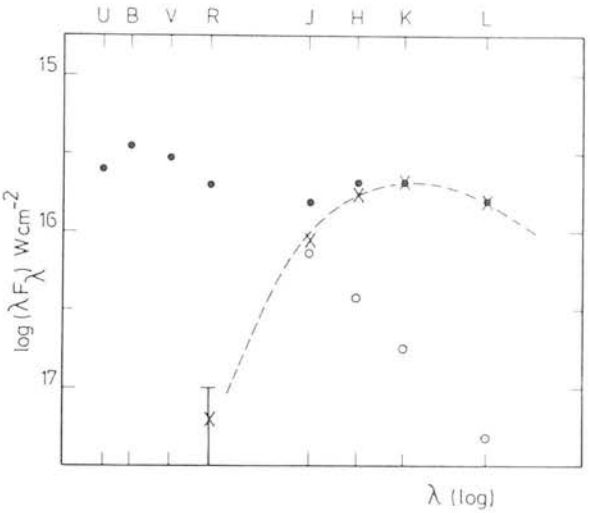


Figure 3. Energy distributions of h 4636 N. Filled circles denote the observed fluxes corrected for interstellar reddening and, in the infrared, the contribution from h 4636 S; open circles the assumed contribution of the underlying star corrected for the circumstellar reddening; and crosses the difference between these two distributions. The point at R is very uncertain being the difference of two almost equal quantities. The broken line represents the flux distribution of a 1600 K blackbody.

component. If the shell does emit as a 1600 K blackbody, this implies either (i) that the grains themselves are at this temperature, restricting their composition to graphite as other grain materials are vaporized at this temperature, or (ii) that the shell's optical depth in the near infrared is low enough for the wavelength dependence of the grains' emissivity to be important. The optical depth in the visible ($\tau_V \approx 0.92\, A_V$) is very nearly unity but the optical properties of circumstellar grains are not well known. If they have the properties of graphite considered by Apruzese (1976), the observed flux distribution can be fitted by grains at a temperature of about 1100 K. On the other hand, using Apruzese's data for silicates, we cannot fit the observed flux distribution with silicates at any temperature at which they would remain solid. These results suggest that the circumstellar shell consists mostly of graphite grains. This is surprising in view of the apparent unevolved status of h 4636 N and requires further investigation with more data on the flux distribution and optical constants.

From their positions in the colour–magnitude diagram, the north and south components appear to have masses of about 8 and 4.5 $M_\odot$ respectively. The separation of the double is about 2500 AU so that, if it is a physical binary, the orbital period of h 4636 is in excess of 30 000 yr which is compatible with the observations of its position angle (Innes 1927; Knipe 1963).

Is the transient X-ray source A1353 – 40 associated with NGC 5367? Soderblom (1976) has suggested star 1 (h 4636) as a candidate on the basis of its Hα emission line profile and we note that the mass of h 4636 N is compatible with the Be star model of a transient X-ray source suggested by Maraschi, Treves & van den Heuvel (1976). In this model, the Be star forms a spectroscopic binary system with a neutron star and the X-ray transients are caused by sudden variations in the rate of mass loss by the Be star and accretion by the neutron star. The initial mass of the system would have been higher than that at present owing to mass loss during the first phase of mass exchange from the precursor of the neutron star to the B star. This model is attractive in our case, particularly if the circumstellar shell around h 4636 N can be identified as overspill from the early mass transfer instead of primordial

material. However, the observed intimate association of the cluster with the whole dust complex argues against it because the initially massive primary could have at least partially dispersed the dust by ionization in the earlier stages of its evolution or by disruption in the subsequent supernova explosion. It should also be remarked that the line profiles described by Soderblom in support for the candidacy of h 4636 are not unusual amongst Be stars, even to the velocity structure which resembles that of 48 Lib (*cf.* Underhill 1960).

Alternatively, the X-ray generation may be a flare phenomenon in cool stars, as suggested by Cooke (1976). Current models (e.g. Crannell, McClintock & Moffett 1974) require magnetic fields of the order of 10^4 gauss for this process, implying anchoring gas pressures higher than 10^6 dyne cm^{-2} and stellar masses below $0.5\,M_\odot$ (Allen 1973). These take more than 10^8 yr to evolve to the main sequence and achieve such pressures (Iben 1965). On the other hand, the central double h 4636 cannot be much older than 2×10^7 yr from the main sequence lifetime of its more massive component. Unless star formation in NGC 5367 began almost 10^8 yr before the formation of the central double, the cluster does not appear to contain a suitable candidate for the X-ray flare star.

We conclude that no star in the NGC 5367–CG 12 complex can reasonably be identified with an X-ray source and recall that the central double lies over a quarter of a degree outside the X-ray error box (Cooke 1976).

Our revision of the distance of NGC 5367 to 630 pc leads to an upward revision of the mass of the cloud (van Till *et al.*) to about $120\,M_\odot$. The total mass of the five stars on the main sequence is about $20\,M_\odot$ and, if we allow another 5–$10\,M_\odot$ for other possible members, it is apparent that about one-fifth of the parent cloud has been converted into stars.

What initiated this star formation? One suggestion, which also provides an explanation for the 10 pc tail of the globule, involves a high-latitude supernova remnant expanding into an inhomogeneous medium. Evidence for this comes from the radiophotographs by Heiles (1976). His Fig. 3 shows a shell or loop of H I centred near $l = 320°$, $b = 30°$ in the velocity range -11.6 to -8.5 km/s. Unfortunately, his horizon excluded the area containing CG 12 but it appears that, if it were completed, the loop would cover that area. Moreover, the head of CG 12 points at the centre of the H I loop suggesting that star formation in the cloud was instigated by the compression caused by the shock front. The flow as the shell continued expanding may have drawn out the tail that makes CG 12 a cometary globule. This would make CG 12 similar to the cometary globules around the Gum nebula in the head of at least one of which, CG 1, there is evidence of star formation (Hawarden & Brand).

The distance to the shell is unknown. Its radial velocity lies in the range in which the shell was observed tangentially, -8.5 to -11.6 km/s. If this is attributed entirely to the differential galactic rotation, the shell appears to be between 660 and 900 pc away and slightly further away than CG 12. The outer radius of the shell is about 23° or between 260 and 360 pc. If the postulated supernova remnant behaved similar to Chevalier's (1974) 'snowplow' model of a supernova shell expanding into low density ($n = 0.01$ cm^{-3}, Chevalier's model C) interstellar medium, we can estimate its age from the relation

$$R = 141\, t_6^{0.32} \text{ pc}$$

between radius and time in Myr. This suggests that the shell has been expanding for between 7 and 19 Myr. The outside of the shell would have passed the position of NGC 5367 some 10^5 yr after the explosion implying, if it was this that triggered the star formation, that the stars in NGC 5367 are between 7 and 19 Myr old. The most massive star on the main sequence, h 4636 N, is no older than 20 Myr from its hydrogen-burning lifetime. The least massive star on the main sequence, star 7, has a mass of about $1.5\,M_\odot$ and would have taken

about 18 Myr to get on to the main sequence although 10 Myr would have been long enough to get close to it (Iben 1965). A more accurate location of this star in the M_{bol}, T_{eff} plane and confirmation of its membership of the cluster should give us a better idea of when star formation there started.

4 Conclusions

Although we do not believe it to contain the transient X-ray source A1353 – 40, NGC 5367 is a very interesting example of intermediate latitude star formation apparently supernova-induced about 10–20 Myr ago. Further spectroscopy and infrared photometry of h 4636 N in order to elucidate the nature of its circumstellar shell should prove very rewarding, as should further study of the other pre-main sequence candidates, stars 5 and 6. The expansion of the supernova remnant also provides a natural explanation of the tail of the cometary globule and the dynamics of this require further study.

Acknowledgments

We are grateful to the PATT for the allocation of observing time on the AAT and SAAO telescopes, to the ANU for time on their 1-m at Siding Spring, to the UKSTU for the use of plates and to staff at all these observatories for their help. Also, PWJLB thanks Bill Zealey and Iain Coulson for help in obtaining the scanner observations and the SRC for a supporting grant, AJL thanks Dr A. R. Hyland for assistance with the infrared observations and the SRC for a Research Fellowship, and PMW thanks Dr M. W. Feast for allowing him to use the *UBV* photometer at very short notice and Malcolm Stewart for assistance with those observations.

References

Allen, C. W., 1973. *Astrophysical quantities,* Athlone Press, London.
Apruzese, J. P., 1976. *Astrophys. J.*, **207**, 799.
Chevalier, R. A., 1974. *Astrophys. J.*, **188**, 501.
Cooke, B. A., 1976. *Nature*, **261**, 564.
Crannell, C. J., McClintock, J. E. & Moffett, T. J., 1974. *Nature*, **252**, 659.
Hawarden, T. G. & Brand, P. W. J. L., 1976. *Mon. Not. R. astr. Soc.*, **175**, 19P.
Heiles, C., 1976. *Astrophys. J. Lett.*, **208**, L137.
Henize, K. G., 1976. *Astrophys. J. Suppl.*, **30**, 491.
Herschel, J. F. W., 1847. *Results of astronomical observations made during the years 1834, 5, 6, 7, 8 at the Cape of Good Hope,* Smith, Elder, London.
Iben, I., 1965. *Astrophys. J.*, **141**, 993.
Innes, R. T. A., 1927. *Southern double star catalogue,* Union Observatory, Johannesburg.
Johnson, H. L., 1966. *A. Rev. Astr. Astrophys.*, **4**, 193.
Knipe, G. F. C., 1963. *Republic Obs. Circ.*, **7**, 48.
Marashi, L., Treves, A. & van den Heuvel, E. P. J., 1976. *Nature*, **259**, 292.
Robinson, L. B. & Wampler, E. J., 1972. *Publ. astr. Soc. Pacific*, **84**, 161.
Sandage, A. R., 1976. *Astr. J.*, **81**, 954.
Soderblom, D. R., 1976. *IAU Circ. 2971.*
Strom, S. E., Strom, K. M., Brooke, A. L., Bregman, J. & Yost, J., 1972. *Astrophys. J.*, **171**, 267.
Underhill, A., 1960. *Stellar atmospheres,* p. 411, ed. J. L. Greenstein, University of Chicago Press.
van Till, H., Loren, R. & Davis, J., 1975. *Astrophys. J.*, **198**, 235.

Astron. Astrophys. 63, 345—352 (1978)

ASTRONOMY
AND
ASTROPHYSICS

Dust Clouds in H II Regions. The Dragon in M 8

P. W. J. L. Brand and W. J. Zealey*
University Department of Astronomy, Royal Observatory, Edinburgh, EH9 3HJ, Scotland

Received January 3, 1977

Summary. A small area of M 8 including an elongated dust cloud has been mapped at 8 μm spacings on a UK 48″ Schmidt plate in the blue, using the COSMOS measuring machine. The map has been calibrated to relative intensity, and scans of intensity made across the cloud boundary. The apparently sharp edge is resolved, and after photographic effects are dealt with the structure of the edge is used to provide a best estimate of dust density in the cloud.

Key words: dust clouds — H II regions — isophotometry

1. Introduction

M 8, after M 42 in Orion is optically the brightest H II region. It is situated approximately 1.4 kpc (Miller, 1968) from the Sun, in the Sagittarius Arm, at $l=6°$, $b=1°.2$. The brightest part, NGC 6523, to the west, surrounds the stars 9 Sgr and 36 Her, both of them O stars, and capable of producing the majority of the ionisation observed. The star cluster NGC 6530 to the east contains several very young stars and may also be considered responsible for some of the ionisation. The whole nebula is embedded in dense un-ionised material which is manifest in reduced star density, and which also contains M 20 to the north and Sharpless 32 and others to the west (Fig. 1a).

We note, following Bok and Reilly (1947) that there are many globules and elephant trunk structures in M 8. In particular, to the southeast where the bright rim structure is particularly complex we note many very small globules, which seem to occur along a north/south lane of obscuration.

Several of these globules have been mapped using COSMOS.

Some of the globules are surrounded by bright rims, a fact which becomes particularly apparent after processing the data from COSMOS (Pratt and Martin, 1975). This will be further discussed in a following paper. The region containing the globules is near a lane of obscuration running parallel to and to the east of the main band of absorption separating NGC 6530 and NGC 6523. Within this region is the elephant trunk we refer to as the 'Dragon' (see Fig. 1b), whose head lies at $18^h1^m41^s$,—24°30′14″ (1950 coordinates).

We have examined the structure of this cloud in some detail, using UK 48-inch Schmidt plates, with a view to understanding better the origin of such clouds. The excellent photometric properties of these plates (IIIaJ + GG 395 filter, bandpass from 3900 Å to 5400 Å), and the positional accuracy of the COSMOS machine, make feasible the determination of the structure of the edges of ionisation bounded clouds, such as the Dragon.

2. Resolution of Sharp Edges of Dust Clouds

Two criteria for edge resolution suggest themselves. The hypothesis used is that adjacency effects are not important. Evidence for this comes from two facts. The first, dealt with in Section 5, is that the developer D-19b together with vigorous agitation has a very small 'chemical spread function', (which determines the magnitude of adjacency effects) as determined by Nelson (1971). The second, which will be expanded on in Section 4, is the independence of the effective intensity profile of star images from stellar magnitude, over a range where saturation and diffraction effects are insignificant.

This being the case, the profile of an edge will be the convolution of the measured star image with the unblurred edge profile of the obscuring cloud.

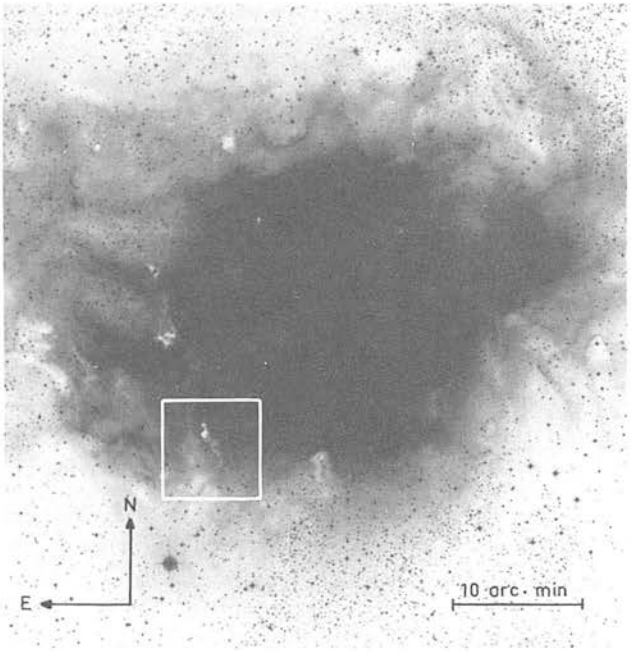
The case of a circular cloud of infinite optical depth will first be considered, with the assumption, borne out below, that star image intensity profiles are gaussian to sufficient accuracy.

Working now entirely in intensity units (effective exposure), suppose the star profile, that is the spread function is

$$S(r)=(1/\pi a^2)\exp(-r^2/a^2) \qquad (1)$$

Send offprint requests to: P. W. J. L. Brand

* W. Zealey is now at the Royal Observatory Edinburgh



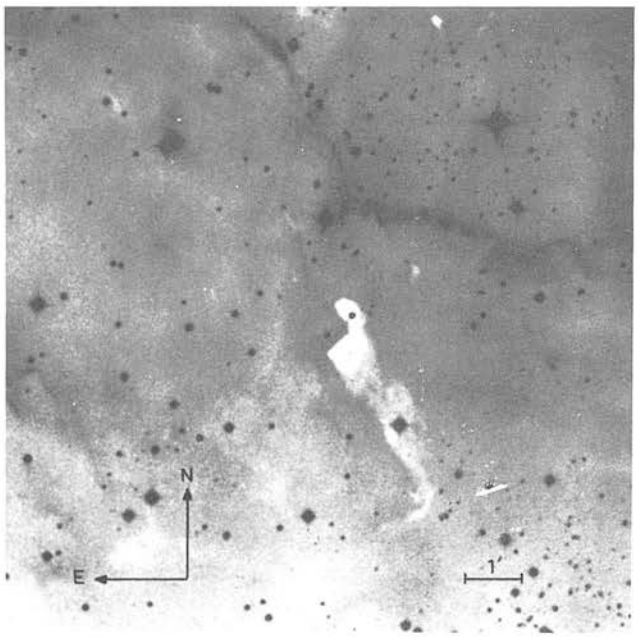


Fig. 1. a The Lagoon Nebula, M 8 indicating the 'Dragon' area of obscuration (U.K.S.T.U. IIIaJ plate). **b** An enlargement of 1a to show the 'Dragon' in more detail

where a will be referred to as the seeing radius, resulting from a convolution of atmospheric, photographic and machine responses. The convolution of this with the 'cloud' profile

$$A_0(r)=1, \quad r \leqq R$$
$$=0 \quad \text{otherwise} \tag{2}$$

is the inverse Fourier transform of the product of the Fourier transforms of these functions, and because of circular symmetry, these are Fourier-Bessel transforms. Use Greek letters to denote the transforms, to get

$$\sigma(k)=(1/2\pi)^2 \exp(-k^2a^2/4)$$
$$\varkappa(k)=RJ_1(kR)/(2\pi k)$$

Putting $S^*A=B(r)$ where * denotes convolution

$$B(r)=2\pi\int_0^\infty \beta(k)J_0(kr)k\,dk \tag{3}$$

which is the Fourier-Bessel transform.

Hence,

$$B(r)=R\int_0^\infty J_1(kR)J_0(kr)\exp(-k^2a^2/4)dk$$

$$\frac{dB}{dr}(r)=-R\int_0^\infty J_1(kR)J_1(kr)\exp(-k^2a^2/4)k\,dk \tag{4}$$
$$=-2(R/a^2)\exp(-[R^2+r^2]/a^2)I_1(2Rr/a^2)$$

In the above, the J_i are Bessel functions, I_1 is a modified Bessel function, and certain properties of these—listed in Erdelyi's Tables of Integral Transforms (Erdelyi, 1954)—have been used in the deduction.

In the limiting cases of $R\to\infty$, $R\to 0$, the profile of a straight edge and of a star are recovered.

The straight edge response is

$$E(x)=[1+\mathrm{erf}(x/a)]/2 \tag{5}$$

Comparing gradients of these profiles, at the maximum gradient points (the half height gradients giving nearly the same answer)

$$\text{Star profile:} \quad -\frac{1}{S_0}\cdot\frac{dS_0}{dr}=0\cdot 86/a$$

$$\text{Edge profile:} \quad -\frac{1}{E_{\max}}\cdot\frac{dE}{dx}=0\cdot 56/a.$$

If the edge profile were broadened for reasons other than seeing, turbidity and machine response, then the ratio of normalized gradients (edge/star) would be less than 0.66.

A corresponding criterion for clouds of finite radius may be deduced from (4). By determining the radius of maximum gradient, and substituting the value in (4), the maximum gradients listed in Table 1 are deduced.

In Table 1, R/a is the ratio of globule radius to seeing radius a defined in Equation (1), and G, the normalised maximum gradient, is given (in units where $a=1$) by

$$G=-\frac{1}{B(0)}\frac{dB}{dr}$$
$$=2R\exp(-r^2)I_1(2Rr)/(\exp(R^2)-1)|_{r=r\,\max} \tag{6}$$

where a further integral property of Bessel functions has been used to deduce

$$B(0)=1-\exp(-R^2/a^2) \tag{7}$$

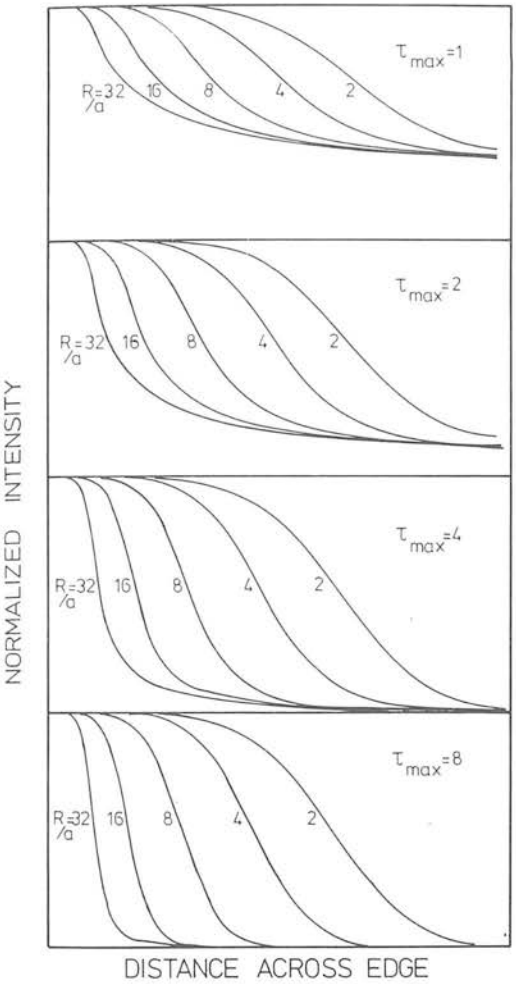


Fig. 2. A series of theoretical profiles through a cylindrical cloud of finite optical depth (Section 3). τ_{max} = max. optical depth; R = cloud radius; a = spread function radius

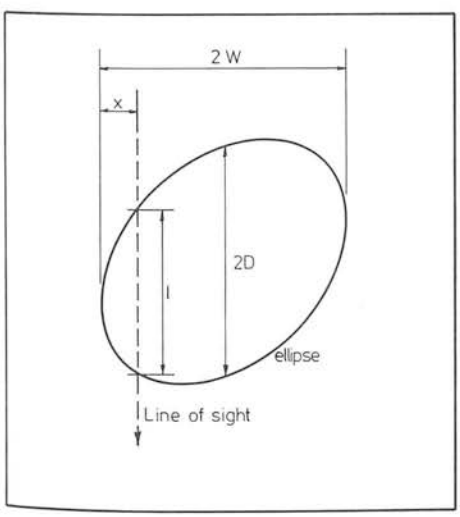


Fig. 3. A cross-section of flattened cloud. $2W$ = projected width of cloud; $2D$ = distance through cloud centre; l = line of sight distance through cloud at a distance X from cloud edge

Table 1. Values of R/a: the ratio of globule radius to seeing radius and G (Section 2, Equation 6) the normalised maximum gradient

R/a	G
0.1	0.856
0.2	0.849
0.5	0.805
0.7	0.761
1.0	0.684
1.2	0.634
1.5	0.585
2.0	0.554
3.0	0.556
5.0	0.561
10.0	0.563

A further criterion for the resolution of an edge may be obtained.

If the edge is not resolved, it is effectively a step function, which is skew-symmetric. A star profile is symmetric, and therefore the convolution of these should be skew-symmetric (that is, a graph of the profile should appear identical when rotated through 180° in its plane). Thus, non-skew-symmetry must mean either that structure in the edge is resolved, or that there is an unresolved symmetric component (line source) at the edge. An example of this might be a bright rim at the edge of a dense cloud. However, it may be shown, by the addition of the error function profile (5) and varying amounts of line profile based on (1), that a bright edge sufficiently bright to influence the gradient measurably, would also cause a readily detectable bump in the edge profile.

Non-skew-symmetry points to edge resolution.

3. A Cloud Model with Finite Optical Depth

If the edge of a cloud is resolved, light must be passing through it. Consider the simplest case of a spherical cloud, or a cylindrical cloud seen perpendicular to its axis, and of uniform density. Suppose it is seen against a bright background, and that scattering into the line of sight is negligible. Then, if the density of dust in the cloud is n_d and the extinction cross-section is σ, the intensity of the background radiation is reduced by a factor $\exp(-\tau)$ where $\tau = n_d \sigma \ell$ is optical depth and ℓ is the chord length through the cloud. Let the cloud radius be R and the distance of the chord from the edge of the cloud by x

$$\ell = 2(x(2R-x))^{1/2}$$

$$I(x) = I_0 \exp(-n_d \sigma \ell)$$

$$= I_0 \exp(-\tau_{max}((x/r)(2-x/r))^{1/2}) \quad (8)$$

where

$$\tau_{max} = 2Rn_d\sigma$$

This gives a very non-skew-symmetric profile. For $\chi \ll R$

$$I(x) = I_0[1-((8R)^{1/2} n_d \sigma) x^{1/2}]$$

Since this profile looks remarkably like some of the measured profiles discussed, a program to convolve this with a stellar spread function (radius a) was run, for a variety of parameters τ_{max} and R/a. Some typical results are shown in Figure 2. It will be seen that a marked lack of symmetry shows for a limited range of these parameters, namely for $R/a > 4$, and for $\tau_{max} < 8$ ($\tau_{max} < 4$ in the case $R/a = 8$).

A limitation on the model is that the radius of curvature of the cloud at the tangential line of sight position may differ from the half width of the cloud. This might be due to small irregularities on the surface. However if smaller than a distance corresponding to the resolution element they will not affect the result, and if much bigger, then contrary to observation they might be expected to affect the straightness of edge of the cloud. The possibility remaining is that the cloud may be flattened or distended along the line of sight, and in this case the parameter determined by the observations is still the optical depth through the centre of the cloud. For suppose the cross-section through the cloud is approximately as in Figure 3, and let the specific intensities due to foreground and background be I_f and I_b respectively; then if I_1 is the specific intensity observed outside the cloud and I_2 is the specific intensity observed at the centre,

$$\mathscr{T} = \frac{I - I_2}{I_1 - I_2} = \frac{e^{-\tau} - e^{-\tau_{max}}}{1 - e^{-\tau_{max}}} \simeq 1 - \frac{\tau}{1 - e^{-\tau_{max}}}$$

where $\tau_{max} = 2n\sigma D$, $\tau = n\sigma\ell$ (see Fig. 3). But from the geometry of the ellipse,

$$\ell = 2D(2X/W)^{1/2} \quad \text{so} \quad \tau = \tau_{max}(2X/W)^{1/2}$$

$$\mathscr{T} = 1 - \frac{\tau_{max}}{1 - e^{-\tau_{max}}} \left(\frac{2X}{W}\right)^{1/2}$$

and for $\tau_{max} > 3$ this result is insensitive to the exact run of density in the centre of the cloud, and moreover produces formally the same relationship as the circular cross-section case. The observational parameter is dust density near the edge of the cloud times dust extinction cross section, times line of sight thickness of the cloud.

4. Star Profiles

To progress further, it is necessary to investigate the spread function, as given by star images. This spread function is the combined spread function due to atmospheric, emulsion and measuring machine effects. Because of saturation effects it is necessary to consider stars sufficiently faint that a one to one relation between density and exposure is obtained. A sample of stars with sufficiently good images in this sense, and in the sense of being unconfused with other star images and on a reasonably uniform background, was selected near the region of the clouds being measured, and elsewhere on the same plate.

The plate in question was measured by COSMOS in mapping mode (for details see Section 6). Star positions were selected, and a 25 by 25 increment map of intensity values at the star image transferred to paper tape (corresponding to 200 μm or 13.″4 across). A program was written to reduce these maps of star images. The corners of the map occupied by right angled isosceles triangles 8 position increments on a side were used to estimate background, represented by a plane which was least squares fitted to the 144 resulting values. This procedure also gives an estimate of the standard deviation per point in the map. The best fit plane is subtracted from the map, and the resulting background-free values coded cyclically 0–9, and the coded map is displayed on a Tektronix VDU attached to the ICL 4130 computer at the Observatory. The background subtracted values were then used to produce a plot of area contained within an isophote of given intensity versus the logarithm of the intensity. Values within two standard deviations of background were ignored.

If a measured star image and hence spread function were gaussian, then the intensity I at a contour containing an area A of the image would be

$$I = I_0 \exp(-A/A_0)$$
$$= I_0 \exp(-N/N_0) \tag{9}$$

where N is the number of measurements with intensity greater than I. A_0 (and hence $N_0 = A/64\,\mu m^2$) is related to the profile radius of Equation (1) by

$$A_0 = \pi a^2$$

So

$$N = N_0 \ln I_0 - N_0 \ln I \tag{10}$$

The histogram of N versus $\ln(1/I)$ is displayed on the Tektronix terminal, and the operator then specifies limits on the graph between which he wishes the data to be fitted by a polynomial. This device enables the operator to eliminate effects due to faint previously undetected stars in the field which show up on the displayed map, and other sources of large scale noise such as a spike which distorts the high intensity end of the histogram.

For a perfectly gaussian profile this plot would be a straight line. The polynomial fit was carried to third order, but except in the case of distorted images, second order would have been sufficient. The profiles were nearly gaussian, with the polyfit second order coefficient always less than 1% of the linear coefficient. The deviation at second order was such that $\log I$ versus A is concave upward, and relates well to work in progress by A.C. Edwards (1975), to whom we are indebted for the (area/log intensity) method.

The linear coefficient was used to calculate the profile radius a for each image. It should be noted that since intensity measurements are relative to the brightest

part of the image, rather than the background, the method is insensitive to background level determination provided, as always was the case, the background does not change by a large fraction over an image diameter.

The mean value for 37 stars on plate C 27 is

$a = 2''.3$

with a standard deviation in this mean of

$\sigma = 0''.2$

The profile radius does not correlate with maximum intensity (star brightness), background intensity, background gradient, or background plus maximum intensity.

The possibility that some of this scatter is due to colour effects is being investigated using a pair of plates.

There is some evidence that the size of the star image depends on the colour of the star, although at this stage it is too early to say much more. Perhaps one third of the scatter about the mean is due to this effect.

5. Adjacency Effects

Adjacency effects in the photographic emulsion may be typified by a chemical spread function applied to the density (not intensity) distribution in the emulsion. Typically, this sharpens edges, in distinction to the optical spread function, which blurs them. Furthermore it is non-linear, the effect being proportional to the ambient density (or more accurately, mass of deposited silver). Nelson (1971) gives an expression for converting the density distribution predicted by using a calibration curve based on large-area density, into the density distribution that would be observed in the presence of adjacency effects. This is

$$D_c(x) = D(x)\,[1 + \beta\{D(x) - D(x) * f(x)\}] \tag{11}$$

where β is the 'strength' of the adjacency effect, and $\beta f(x)$ is the chemical spread function, with $\int_{-\infty}^{\infty} f(x)dx = 1$.

Since we shall show adjacency effects are small, a very approximate relationship between differences in density and exposure may be used, and we will use proportionality. Using the edge profile $E(x)$ of Equation (5) and a chemical spread function

$$f(x) = \exp(-|x/b|)/2b \tag{12}$$

which closely approximates those given by Nelson (loc. cit.), we get

$$\begin{aligned} D_c(x) = D(x)\,[1 &+ (\beta D_0/2) \\ &\cdot\{\exp(-x/b)\,[1 + \operatorname{erf}(x/a - q/2b)] \\ &- \exp(x/b)\,[1 - \operatorname{erf}(x/a + a/2b)]\} \\ &\cdot \exp(a^2/4b^2)] \end{aligned} \tag{13}$$

where D_0 is the density amplitude of the edge. The 'adjacency strength' β is difficult to estimate, but for vigorous developers such as Kodak D-19 which was used to develop the plates it is known to be small, and Nelson (loc. cit.) gives figures indicating $\beta = 0.1$ for *D*-19 used with aerial films. Using this value, and $a = 30\,\mu m$, $b = 30\,\mu m$, it may be seen that the density correction, even for a density contrast D_0 of unity, is never much greater than 10%. Adjacency effects may be ignored, and the measured cloud profiles may be used with the star profiles to ascertain the true profile of the cloud.

6. Measurement of Cloud Profiles

An area 11′ by 22′ surrounding the Dragon within the boundary of M8 was selected for study. A 42 minute, blue, U.K. 48-inch Schmidt plate (C27, IIIaJ + GG395 filter) containing this region was mapped using COSMOS. The adjacent calibration step wedge on the plate was also mapped. In this way the transmission of each element, at 8 μm intervals, was digitised into one of 128 COSMOS '*T*' levels. The resulting data was reduced into a Fortran compatible format for analysis on the ICL 1906*A* of the Atlas Computer Laboratory.

A contour map of the whole area was used to identify regions to which the preceding theory could be applied. The Dragon appears over a 4′ by 5′ area as a dark sinuous lane of obscuration, seen against the slowly varying emission of the H II region. (Fig. 4).

Intensity cross sections, arranged to cut the edges of the obscuring lane along lines of maximum slope, were displayed. These cross sections (Fig. 5) were also arranged to avoid the bright rim seen near the tip of the lane and to cut the edges lying parallel to the longer axis of the lane.

The calibration was carried from a histogram of the number of elements with transmission '*T*' within the image of the step wedge, versus '*T*'. The resultant peaks in this distribution enabled the '*T*' levels for each step to be found and associated with the relative intensity stated for that step. The relationship between '*T*' and intensity was then represented by a third order polynomial.

At this stage areas containing stellar images for profile analysis were also mapped, calibrated and output.

The observed edge profiles are non-skew-symmetric, indicating that the structure of the edge is resolved (Section 2).

7. Fitting the Model to the Observations

For each cross section through the obscuring lane the best fitting curve from the model described in Section 3 was found.

A series of theoretical curves for various $Q = R/a$ and τ_{max} were calculated from Equation (8) convolved with a

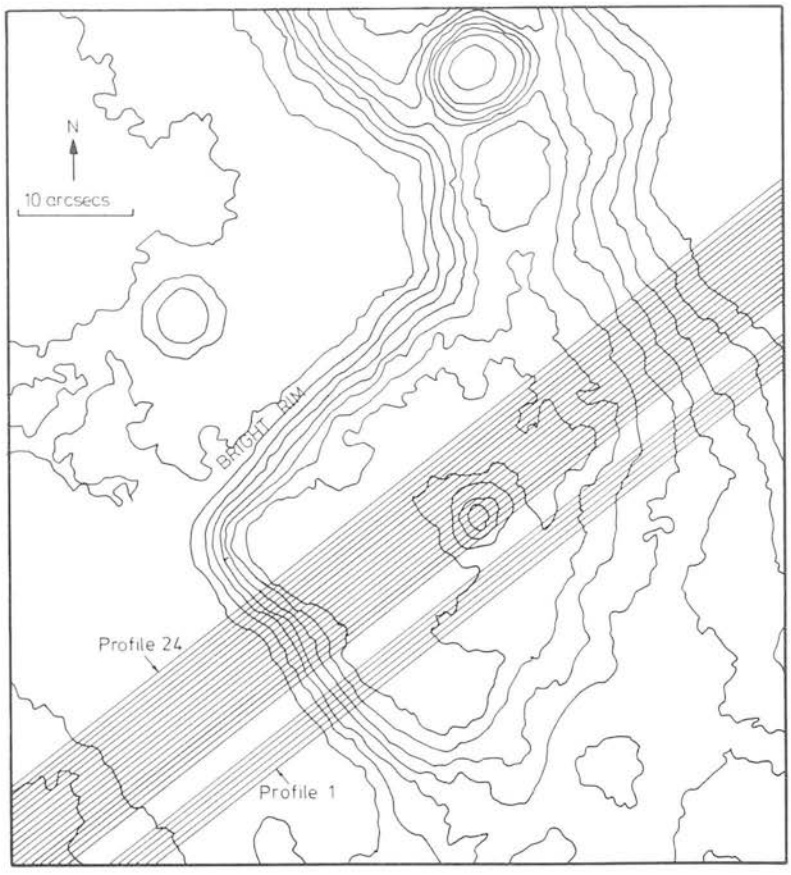


Fig. 4. Isotransmission contours of the 1 mm × 2 mm area of the Dragon as mapped by COSMOS indicating *I*. (i) the bright rim; (ii) the sections used in the analysis

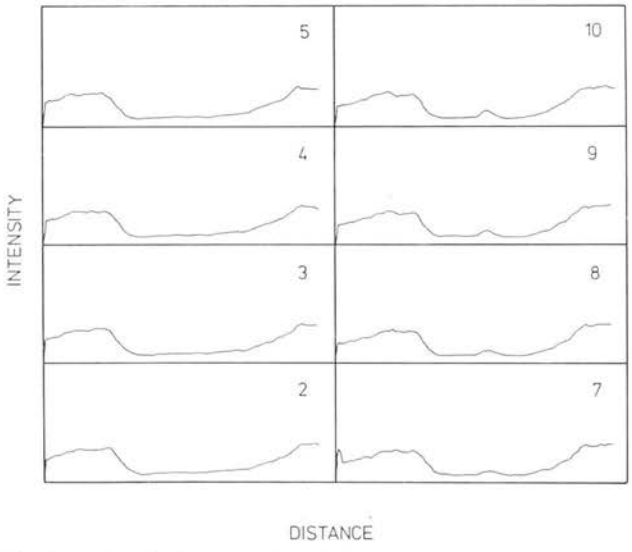


Fig. 5. Sections in intensity through the 'Dragon'

star profile. To each cross section all theoretical profiles of identical Q were compared.

In order to match the shape of the profiles, independent of nebular background variations, both observed and theoretical profiles were normalised. The intensity found in regions of profile unaffected by absorption (Upper Background) was normalised to unity, while the intensity of regions most affected by absorption (Lower Background) was normalised to zero. To counter the effects of slowly varying bright nebular background the Upper Background was linearly interpolated between two points clear of the obscuration and the Lower Background similarly calculated from two points within the obscuration (Fig. 6).

The observed profile was scaled in X to match the scale of the theoretical curve. The offset in X between the observed and theoretical profiles was calculated from the distribution of the difference between the X positions on the observed and theoretical profiles at which a given intensity was attained. The mean of such a distribution is the X offset between the profiles. The standard deviation is then a measure of the similarity between the profiles (Fig. 6).

Using this offset the observed and theoretical profiles were matched in X. The goodness of this fit was estimated from the variance

$$V_1 = \frac{1}{N}\sum (\text{Theoretical Intensity } (X) - \text{Observed Intensity } (X))^2.$$

A third order polynomial was also fitted to the observed profile and a similar variance calculated

$$V_2 = \frac{1}{N}\sum (\text{Polynomial } (X) - \text{Observed Intensity } (X))^2.$$

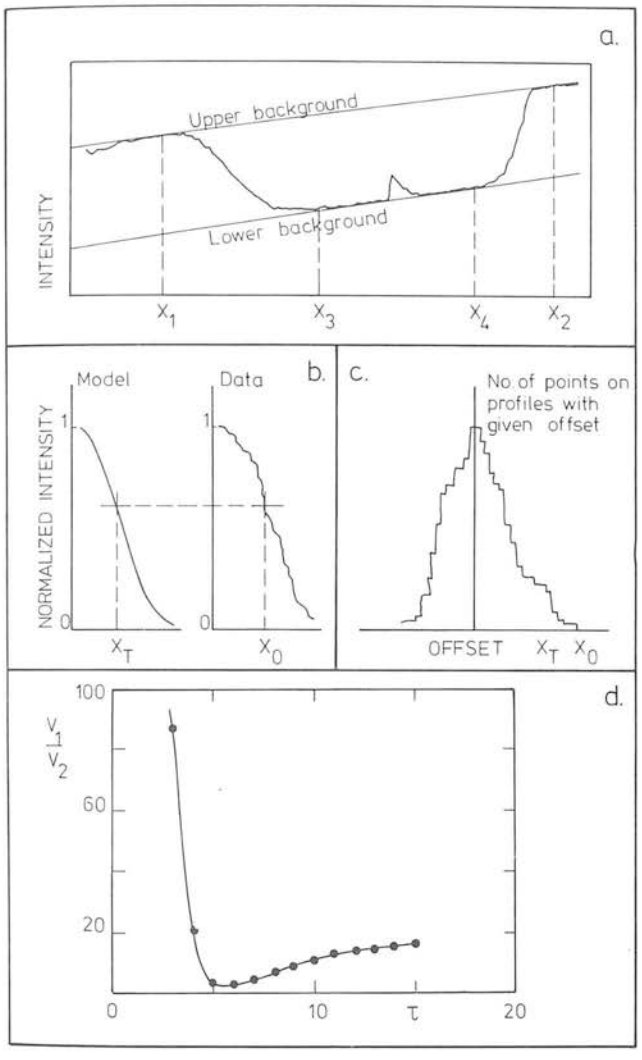


Fig. 6a—d. The method of fitting the theoretical model to the observed edge profiles (Section 7). **a** The normalisation procedure. **b** and **c** The calculation of the offset between the model profile and the observed profile. **d** The determination of the best fitting model by minimising the ratio of variances V_1/V_2 with respect to the maximum optical depth of the model

The ratio of variance V_1/V_2 is a minimum for the best fitting theoretical profile.

8. Results and Discussion

The results presented in Table 2 have a mean optical depth of 6.6 with a standard deviation of 1.0. There appears to be a trend towards higher optical depths as we approach the shoulder of the 'Dragon' to the north of profile 24. Observed profiles north of this shoulder are affected by a bright rim and cannot be fitted by the preceding methods.

By assuming a simple, single species grain model with (*a*) a gas to dust mass density ratio of 100:1 (*b*) a dust specific gravity of unity, and extinction cross

Table 2. The optical depth of theoretical models best fitting cross sections through the Dragon cloud

Number	Max. optical depth	V_1/V_2	Density (cm^{-3})
1	6	1.5	6300
2	6	2.3	6300
3	6	1.5	6300
4	6	2.4	6300
5	6	6.5	6300
6	7	1.5	7400
7	7	1.0	7400
8	6	2.7	6300
9	6	4.1	6300
10	6	2.4	6300
11	5	2.7	5300
12	5	1.2	5300
13	5	1.6	5300
14	6	2.4	6300
15	7	1.2	7400
16	7	4.7	7400
17	7	1.2	7400
18	9	1.2	9500
19	8	1.7	8500
20	8	1.3	8500
21	8	3.9	8500
22	7	3.4	7400
23	7	1.5	7400
24	8	1.7	8500

section of $3\,10^{-10}\,cm^2$ i.e. an effective radius of 0.1 μm, we find for dust opacity

$$k = 5\,10^{-22}\,n_H$$

where n_H is the atomic number density (atoms cm^{-3}), and n_d is the grain number density. Since $\tau_{max} = 2Rk$ we deduce (with $R = 0.13$ pc for the Dragon)

$$n_H = 1057\,\tau_{max}\,cm^{-3}$$

as the atomic density at the edge of the Dragon dust cloud. This quantity is given as the fourth column of Table 2.

If this dust cloud is within the outer boundary of M8, and is surrounded by dilute ionised gas, it will be surrounded by a D-type ionisation front (Osterbrock, 1974) almost certainly weak D, and therefore the density contrast at the boundary will be between 2.0 and 1.0 times the ratio of temperatures inside and outside the cloud. For a temperature ratio of 100 (e.g. $T_{HII} = 10^4$K and $T_{HI} = 100$ K) this suggests a density contrast of 100 to 200.

The absolute surface photometry of Pronik (1960) and Ishida & Kawajiri (1968) indicate that the emission measure near the Dragon is of the order of 3000 cm^{-6} pc which (assuming a line of sight depth of H II of say 3 pc) suggests an electron density (and therefore a proton density) of approximately 33 cm^{-3}. The ratio of the mean predicted cloud density to this is 210, in the range of values of density predicted for D-type fronts.

A further consequence of a D-type ionisation front is that it will have been preceded into the cloud by a shock, which may have increased the interior density by a considerable amount. However we can assign a lower limit to the mass of the cloud by assuming that its mean density is equal to the mean density found by us, and that it approximates to a cylinder of radius 0.13 pc and length 1.3 pc. In this way we find a lower limit to the mass, of 9 solar masses.

Acknowledgements. We should like to thank the UK Schmidt Telescope Unit for the loan of non-survey plates, Dr. Pratt, and Messrs Martin, Williams and Barclay for measuring the plates on COSMOS, and the SRC for providing funds for this work.

References

Bok, B., Reilly, E.: 1947, *Astrophys. J.* **105**, 255
Edwards, A.C.: 1975, private communication
Erdelyi, A.: 1954, Tables of Integral Transforms, McGraw Hill, New York
Foukal, P.: 1969, *Astrophys. Space Sci.* **4**, 127
Goy, G.: 1973, *Astron. Astrophys. Suppl.* **12**, 277
Ishida, K., Kawajiri, N.: 1968, *Publ. Astron. Soc. Japan* **20**, 95
Miller, J.: 1968, *Astrophys. J.* **151**, 473
Nelson, C.N.: 1971, *Photogr. Sci. and Eng.* **15**, 82
Osterbrock, D.: 1974, Astrophysics of Gaseous Nebulae, Freeman, San Francisco
Pratt, N., Martin, R., Alexander, W.G., Walker, G.S., Williams, P.R.: 1975, Image Processing Techniques in Astronomy Proceedings, p. 217, Ed. C. de Jager and H. Niewenhuijzen, Reidel
Pronik, W.I.: 1960, *Bull. Crimean Obs.* **23**, 3
Sharpless, S.: 1959, *Astrophys. J. Suppl.* **4**, 257
Walker, M.: 1957, *Astrophys. J.* **125**, 636

Mon. Not. R. astr. Soc. (1983) **203**, 215–222

Cometary Globule 1

P. W. J. L. Brand *University of Edinburgh, Edinburgh EH9 3HJ, Scotland*
T. G. Hawarden, A. J. Longmore and P. M. Williams
Royal Observatory, Edinburgh EH9 3HJ, Scotland
J. A. R. Caldwell *South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape, South Africa*

Received 1982 June 29; in original form 1982 January 18

Summary. The head of Cometary Globule 1, a dense dust cloud with a faint luminous tail, is illuminated by a variable F2 IIIe star. The age and distance of this star, and hence of the globule, are discussed. It is proposed that the form of the globule is described well by Woodward's model of a spherical cloud through which an initially plane shock has passed. This seems to be the only means to eject dense material downstream as has happened in the tail of Cometary Globule 1. If the illuminating star was formed by the implosion of the cloud and has since drifted to the front edge, comparison with the models suggests that the implosion occurred 1.2 ± 0.6 million years ago.

1 Introduction

The cometary globules (Hawarden & Brand 1976, hereafter HB) which have heads pointing to an apparent centre in the Gum Nebula (HB; Zealey *et al.* 1982) form part of a class of objects which may be quite general in the interstellar medium, but which are most clearly delineated against the Gum Nebula. These globules with tails may be the most probable configuration for cool clouds in an interstellar medium whose dynamics are dominated by supernova explosions. The tails, streamers in the wind, are normally too faint to see against the irregular galactic background and only the heads are visible as dense cloudlets, possibly with emission nebulosity associated with recent star formation embedded in them. NGC 5367 (Williams *et al.* 1977) is an example of such a cloud that 'became' a cometary globule (number 12 in HB) on the SRC IIIa-J Schmidt deep sky survey. In this paper we use observations of the star at the head of cometary globule 1(CG1) in the list of HB to set limits on the evolution of the globule and suggest that the evolution resulted from a blast wave colliding with an initially approximately spherical cloud. CG1 is shown in Plate 1.

2 The star in CG1

The star CD −44°3318 in the head of CG1 is a suspected variable (SVS 1025, Kukarkin *et al.* 1951) and emission line star (He 3 – 32, Henize 1976). A spectrum of the star and surrounding nebulosity was observed with the Wampler–Robinson image-dissector scanner on the Anglo-Australian telescope in 1975 December and is shown in Fig. 1. Only the Hα line is seen in emission. The spectrum was classified F2 IIIE from the G band, Ca II *H* and *K* (where there was some possible emission), Ca I λ 4227, Mn I λ 4030 – 4 and, Ti II λ 4444, (Morgan, Abt & Tapscott 1978; Yamashita, Narai & Norimoto 1977). A recent observation of the blue part of the spectrum at much higher resolution, provided by Drs Cooke and Kelly, confirms that the *H* and *K* calcium lines are filled in by emission, and also confirms the spectral type.

Photometry in *UBVRI* and *JHKL* were taken on different dates in 1976 using the 1-m ANU telescope at Siding Spring Observatory. Owing to difficulty in interpreting the combined energy distribution and the suspected variability, further *UBVRI* photometry was secured in 1981 March and April, using the 1-m Elizabeth reflector at the SAAO Sutherland station. The data are given in Table 1 and confirm the suspected variability of the star although they are insufficient to enable us to classify the type of variation. The locus of the star in the colour – magnitude plane during its variation is most unusual. There is very little change in colour while the flux varies by over a magnitude and what little variation there is, is in the sense of bluer $(B-V)$ at fainter magnitudes. This gradient is quite unlike that $[\Delta V/\Delta(B-V) \sim 3.3]$ shown by a pulsating variable during its cycle, allowing us to eliminate classifications of this type. Although we cannot yet rule out an eclipsing binary, the absence of a well-defined maximum rules out all but the closest separations. Given the presence of the star in the cometary globule and the infrared excess demonstrated by the *JHKL* photometry, we consider the star to be an irregular variable owing its variation to extinction by moving dust clouds in the remnants of the prestellar nebula. Such stars show gradients $R' = \Delta V/\Delta(B-V)$ ranging from near the interstellar value of the total-to-selective

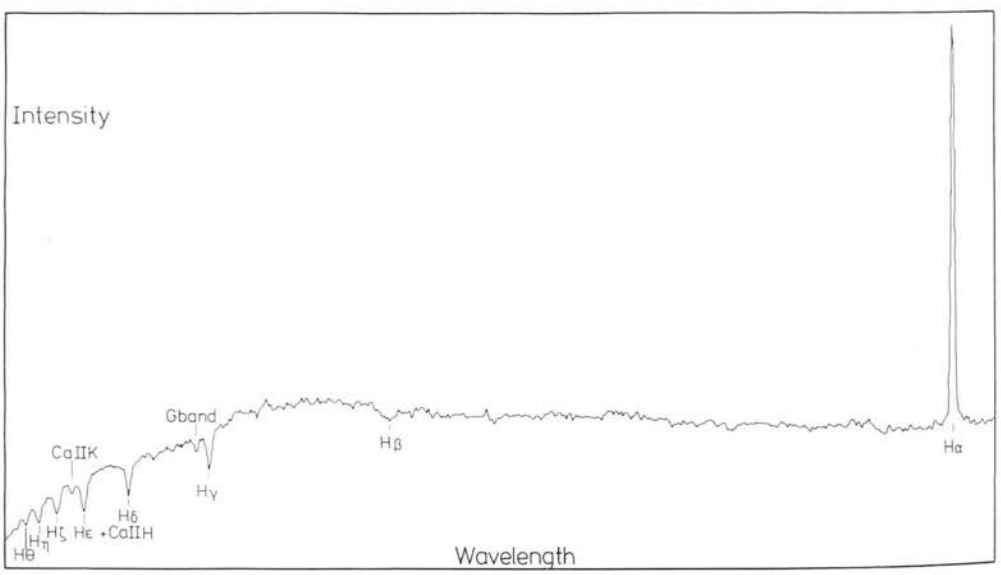
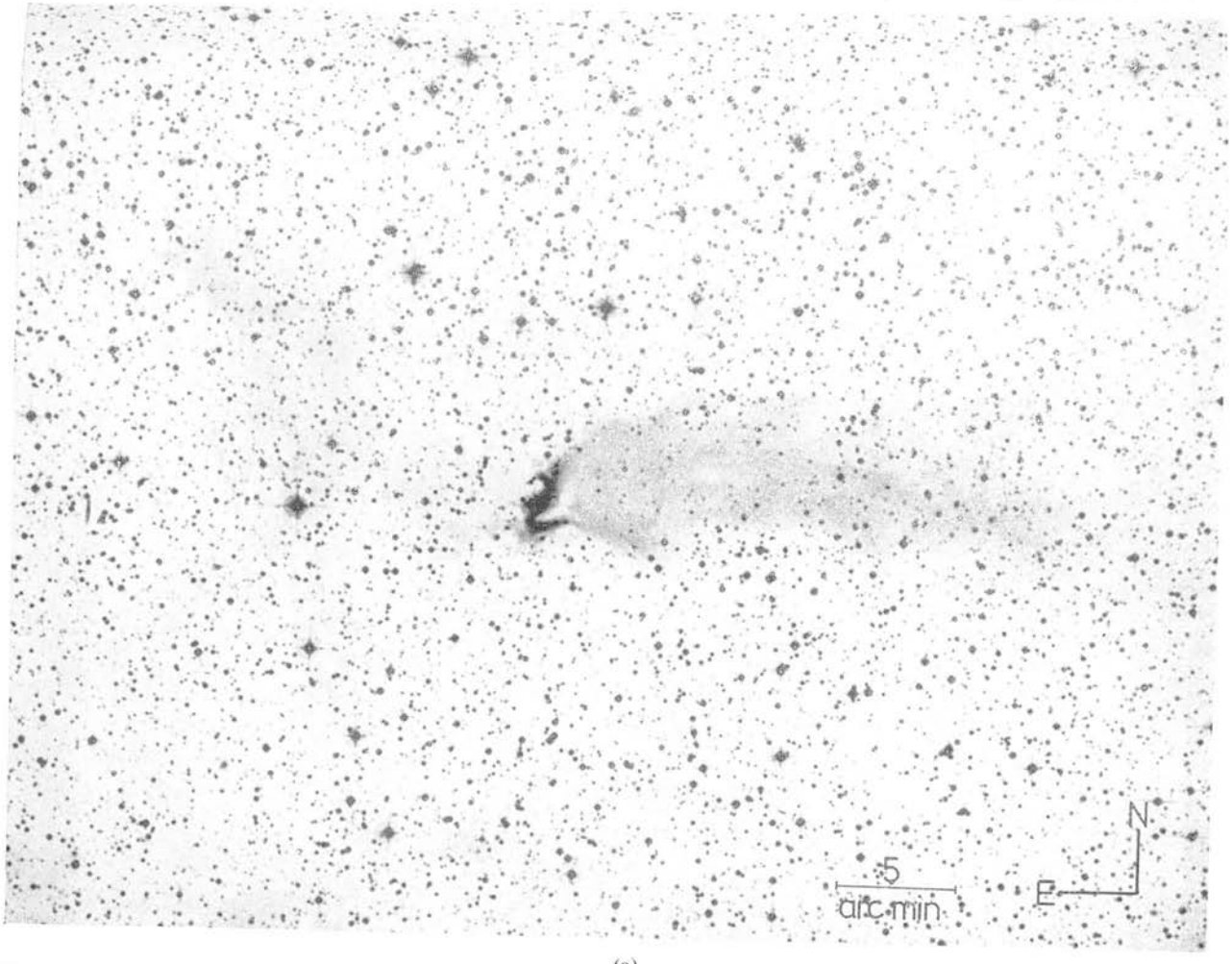
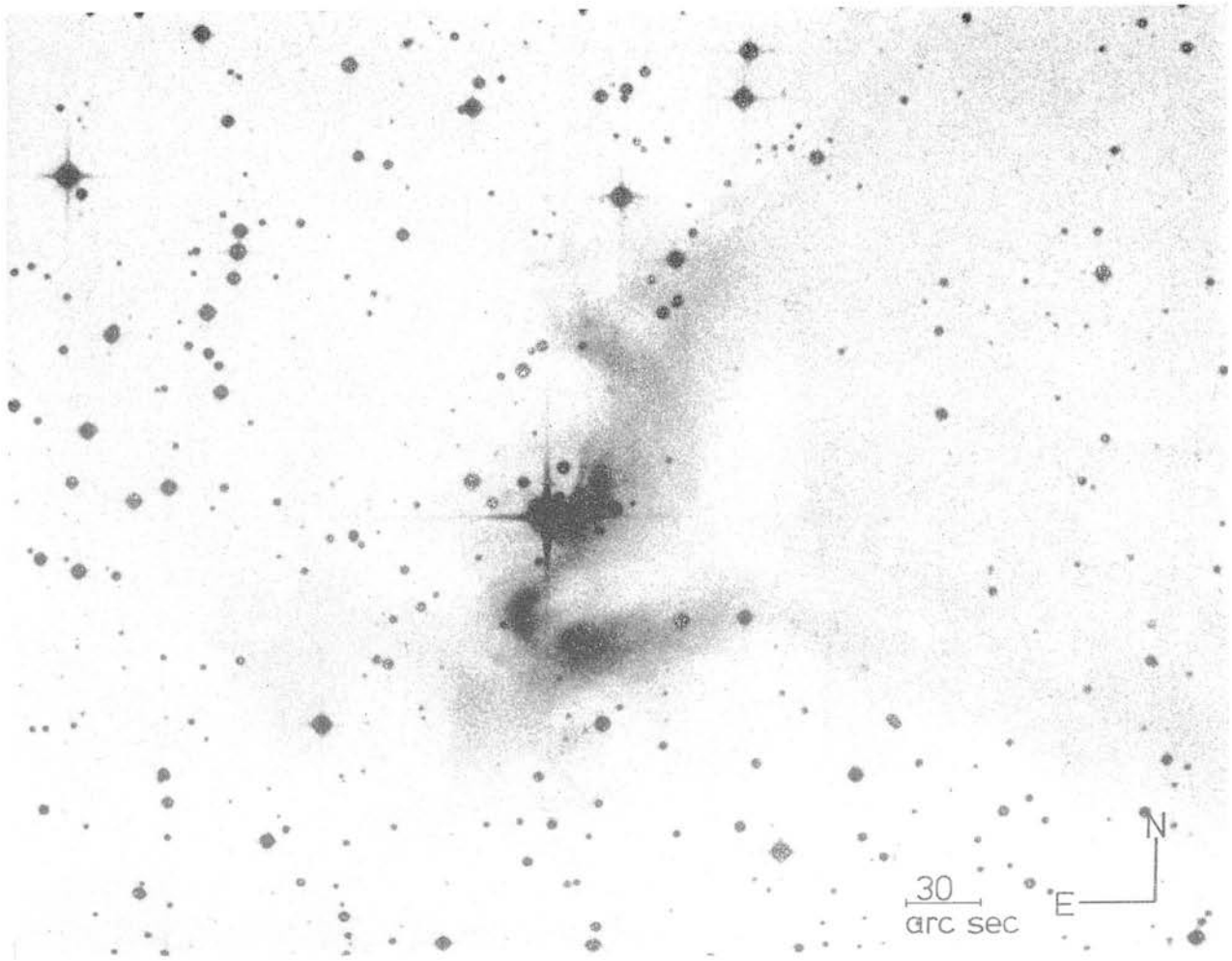


Figure 1. The spectrum of He 3 – 32 in 1975 December. This spectrum is on a linear scale, with a resolution of 10 Å.



(a)

Plate 1. Cometary Globule 1. (a) The complete globule, (b) The head, showing (with diffraction spikes) the star He 3−32 embedded in the head.



(b)

Table 1. Photometry of He – 3 32 (– 44°3318).

Date	V	$U-B$	$B-V$	$V-R$	$V-I$	$V-J$	$V-H$	$V-K$	$V-l$
1976	10.14	0.13	0.53	0.31	0.64	2.13	3.20	4.29	5.64
1981 March 12	10.890	–	0.373	0.338	0.753	–	–	–	–
1981 March 13	10.694	–	0.473	0.404	0.832	–	–	–	–
1981 March 14	10.630	–	0.465	0.396	0.823	–	–	–	–
1981 March 15	10.580	–	0.474	0.400	0.822	–	–	–	–
1981 April 8	10.514	0.022	0.484	0.404	0.842	–	–	–	–
1981 April 9	10.744	–0.016	0.454	0.413	0.851	–	–	–	–
1981 April 16	11.158	–0.076	0.428	0.427	0.876	–	–	–	–
1981 April 17	11.123	–0.082	0.445	0.424	0.895	–	–	–	–

The data given to three decimal places are accurate to approximately 1 per cent.

extinction to several times this value (e.g. Tjin A. Djie & Thé 1978; Walker 1980). This range reflects the composition and grain size distribution of the circumstellar clouds; the larger values of R' being due to larger grains.

We consider the variation of the CGI star in this framework. The lack of increased reddening when the star is dimmed suggests that the grains are relatively large while the range of variation indicates that the clouds are optically thick. It is reasonable to identify the excess infrared radiation measured in *JHK* and *L* with these clouds. Although the optical thickness prevents our inferring the wavelength dependence of the grains' emissivity (and hence their composition) from the infrared emission, it allows us to treat the grains as black bodies. The *JHKL* photometry, corrected for the stellar contribution, indicates a characteristic grain temperature of 1300 K. We can use this temperature and that of the star (6700 K from the spectral type) to determine the distance of the emitting grains from the star by requiring the grains to be in thermal equilibrium with the stellar radiation. This leads to a characteristic cloud distance of 13 stellar radii for the hottest components, whose radiation dominates the *JHKL* measures. Further photometry at longer wavelengths is needed to infer the contribution of cooler material. It would be useful to obtain polarimetry and further photometry in the visible and infrared in order to provide tighter constraints on possible models of the irregular nebula believed to be responsible for the light variation.

The time-scale of one day in the photometric variability suggests a characteristic cloud size (assuming for the purpose of the estimate that the star is of solar mass) of two solar radii, and characteristic velocity of several tens of km s^{-1}. (The cloud size (l) and velocity (v) are given by

$$l/R_\odot = 14(\alpha M/M_\odot)^{1/3} \tag{1}$$

$$v = 116(\alpha M/M_\odot)^{1/3} \text{ km s}^{-1} \tag{2}$$

where α is the ratio of cloud size to distance from the star and M is the star's mass. $M_\odot$ and $R_\odot$ are solar mass and radius respectively.) The cloud size must be comparable with or greater than the size of the star, which we presume lie above the main sequence and therefore are of a radius larger than that of the Sun. It is satisfying that the cloud distance derived from equation (1) with $\alpha \simeq 0.1$ agrees with the distance derived above from the *JHKL* photometry. Fig. 2 is a theoretical Hertzsprung–Russell diagram taken from Cohen & Kuhi (1979) on which the range of temperatures inferred from the spectral type is superposed. If we presume that the luminosity class is that derived from the spectral features (admittedly a risky presumption; Welin 1978), then the luminosity of the star lies within limits defined by the box in Fig. 2, allowing an error of one-half in luminosity class. The box defined by these

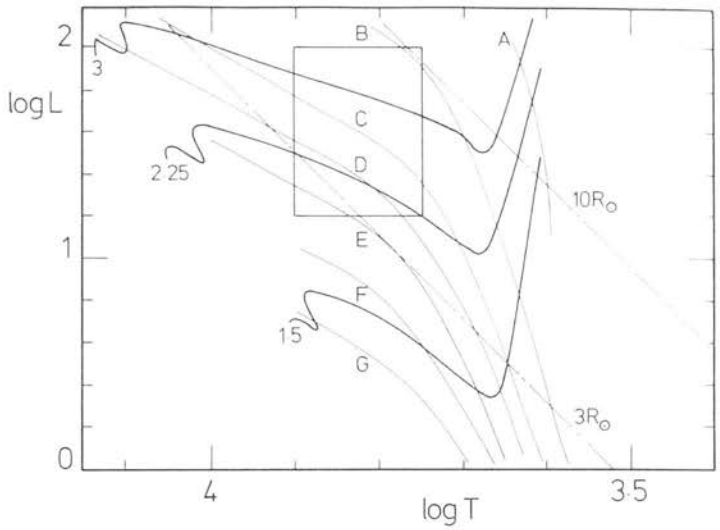


Figure 2. Part of the HR diagram of Cohen & Kuhi (1979). The dashed lines labelled A, B, C ... are isochrones. A = 0.03 Myr, B = 0.3 Myr, C = 1.0 Myr, D = 1.75 Myr, E = 3 Myr, F = 6 Myr, G = 10 Myr. The box shows limits on the position of He – 3 32.

estimates, and by allowing the range of probable spectral type to be F2 to G0, fixes the possible range of masses, radii and ages of the star. These are $2M_\odot$ and $2.5\,R_\odot$ at 4×10^6 yr to $3.5\,M_\odot$ and $10\,R_\odot$ at 3×10^5 yr.

Finally, if we (a) take the colour excess to be $E(B-V) = 0.2$, (b) assume the ratio of total to selective absorption is 3.3 (Allen 1973), (c) take the maximum observed V magnitude, and (d) estimate $M_v = 1.2 \pm 0.6$ (Allen 1973), we derive a distance of approximately 500 pc, with very large uncertainty. This compares with a distance of 400 ± 60 pc derived for the distance to the Gum Nebula by Brandt *et al.* (1971).

3 The Gum Nebula

Reynolds (1976b) has fitted a scaled version of the supernova remnant (SNR) model by Chevalier (1974) to his own Fabry–Perot observations (Reynolds 1976a). He obtains a best fit for an $E_0 = 5 \times 10^{51}$ erg initial explosion energy in an ambient density of 0.25 H atoms cm^{-3}, at an age of 1.0×10^6 yr. This fits the velocities he observed in the range 10–30 km s^{-1}, and a radius of 125 pc (assuming a distance of 400 pc). In our own scaled fit to Chevalier's models we derive an expression for the blast wave radius (R) in parsecs:

$$R = 36\, E_0^{0.2}\, n_0^{-0.25}\, t^{0.31} \tag{3}$$

where E_0 is the total energy of the initial explosion in units of 10^{50} erg, n_0 is the ambient number of H atoms per cm^3 prior to the explosion, and t is the elapsed time since the supernova explosion in units of 10^6 yr.

From this, one can find the relations

$$t = 0.30\, R/v_b \tag{4}$$

$$E_0 = 1.2 \times 10^{-7}\, v_b^{1.54}\, R^{3.43}\, n_0^{1.25} \tag{5}$$

where v_b is the blast wave velocity in km s^{-1}. For $10 \text{ km s}^{-1} < v < 30 \text{ km s}^{-1}$, $n_0 = 0.25$ cm^{-3}, $R = 125$ pc, we get $1.2 \text{ Myr} < t < 3.7$ Myr, and $10^{51} \text{ erg} < E_0 < 6 \times 10^{51}$ erg. The McKee & Ostriker (1977) model cannot be fitted in its standard form, without such a low hot gas

filling factor that the scheme in which their model is assumed to work not operate, and the evolution would appear rather like that described by Chevalier. Thus it appears that the Gum Nebula remnant resulted from an explosion (or several explosions) of energy between 10^{51} erg and 6×10^{51} erg some 1 to 4 Myr ago in a medium whose density is 0.25 cm^{-3} as derived by Reynolds (1976b).

4 The Cometary Globule

The head of the globule is at RA = $7^h\,18^m.6$, Dec = $-44°\,32'$ (1950) and the position angle of the tail is 270° (HB). The tail is 25 arcmin long, while the characteristic size of the head is 2 arcmin. The head is opaque, and most importantly there are parts of the tail (approximately 5 and 10 arcmin from the head) which also appear to be opaque. A possible explanation for the tail is that it is ablated or ionized material being swept away from the dense core of the cloud by a wind or radiation originating in the central region of the SNR. However, the opaque parts of the tail suggest lumps as dense as the core of the cloud have been driven downstream, and these are not likely to be separated from the cloud by a wind or radiation. A means must be found for driving dense gas out of a cloud.

Such a means has been already demonstrated in the calculations by Woodward (1976, 1979) who computed the effects of a blast-wave hitting an initially spherical cloud. The configuration, in its late stages, closely resembles a cometary globule. The tail has been produced mechanically, and therefore dense material can be brought 'downwind' of the globule. Interestingly, the detailed fragmentation of the sides of the imploded cloud which is demonstrated in Woodward's models, appears to show up in the case of Cometary Globule 12 (HB, Williams *et al.* 1977).

Heathcote & Brand (1982) have investigated analytic approximations to various parts of Woodward's model, and summarize the major phases of evolution of the cloud in their table 1. Briefly, the events separating the phases are as follows: (a) the SNR blast-wave strikes the front of the spherical cloud (which was initially in pressure equilibrium) and initiates a forward-moving shock within the cloud; (b) the blast-wave sweeps past, and on reaching the back of the cloud, initiates a shock into the cloud in the opposite direction; (c) the internal shock propagating from the front of the cloud collides with the reverse shock. Both of these shocks are approximately planar, and cause the cloud to 'pancake'; (d) there follows a re-expansion phase, returning the outer parts of the cloud to pressure equilibrium. This last phase is the most complex, including as it does the lateral compression (much weaker than that of the two main shocks), flow due to instability of cloud material to a dense clump, and the propagation of the original forward moving internal shock into the intercloud medium while a rarefaction wave causes dense cloud material to stream in a broad fan or tail in the direction of motion of the blast-wave. This phase produces the phenomena of interest, namely gravitationally unstable material in a dense clump or 'head', and a 'tail' of cloud material considerably denser than the intercloud medium. A comparison of Woodward's (1976, 1979) diagrams with photographs of cometary globules (particularly CG1 and CG12, HB) provides a persuasive argument that cometary globules are the late stages of shocked clouds.

Heathcote and Brand (1982) show in their table 1 that the times Δt_c for maximum cloud compression (event c above) and Δt_e for complete re-expansion (event d), measured from first contact between cloud and blast-wave, are

$$\Delta t_c = (0.64/M_b)\, t_0 \tag{6}$$

and

$$\Delta t_e\ (0.72\, t_0 + \Delta t_c). \tag{7}$$

In these formulae M_b is the Mach number of the incident shock, and t_0 the sound crossing time, the ratio of initial cloud radius to isothermal sound speed in the cloud. Furthermore, the length of the cloud (head plus tail) at the end of the expansion phase is approximately 60 per cent of the original cloud diameter. This implies that with a distance of 400 pc (Section 2) the original cloud had a diameter of 5 pc (this corresponds to a mass of $14 M_\odot$ if the initial density is $10\,\mathrm{cm}^{-3}$, typical for small isolated clouds).

To estimate the Mach number of the incident shock, we note that the distance to the globule from the assumed blast centre is approximately 80 per cent of the current SNR radius R. Then (from equation 5) the velocity with which the cloud was struck is $(0.8)^{-2.23} = 1.6$ times the current value (estimated to be 10–30 km s^{-1} by Reynolds 1976a). By assuming a value for mean interstellar pressure of 3000 K cm^{-3} (e.g. typical of the range discussed by McKee & Ostriker 1977) we find M_b lies in the range 1–3, very approximately.

We can also make a rought estimate of the cloud sound crossing time t_0 to be of the order of 2×10^6 yr by assuming its temperature before interaction was 100 K, a value not untypical for a cloud of its size.

5 The sequence of events

We can now place the events in time sequence, and establish limits on the time intervals involved. We define t_R as the age of the SNR now, related to the SNR size by equation (3). Then we define t_* as the age of the star, and t_0, Δt_c, Δt_e are already respectively defined (*cf.* equations 6 and 7) as the original cloud sound crossing time, the time during which the cloud is compressed, and the time taken to re-expand to pressure equilibrium, from first contact with the blast wave. With a tail length 0.6 times the original cloud diameter, and a distance to the cloud of D pc, the cloud crossing time in units of 10^6 yr is

$$t_0 = (2.4\ \mathrm{km\ s^{-1}}/c_{\mathrm{cloud}})\,(D/400\ \mathrm{pc}) \tag{8}$$

where c_{cloud} is the original cloud soundspeed.

The age of the SNR when it hit the cloud, at 0.8 times the present SNR radius, is (equation 3) $0.49\,t_R$. In the time since then, the cloud has collapsed and a star has formed. Hence,

$$0.51\,t_R = \Delta t_c + t_*. \tag{9}$$

Table 2. The sequence of events leading to CG1, derived from equations (3), (6) and (7).

Event	Epoch
Supernova explodes	0
SNR hits cloud (at $R = 0.8 R_{\mathrm{now}}$)	$0.49\,t_R$
Star forms at max collapse	$0.49\,t_R + (0.64/M_b)\,t_0$
Cloud re-expanded	$0.49\,t_R + (0.64/M_b + 0.72)\,t_0 = t_{\mathrm{exp}}$
Now	$t_{\mathrm{exp}} + t_{\mathrm{delay}} = t_R$

Constraints (used to define lines in Fig. 3).

$T_{\mathrm{cloud}} < 4000$ K, equation (8)	$t_0 > 0.6$ Myr
Upper limit to SNR age	$t_R < 3.7$ Myr
$M_b = 2$, equations (9), (10)	$t_R > 2\,t_0$
Equation (9)	$t_R > 2\,t_*$
$t_* \geqslant \Delta t_e$	$t_* > 0.72\,t_0$

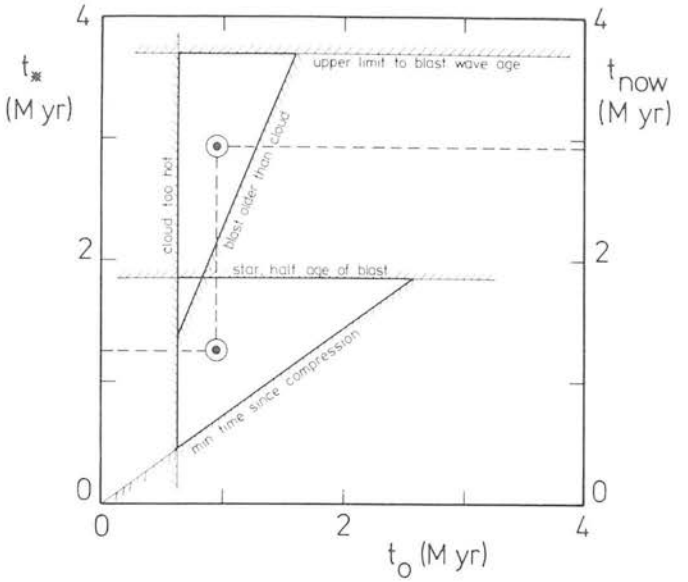


Figure 3. The various constraints on t_*, t_R and t_0 are shown here, and discussed in Section 5. The optimum solutions are shown as circled dots, referred to the relevant axes.

Furthermore,

$$0.51\, t_R = \Delta t_e + \Delta t_{delay} \tag{10}$$

where Δt_{delay} is the time since the cloud completed its re-expansion.

The sequence of events is given in Table 2, together with the constraints that are marked on Fig. 3.

Now we can see that with the initial estimates from earlier sections of $t_* = 0.3$–4 Myr, $t_{now} = 1.2$–3.7 Myr, $M_b = 1$–3, $D = 300$–500 pc, $c_{cloud} = 1$–3 km s^{-1}, we have a region of self-consistency where

Age of the star	$t_* = 1.2 \pm 0.6$ Myr
Cloud sound crossing time	$t_0 = 0.9 \pm 0.3$ Myr
Age of SNR	$t_R = 2.9 \pm 0.8$ Myr
Time since cloud re-expansion	$t_{delay} = 0.6 \pm 0.8$ Myr.

This is illustrated in Fig. 3.

6 Conclusion

The sequence of events leading to the creation of a cometary globule has been followed in some detail, based on the simple and attractive assumption that cometary globules are shocked clouds. The similarity of appearance of CGs and Woodward's (1976, 1979) cloud models, the existence of a blast-wave that could be responsible for the formation, the presence of a recently formed star in the head of the globule, and the detailed matching of derived ages and sizes, all argue strongly in favour of the shock origin of cometary globules, and of such globules as the sites of star formation. This process appears sparse as a star generator in the

Gum Nebula. However, in regions of density higher than that of the Gum Nebula, cometary globules may be a significant source of low mass stars.

Acknowledgments

We thank D. Allen with help in obtaining the spectrum in Fig. 2, J. Cooke and B. Kelly for providing us with another spectrum of CG1, and PATT and ANU for awarding the telescope time, and the referee for helping to improve the presentation of this result.

References

Allen, C. W., 1973. *Astrophysical Quantities*, Athlone, London.
Brandt, J. C., Stecher, T. P., Crawford, D. L. & Maran, S. P., 1971. *Astrophys. J.*, **163**, L99.
Chevalier, R. A., 1974. *Astrophys. J.*, **188**, 501.
Cohen, M., Kuhi, L. V., 1979. *Astrophys. J.*, **227**, L105.
Hawarden, T. G. & Brand, P. W. J. L., 1976. *Mon. Not. R. astr. Soc.*, **175**, 19P.
Heathcote, S. R. & Brand, P. W. J. L., 1982. *Mon. Not. R. astr. Soc.*, **198**, 551.
Henize, K. G., 1976. *Astrophys. J. Suppl.*, **30**, 491.
Kukarkin, B. V., 1951. *Catalogue of Stars Suspected in Variability*, Moscow.
McKee, C. F. & Ostriker, J. P., 1977. *Astrophys. J.*, **218**, 148.
Morgan, W. W., Abt, H. A. & Tapscott, J. W., 1978. *Revised MK Spectral Atlas*, Yerkes, University of Chicago.
Reynolds, R. J., 1976a. *Astrophys. J.*, **203**, 151.
Reynolds, R. J., 1976b. *Astrophys. J.*, **206**, 679.
Tjin, A. Djie, H. R. E., Thé, P. S., 1978. *Astr. Astrophys.*, **70**, 311.
Walker, M., 1980. *Publ. astr. Soc. Pac.*, **92**, 66.
Welin, G., 1978. In *Protostars and Planets*, ed. Gehrels, T., University of Arizona Press.
Williams, P. M., Brand, P. W. J. L., Longmore, A. J. & Hawarden, T. G., 1977. *Mon. Not. R. astr. Soc.*, **181**, 709.
Woodward, P. R., 1976. *Astrophys. J.*, **207**, 484.
Woodward, P. R., 1979. In *Large Scale Characteristics of the Galaxy, IAU Symp. 84*, p. 159, ed. Burton, W. B., Reidel, Dordrecht, Holland.
Yamashita, Y., Nariai, K. & Norimoto, Y., 1977. *Atlas of Representative Stellar Spectra*, University of Tokyo.
Zealey, W. J., Ninkov, Z., Rice, E., Hartley, M. & Tritton, S. B., 1982. *Astrophys. Lett.*, submitted.

Astron. & Astrophys. 38, 363–371 (1975)

Cloud Structure in the Galactic Plane: A Cosmic Bubble Bath?

P. W. J. L. Brand and W. J. Zealey
Department of Astronomy, University of Edinburgh

Received July 23, revised September 16, 1974

Summary. An examination of photographs of the Milky Way has given strong indications of large scale loop-shaped dust and gas structures. A list of these, many of them already noted by others, is provided. In addition it is suggested that a large part of the interstellar medium is in the form of filaments, and that these filaments may be related to the loop structures. It is shown that much of the interstellar medium may have been processed through supernova blast waves, and that recent estimates of energetics of these provide ample kinetic energy for the observed motions in the interstellar medium. Whether or not they are due to supernova explosions, many examples of dust shells round regions of recent star formation are known in external galaxies. If these shells extend out of the plane of the galaxy, they will be subject to Rayleigh-Taylor instability and may fall back as streamers of material. We suggest that a process of diffusion of the dust to the cores of these streamers may preserve filamentary structure in the galactic plane.

Key words: supernova remnants — interstellar clouds — interstellar dust — filaments

1. Introduction

The optical appearance of the galactic plane is chaotic at first glance due to the apparent randomness of structure in the gaseous emission and dust clouds in the plane. A closer examination strongly suggests two facts. The first is that the dust lanes possess a measure of coherence, over several degrees in many cases, the most striking of these structures being loops.

The second fact is that the dust clouds fall naturally into two categories—sharp-edged and diffuse clouds. As a general rule sharp-edged clouds (of which bright rims are an extreme example) are closely associated with H II regions or star clusters and evidently have been or are being influenced by these. The remainder of the dust appears to have a structure best described as filamentary, whenever it is seen at sufficiently small optical depth against a bright background. It is at present very difficult to quantify this, although it is hoped that large scale star counts and measurements of background using the COSMOS measuring machine which will be in operation at the Royal Observatory Edinburgh in the near future, will provide the necessary data. This phenomenon appears to be not merely local, since it may be seen at the edges of the dense clouds in which the H II regions M 8, M 20 and M 17 are embedded in the Sagittarius Arm, and may readily be verified by an examination of the Palomar Sky Survey charts. An example of a region at $l = 358°$, $b = 4$ containing non-sharp edged clouds lying in filaments is shown in Fig. 1.

Thus there is some evidence that except near regions of recent or on-going star formation, the cloudy structure of the interstellar medium is essentially filamentary. The overall topography of the filaments suggests the same tangled structure as that of a magnetic field in energy equipartition with the gas. Whether the filaments outline the magnetic field or not, the usual assumption of constant gas-to-dust ratio produces an embarrassing situation, since filaments in such cases should be unstable to lateral dissipation on a small timescale. The problem of how filaments are generated will be returned to in Section 5. We suggest that certain puzzling structures such as the filaments extending perpendicular to the edge of the dark cloud in IC 434 (surrounding Horsehead Nebula) which pose theoretical difficulties if dynamically a part of the interaction between the dark cloud and the radiation incident on it (Goldsworthy, personal communication), are clouds of the type discussed, silhouetted against the uniform brightness of IC 434. A similar situation occurs in the region to the south of NGC 7000

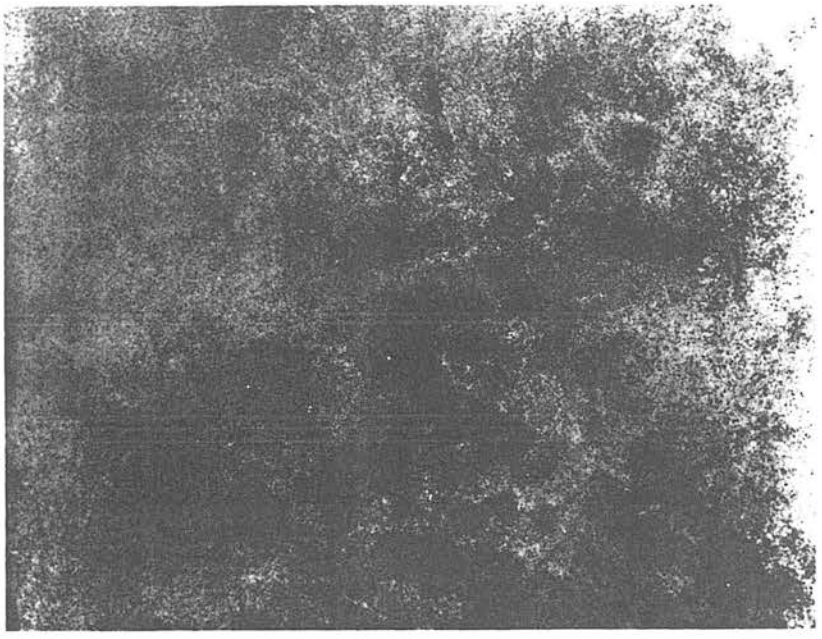


Fig. 1. An enlargement from the red Palomar Sky Survey, showing an area approximately $2^\circ \times 3^\circ$ at $l = 358^\circ$, $b = 4^\circ$. The dust structures are predominantly elongated and appear threadlike rather than as distinct, localised blobs

("mid atlantic" with respect to the North America nebula) where the obscuration appears as "cross-hatching".

We shall discuss the possibility that there is a connection between loops and filamentary structure in Section 5.

2. Loop-type Structures of Dust and Gas

H II regions showing well defined loop structure were first commented on by Gum and de Vaucouleurs (1953). The most obvious of the galactic H II regions were Barnard's Loop complex, the Scorpius OB1 region, the Vela/Puppis nebulosity and the λ Orionis nebula. Further loop structures have been noted by Morgan *et al.* (1955) in Monoceros and Cygnus; Westerlund (1960) has noted that HD 148937 is located near the centre of a series of ring-like nebulae.

Radio continuum emission loops or spurs have been observed centred at intermediate latitudes and extending over tens of degrees (Haslam *et al.*, 1971; Berkhuijsen *et al.*, 1971). These coincide with H_α emission filaments (Meaburn, 1967) and are considered to be old supernova remnants lying within 300 pc of the Sun (Berkhuijsen, 1973).

The connection between these loops and associations or other objects was pursued, using the Atlas of H_α Emission in the Southern Milky Way (Rodgers *et al.*, 1960a), a photomosaic of some forty Palomar charts and overlays of positions of (i) associations (Ruprecht, 1964), (ii) supernova remnants (Downes, 1971), and (iii) X-ray sources (Giacconi *et al.*, 1974). In addition, indications of loop structure were searched for in the dark cloud atlases of Lynds (1962) and Khavtassi (1960). Radio H I surveys and H_α photos from the literature were also studied. The only major correlations found were with the positions of associations, and the results of this search are described in the appendix.

The paper by Sivan (1974) which appeared as this paper was submitted contains evidence of further H_α emission loops.

3. Shells in External Galaxies

Lynds has been carrying out a systematic survey of dust in external spiral galaxies (Lynds, 1970, 1972) which demonstrates the fact that the bright (giant) H II regions occurring just outside the thin spiral dust lanes are in regions of high obscuration. What is remarkable in the majority of cases is that the apparent diameter of the obscuring region is very little greater than the diameter of the H II region itself. This suggests that the boundary of the observed region is determined by the dynamical activity of the H II region, forming a shell of neutral gas and dust round it, since it would stretch coincidence to find ionisation-bounded H II regions with radius determined by the outer edge of the clouds in which they formed. Further work remains to be done, since it is not easy to be sure of the exact outer boundary of the obscuring clouds. Several examples of this phenomenon are shown in Fig. 2 which are enlargements from copies of plates taken for the ESO/SRC Sky Survey. Figure 2 also shows an edge-on galaxy from the same source, with dust lanes extending in a sinuous fashion to fairly high z values (one kiloparsec if the galaxy is the same

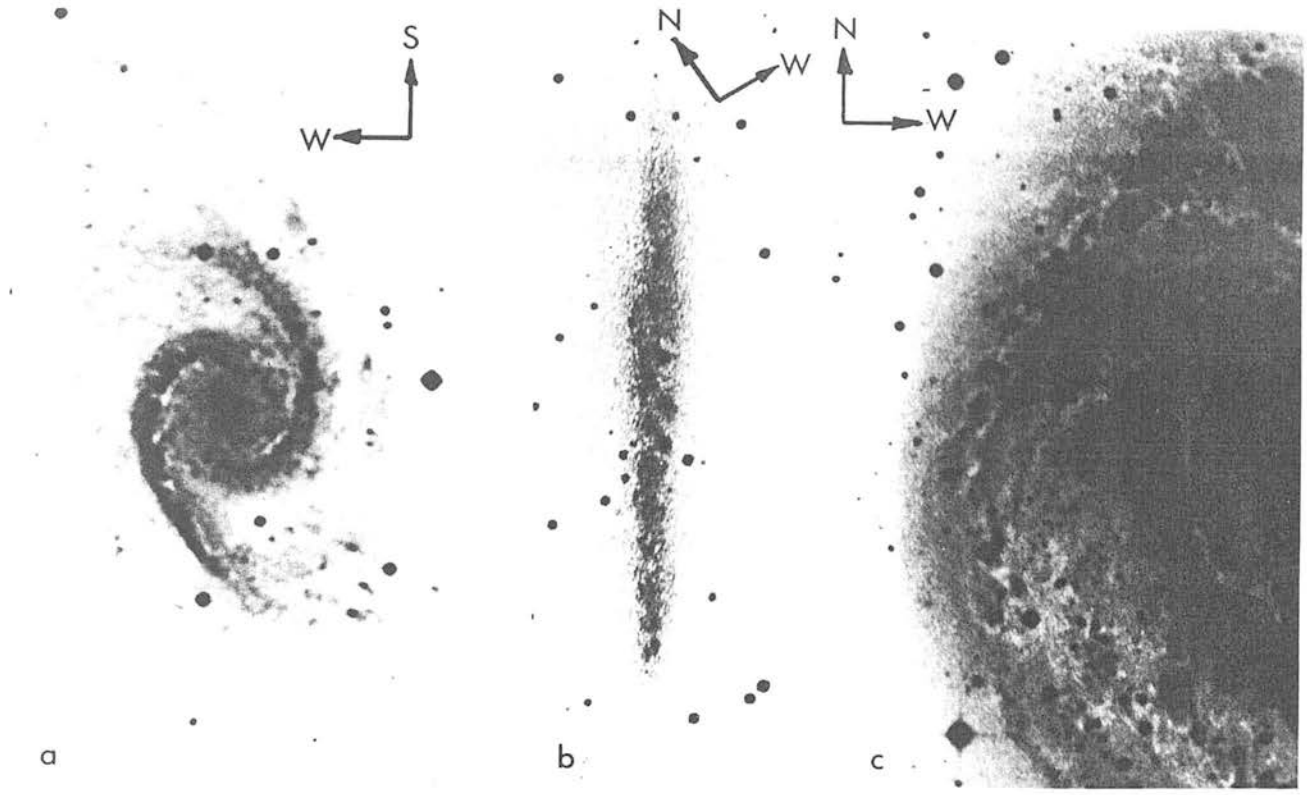


Fig. 2a—c. Some examples of dust structures in external galaxies. (a) NGC 1566 showing a particularly bright emission region, west of the nucleus, surrounded by a thin dust lane. (b) IC 5052 is seen edge-on with thin lanes of obscuration extending large distances from the plane. If the galaxy is not truly edge-on, then it appears to be covered with enormous irregular dust lanes. (c) NGC 5236 showing a series of dust shells surrounding emission regions. Copies from ESO/SRC Sky Survey

size as our own). These structures are most easily explained as segments of dust shells. Dust extending to large distances out of the plane has also been noted by Vorontsov-Velyaminov (1964) in several external galaxies, and there are examples in the Hubble Atlas (Sandage, 1961).

Shell structures in the Magellanic Clouds have been known for a long time (e.g. Lindsay, 1953 discussing observations by Innes in 1918). In the LMC there is a remarkable example discussed by Westerlund and Mathewson (1966), which is apparently 1 kpc in diameter. The northern half consists of a neutral hydrogen loop detected by McGee and Milton (1964) in which several supernova remnants are embedded, while the southern part is a dramatic arc of blue stars noted by Nail and Shapley (1953). Hindman (1967) has remarked on three similar features in the SMC detected in H I. On the same gigantic scale are the apparent dust rings noted by Hayward (1964) in several spiral galaxies.

4. Formation of Shells

Several methods of expelling shells of material from associations of young stars may be envisaged. In order of increasing violence these are:

a) radiation pressure by starlight on the dust component;

b) late stages of expansion of cluster H II regions;

c) supernova explosion.

All of these are related to the existence of high mass stars, and considerations of size and age do not discriminate effectively between the possibilities. In addition, the processes are not exclusive.

Mechanism (*a*) has been considered by O'Dell *et al.*, (1967) in connection with Barnard's Loop. Momentum from the near UV photons emitted by stars in a central cluster is transferred to the gas via radiation pressure on the grains, which are constrained to co-move with the gas by high viscosity.

The gas is swept beyond a radius r given approximately by

$$r = 0.15\,(L/L_{\odot})^{1/3} \quad (t/10^6\ \mathrm{yr})^{2/3}\ \mathrm{pc}\,. \tag{1}$$

Thus for $10^7\,L_{\odot}$ in 3×10^6 years, $r = 64$ pc. Mechanism (*b*) is described (Field, 1970) by the approximate formula

$$r = 32\zeta\,[1 + (t/10^6\ \mathrm{yr})/1.3\,\zeta]^{4/7} \tag{2}$$

where $\zeta = S/(10^{48}\,n^2)$ and S is the number flux of ionising photons, n is the hydrogen number density. For

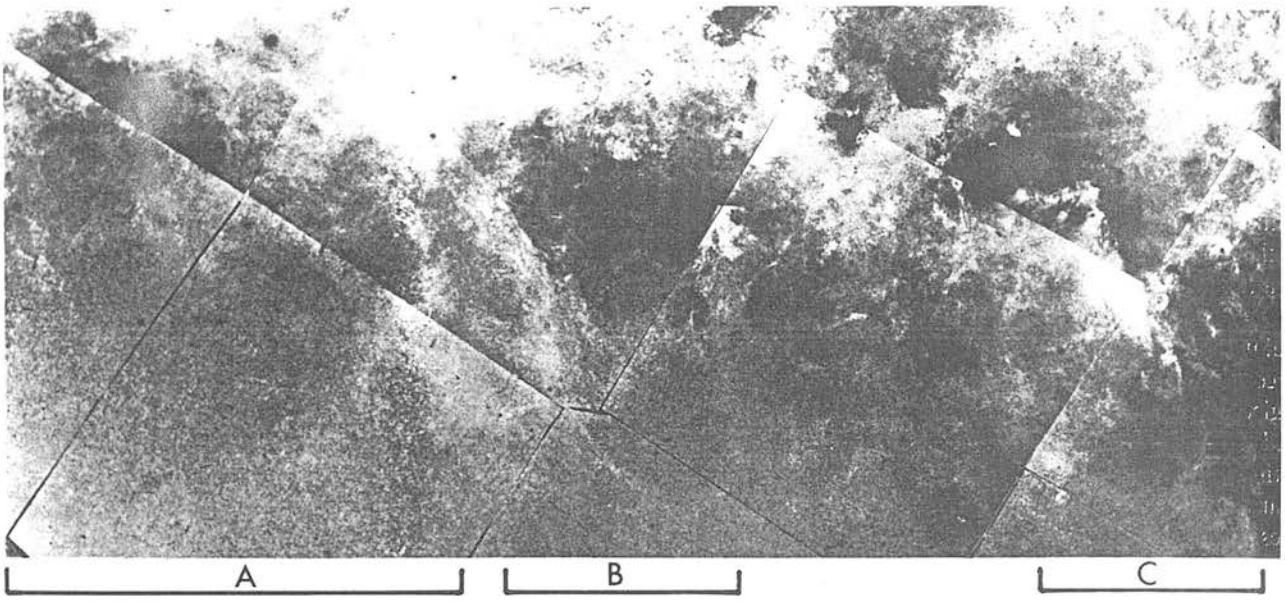


Fig. 3. A sequence of dust structures *A*, *B* and *C* extending below the galactic plane, showing successive stages of dissipation of giant shells blown out of the plane. Structure *A* is part of the loop system around Sct OB2

example, 10 O7 stars (Churchwell and Walmsley, 1973) in $n = 10\,\mathrm{cm}^{-3}$ at 10^6 years give $r = 70$ pc.
Mechanism (*c*) is described by a snowplough model (Woltjer, 1972) with

$$r \doteq (\varepsilon)^{1/4}\,(t\ \mathrm{yr})^{1/4}\ \mathrm{pc} \qquad (3)$$

where $\varepsilon = (E/10^{50}\,\mathrm{erg})/(\varrho/1.6 \times 10^{-24}\,\mathrm{g\,cm}^{-3})$, and with $\varepsilon = 10$, after 10^6 years $r = 56$ pc. Since the shell in all three cases is swept-up mass, and velocities for ages of a million years are all low, it is not obvious that one may discriminate between the possibilities. It may be that for any association these will occur in sequence, one or two high mass stars forming SN remnants, the lower mass stars then forming a large H II region in the low density swept-out region, and yet lower mass stars in much greater numbers cooperating in sweeping grains to the outside of the association.
The very large (1 kiloparsec) shells are best fitted by supernova explosions of $E = 10^{53}$ erg in $n = 0.1\,\mathrm{cm}^{-3}$ material.

5. Condensation of Filaments

So far no generally satisfactory theory of filaments exists. The dynamical problem mentioned in the first section is that the dissipation time for a filament of low optical depth will be given by the ratio of its thickness to the local sound speed, if it represents simply a lateral condensation in a uniform medium. If indeed the filament is aligned along the magnetic field lines then the problem is enhanced by flux-freezing, since the compressed magnetic field lines cooperate with gas pressure to dissipate the filament. In fact, any condensation in the general interstellar medium ought to occur preferentially perpendicular to the field lines, giving "pancake" configurations, contrary to the observations discussed. The one means to overcome this tendency is to heat the gas sufficiently that the dynamics is little affected by the magnetic field, while the non-isotropic conductivity remains. This process has been exploited by McCray *et al.* (1972) to explain the filamentary structure in the Gum Nebula, following the analysis of condensation instability by Field (1965). However, as soon as the gas heated by a burst of UV or X-radiation or of cosmic rays, or by the passage of a hypersonic shock, cools to some 10000 K, the filaments will dissipate, *unless* a mechanism can be found which will maintain the dust against the dissipation.
Another possibility for producing filaments on a large scale is the Rayleigh-Taylor instability associated with the gravitational field of the galactic plane. The time-scale for this instability is of the order of 10^7 years. To operate, a density inversion out of the plane of the galaxy is required, and one would then expect to see streamers of material falling in towards the plane, the motion moderated by the magnetic field to a greater or lesser extent.
We believe that there exists a mechanism that will tend to preserve the filamentary structure created in either of the above processes. The idea of forced diffusion was introduced by Reddish (1971) to account for a high degree of circumstellar reddening in OB stars which seemed to be devoid of local H II regions. Radiation from the star desorbs atoms or molecules from grains towards the star, whereas adsorption is isotropic. Thus

a net momentum is given to the grain by the arrival of the photon, far in excess of the momentum carried by the photon. Eventually a limiting drift velocity through the gas is achieved, and the dust migrates with respect to the gas.

We have used the radiation field given by Witt and Johnson (1973) to calculate the trajectories of particles shielded on one side by a filament of optical depth τ (Brand, 1975).

The dust grain is subjected to subsonic drag force, radiation pressure, and the desorption momentum force compared with which radiation pressure may be neglected in present circumstances. The UV radiation field near the galactic plane is 0.01 photons cm^{-3}, and taking the mean energy per photon to be 6 eV, with the assumption that the majority of the energy of the photon goes to the desorbed particle, the desorption velocity is $V = 20\,km\,s^{-1}$ for hydrogen molecules. Conditions are assumed so that Coulombic drag is not dominant (i.e. not in an H II region) and the desorption rate equals the adsorption rate. A grain at the edge of a dense cloud is in an anisotropic radiation field, which causes more desorption from one side than the other, and if the cloud has a very large optical depth, the anisotropy of desorption of momentum is such that momentum imparted to the grain per second is one sixth of the total momentum desorbed per second.

In these circumstances, the equation of motion of the grain is

$$m_g \dot{v} = \dot{m}_d [V/6 - A v] \tag{4}$$

where m_g is the grain mass, v the grain velocity, $\dot{m}_d$ the mass desorption rate and A is approximately the ratio of the coefficient K to the sticking probability (which is approximately unity). Thus since K is of the order unity, so is A. K is given by

$$F_{\text{drag}} = K \pi a^2 \varrho c_{\text{ad}} v \tag{5}$$

where a is the grain radius and ϱ and c_{ad} are the density and adiabatic sound speed of the ambient gas.

The solution of the equation of motion is

$$v = v_0 [1 - \exp(-t/t_d)] \tag{6}$$

where $v_0 = V/6A$, approximately 2 km s^{-1} and $t_d = m_g/A\dot{m}_d$. Since $\dot{m}_d$ equals the arrival rate of photons times a desorption probability (estimated at 0.1) times the desorbed particle mass, we can calculate $t_d = 6000$ years for grain radius $a = 0.1$ μm. Thus the time to reach terminal velocity may be too great for the process to be effective during the McCray mechanism, since the cooling time (time for dissipation of filaments) may be as small as, or smaller than t_d.

In the case of the Rayleigh-Taylor instability, however, the falling filaments are subjected to the process of forced diffusion at the full rate for most of their lifetimes. By the time the filaments have reached the galactic plane, they may have dissipated as far as the gas is concerned, but the dust cores will be permanent entities. The magnetic field will be stretched along the filament as a consequence of flux freezing and so the filaments will approximately map the field lines.

Recent estimates of Type II supernova energetics (Woltjer, 1972, 1974) and of numbers in our galaxy (Cox and Smith, 1974; Ilovaisky and Lequeux, 1972; compare also Tammann, 1974) indicate that a reasonable proportion of the interstellar medium (perhaps most of it) at one time or another is processed through a supernova. Some of the material will be lifted out of the plane, coherently since timescales of remnant expansion are much smaller than Rayleigh-Taylor timescales, and will then be susceptible to the instability.

6. Conclusion

We have presented further evidence for two types of structure in the interstellar medium, loops (appendix) and dust filaments. Given that supernova expansion processes much of the interstellar medium, we suggest that both types of structure can be explained as old supernova shells in various stages of dissipation and re-entry to the galactic plane. The dust components of such filaments formed either during the supernova expansion or by Rayleigh-Taylor instability while falling into the plane, may be stabilised by the "dust migration" process presented here.

We hope with future work on the ESO/SRC Sky Survey plates using the COSMOS measuring machine to be able to quantify the filamentary nature of the dust clouds in the plane, and possibly to study the large scale geometry of them in relation to regions of star formation such as H II regions.

Acknowledgements. We should like to thank the U.K. 48-inch Schmidt Telescope Unit for help and for the supply of one of the plates used here, and the Royal Observatory Edinburgh for the use of glass copies of the ESO Schmidt plates. We thank Miss M. Lumsden for help in the preparation of the typescript. One of us (W.J.Z.) is supported by an SRC grant.

Appendix

Here we describe 28 possible shell structures (Table 1) seen as emission regions, absorption lanes or radio features. Each formation is named by the constellation in which it is found and given a running number within that constellation.

(i) ARA *1 and 2*

Westerlund (1960) drew attention to the complex of dark and bright nebulosities centred near the star HD 148937. NGC 6164, 6165 lie within 3′ of this member of Ara OB 1 association while 14′ to the south east lies an arc of nebulosity. Also centred on HD 148937 there exists a nearly circular arc of bright nebulosity 2°

Table 1. Loop structures described in appendix

Feature	l^{II}	b^{II} (degrees)	diameter	Possible associations	D (pc)	d (pc)
Emission Loops						
CYG 3	73	2	2	Cyg OB3	2300	80
CYG 1	76	1	3	Cyg OB1	1700	89
CYG 2	78	1.5	2	Cyg OB9	1170	41
CYG 4	80	1	15	Cyg OB2	1500	402
ORI 1	206	−19	14 × 12	Ori OB1	500[a])	120 × 105
					300[b])	75 × 64
MON 2	206	0	3	Mon OB2	1400	73
MON 1	208	1	4	Mon OB1	715	50
CMA 1	223	0	3	CMa OB1	1315	69
ARA 1	336	0	2	Ara OB1	1300	45
SCO 1	343	0	4	Sco OB1	1400	100
SCO 4	349	−1	1	RCW 123		
Dust Loops						
SCU 1	25	−3	10	Sct OB2	730	128
CYG 5	80	1	8	Cyg OB2	1500	210
CEPH 1	103	4	6	Cep OB1	700	74
PER 1	164	−13	15	Per OB2	400	107
ORI 2	195	−11	7	λ Ori	400	50
CAR 1	314	−1	10			
ARA 2	337	0	5	contacts ARA 1	1300	113
SCO 2	343	0	7		1400	171
SCO 3	343	0	14		1400	349
SCO 5	350	−1	8			
Other Loops						
VUL 1	60	0	4	21 cm Radio	2000	140
LOOP II	100	33	91	Continuum	110	175[c])
LOOP III	124	16	65	Continuum	150	200[c])
ORIGEM	195	0	4.6	Continuum	1000	100[c])
LOOP IV	315	49	40	Continuum	250	210[c])
LOOP I	329	18	116	Continuum	130	230[c])
LUPUS I	330	15	4.6	Continuum	400	32[d])

[a]) Ruprecht (1964).
[b]) Johnson (1964).
[c]) Values from Berkhuijsen (1973).
[d]) Values from Spoelstra (1973).

in diameter intersecting NGC 6188. Here also are visible dust lanes and in the H_α atlas (Rodgers *et al.*, 1960a) this outer circle of nebulosity appears to be ringed by dust. A further feature of these prints is the dark ring, ARA 2, 5° across contacting ARA 1 in the south west.

Both the Ara OB 1 association and that of Sco OB 1 lie at 1400 pc (Ruprecht, 1964). It seems likely that the structures ARA 1 and ARA 2 are associated in space with those of SCO 1, 2 and 3 described later.

(ii) CAR *1*

A large dust arc, almost a ring 10° in diameter can be seen in the H_α atlas (Rodgers *et al.*, 1960a) and on the prints of the Georgetown Atlas (1952). No associations are known in the area. It seems possible that there is some connection between it and the Coalsack.

(iii) CEPH *1*

In this area semi-circular filamentary dust lanes can be seen forming the southern half of a 6° ring. The northern half is traced by faint emission nebulosity. This area van den Bergh (1967) classifies as a fossil H II region, and here Simonson III (1973) notes a deficiency in neutral hydrogen. To the west the dust seems to interact with IC 1396 where bright rims appear. This nebulosity is placed at 820 pc and associated with Cep OB 2 (Ruprecht, 1964). Blaauw (1961) finds 2 runaway stars 6809 Cyg and λ Ceph originated here.

(iv) CMA *1*

An optical nebula can be seen in the H_α atlas (Rodgers *et al.*, 1960a), taking the appearance of a half-loop 3° in diameter. In the same area at 1315 pc lies the associ-

ation CMa OB 1 and the cluster NGC 2353 (Ruprecht, 1964). The association is included in van den Bergh's (1966) list of reflection associations.

(v) CYG *1*

Seen on the Palomar Sky Survey prints as an almost complete ring of H_α filaments 3° in diameter, this region lies in the same direction as Cyg OB 1. Within the ring lies the Wolf Rayet star HD 192163 and its large associated nebula NGC 6888. Also in the region is found P Cyg and it is believed that the majority of OB stars here lie at 1700 pc (Ruprecht, 1964).

(vi) CYG *2*

On Palomar Survey prints there exists a 2° filamentary emission structure. Radio observations (Dickel *et al.*, 1969; Winzer and Roberts, 1971) show a ringlike emission region coinciding with the optical ring. At 1170 pc in this region lies the association Cyg OB 9.

(vii) CYG *3*

Seen close to CYG 1 this region suggests a filamentary H_α ring 2° in diameter.The northern edge is coincident with an absorption lane also seen to affect CYG 1. In this area lies Cyg OB 3 at 2300 pc.

(viii) CYG *4 and 5*

The extensive H_α emission filaments existing up to 7° from the galactic plane have been mentioned by Dickel and Wendker (1969). The central region of this filamentary ring is close to the heavily obscured association Cyg OB 2. Evidence for an 8° dust ring can be seen in the lane crossing NGC 7000 in the east and is particularly well seen in Morgan *et al.* (1955) prints. Johnson (1971) has compared the whole Cygnus complex with that seen in Vela.

(ix) MON *1 and 2*

The first complete description of the Monoceros complex is found in Morgan *et al.* (1955). NGC 2244, "the Rosette", is seen projected on to the southern edge of a gaseous emission ring 3° in diameter. Part of a second 4° emission ring is seen centred on NGC 2264 to the north.

The accepted distance of NGC 2264 is 715 pc and it is part of the association Mon OB 1. NGC 2244 has been placed at 1400 pc (Ruprecht, 1964) although more recent work finds a distance closer to 800 pc (Morgan, Osterbrock *et al.*, 1965). It is part of the Mon OB 2 association.

(x) ORI *1*

The main feature of this complex is the structure known as Barnard's Loop. H_α filter photography reveals an elliptical ring of nebulosity 14° × 12° along its major axes, (Morgan *et al.*, 1955). Ultraviolet observations led O'Dell and York (1967) to find evidence for scattering within the Loop, as well as for a possible neutral hydrogen region outside it. The Loop observed in the ultraviolet has an extent 19° × 14°. Isobe (1973) too finds that the H II loop occurs at the inner edge of the H I region. Gordon (1970) finds the neutral hydrogen to be in the form of a shell rotating about $l = 209°$, $b = -19°.4$ and that this is not the centre of the Loop.

Projected against the Loop's centre are the nebulosities surrounding the Trapezium and σ Ori. The former is the site of several compact H II regions and is believed to be very young. The latter is much obscured but exhibits a Strömgren sphere 2° in diameter (Morgan *et al.*, 1955). The association Ori OB 1 is found in this region at a distance of 300 pc (Ruprecht, 1964). Within the association exist at least 3 sub groups: the Trapezium, the Belt and a third which is the oldest of the groupings (Blaauw, 1964). Blaauw (1961) suggests that the association Ori OB 1 is the site of formation of several runaway stars, including AE Aur and μ Col. From this one might infer supernova events within this region 4.9, 2.7 and 2.2 million years ago.

(xi) ORI *2*

Centred on λ Ori this complex is best seen on the plates of Morgan *et al.* (1955). Surrounding the 7° spherical H_α nebulosity Wade (1958) finds a dark obscuring region $1\frac{1}{2}°$ thick which may also be seen in Khavtassi's atlas (1960). Coincident with the ring, Wade (1958) finds a shell of neutral hydrogen with a mass of about $45000\,M_\odot$ expanding at $8\ \mathrm{km\ s^{-1}}$. The mass of H II is about $2000\,M_\odot$. In the densest parts of the obscuring ring are found H_α emission stars (Wade, 1958).

If λ Ori is placed 400 pc distant, the diameter of the shell is 50 pc.

(xii) PER *1*

Maps of dust distribution by both Lynds (1969) and Khavtassi (1960) reveal an area 10° in diameter, of reduced obscuration surrounded by a ring of obscuring matter. Heeschen (1951) using star counts finds most of the obscuration to be at 200–300 pc. Observations of neutral hydrogen show a minimum about $l = 164°$, $b = -17°$ (Sancisi, 1970). Within the dust ring at 300–400 pc lies the OB association Perseus OB 2 (Ruprecht, 1964) and associated with it the emission nebula NGC 1499. Van den Bergh (1966) states that Perseus OB 2 is an "*R*" association.

(xiii) SCO *1, 2, 3*

The constellation of Scorpius is a complex region when viewed in H_α light. Gum (1955) drew attention to the "large roughly semicircular loop" centred on NGC 6231, commenting that this was the nucleus of the "O" association Sco OB 1. In the Bok, Bester and Wade (1955) catalogue the whole nebulosity is almost a closed ring (BBW 31100) 4° in diameter.

Outside this central emission ring are two further dust shells 7° and 14° in diameter. The dust appears filamentary on the Palomar Sky Survey. To the east of the H_α ring lies IC 4628, said by Westerlund (1960) to be like a supernova remnant. Within 2° of NGC 6231 Morgan *et al.* (1953) find 70 OB stars belonging to Sco OB 1, for which Ruprecht (1964) gives a distance of 1400 pc. Connected with this association are 3 runaway stars, HD 41534, 152408, 151397 (Blaauw, 1961) and also nearby is the source Sco X 2 (Braes and Hovenier, 1966).

(xiv) SCO *4 and 5*

SCO 4 is seen as a ringlike emission nebula on red Sky Survey plates. Noted by Rodgers *et al.* (1960b) as RCW 123 it is described in their catalogue as ring-like. Surrounding this nebulosity is a 8° diameter dark ring of material well seen on plates by Steavanson (RAS slides nos. 427, 428). This appears to pass between NGC 6334 and 6357 (Lindsay, 1955), both situated at about 1600 pc. Thus the ring structure is at or within that distance. Like ARA 1 and 2, it may be situated close in space to SCO 1, 2, 3.

(xv) SCU *1*

A series of dust arches (possibly 3 different events) showing clearly in the H_α atlas (Rodgers *et al.*, 1960a) on Plate 5, surrounding Scutum OB 2 in the plane.

(xvi) VUL *1*

Finally we mention the "explosive event" observed at 21 cm by Katgert (1969). The H I loop structure is 4° in diameter and at a probable distance greater than 2000 pc. In this direction lies Vul OB 1 at 2050 pc and the Palomar Sky Survey prints reveal much filamentary dust extending for several degrees.

References

Bergh, S. van den 1966, *Astron. J.* **71**, 990
Bergh, S. van den 1967, *Roy. Astron. Soc. Canada* **61**, 23
Bergh, S. van den 1972, in A. Beer (Ed.), Vistas in Astr. **13**, 265
Berkhuijsen, E. M., Haslam, C. G. T., Salter, C. J. 1971, *Astron. & Astrophys.* **14**, 252
Berkhuijsen, E. M. 1973, *Astron. & Astrophys.* **24**, 143
Blaauw, A. 1961, *Bull. Astron. Inst. Neth.* **15**, 280
Blaauw, A. 1964, *Adv. Astron. & Astrophys.* **2**, 213
Bok, B. J., Bester, M. J., Wade, C. M. 1955, Harvard College Repr. No. 416
Brand, P. W. J. L. 1975, *Astrophys. Space Sci.* (in press)
Braes, L. L. E., Hovenier, J. W. 1966, *Nature* **209**, 360
Churchwell, E., Walmsley, C. M. 1973, *Astron. & Astrophys.* **23**, 117
Cox, D. P., Smith, B. W. 1974, *Astrophys. J. Letters* **189**, L 105
Dickel, H. R., Wendker, H., Bieritz, J. H. 1969, *Astron. & Astrophys.* **1**, 276
Downes, D. 1971, *Astron. J.* **76**, 305
Field, G. B. 1965, *Astrophys. J.* **142**, 531
Georgetown 1952, Georgetown College Obs. Photographic Atlas of the Southern Milky Way
Giacconi, R., Murray, S., Gursky, H., Kellog, E., Schreier, E., Matilisky, T., Koch, D., Tananbaum, H. 1974, *Astrophys. J. Suppl.* **27**, 37
Gordon, C. P. 1970, *Astron. J.* **75**, 914
Gum, C. S. 1955, *Mem. Roy. Astron. Soc.* **67**, 155
Gum, C. S., de Vaucouleurs, G. 1953, *Observatory* **73**, 152
Haslam, C. G. T., Kahn, F. D., Meaburn, J. 1971, *Astron. & Astrophys.* **12**, 388
Hayward, R. 1964, *Publ. Astron. Soc. Pacific* **76**, 35
Heeschen, D. S. 1951, *Astrophys. J.* **114**, 132
Hindman, J. V. 1967, *Australian J. Phys.* **20**, 147
Hovaisky, S. A., Lequeux, J. 1972, *Astron. & Astrophys.* **20**, 347
Isobe, S. 1973, in J. M. Greenberg and H. C. Van de Hulst (Eds.) Interstellar Dust and Related Topics, IAU Symposium No. 52, p. 433
Johnson, H. M. 1964, Trans. IAU 12 B, 446
Johnson, H. M. 1971, NASA SP 332, The Gum Nebula and Related Problems, p. 12
Katgert, P. 1969, *Astron. & Astrophys.* **1**, 54
Khavtassi, J. S. 1960, An Atlas of Galactic Dark Nebulae, Abastumani Astrophys. Obs.
Lin, C. C. 1970, in W. Becker and G. Contopoulos (Eds.), The Spiral Structure of Our Galaxy, IAU Symposium No. 38, p. 377
Lindsay, E. M. 1953, *Irish Astron. J.* **2**, 210
Lindsay, E. M. 1955, *Irish Astron. J.* **3**, 182
Lynds, B. T. 1962, *Astrophys. J. Suppl.* **7**, 1
Lynds, B. T. 1969, *Publ. Astron. Soc. Pacific* **81**, 496
Lynds, B. T. 1970, in W. Becker and G. Contopoulos (Eds.), The Spiral Structure of Our Galaxy, IAU Symposium No. 38, p. 26
Lynds, B. T. 1972, in D. E. Evans (Ed.), External Galaxies and Quasi-stellar Objects, IAU Symposium No. 44, p. 56
McCray, R., Schwarz, R. F., Stein, J. 1972, *Astrophys. J. Letters* **177**, 75
McGee, R. X., Milton, J. A. 1964, IAU/URSI Symposium No. 20, 289
Meaburn, J. 1967, *Z. Astrophys.* **65**, 93
Morgan, W. W., Gonzalez, G., Gonzales, G. 1953, *Astrophys. J.* **118**, 323
Morgan, W. W., Strömgren, B., Johnson, H. M. 1955, *Astrophys. J.* **121**, 611
Morgan, W. W., Hiltner, W. A., Neff, J. S., Osterbrock, D. E. 1965, *Astrophys. J.* **142**, 974
Nail, V. McK., Shapley, H. 1953, *Proc. Nat. Acad. Sci. Am.* **39**, 358
O'Dell, C. R., York, D. G., Henize, K. G. 1967, *Astrophys. J.* **150**, 835
Öpik, E. J. 1953, *Irish Astron. J.* **2**, 219
Parker, E. N. 1966, *Astrophys. J.* **145**, 811
Reddish, V. C. 1971, *Nature* **232**, 40
Rodgers, A. W., Campbell, C. T., Whiteoak, J. B., Bailey, H. H., Hunt, V. O. 1960a, An Atlas of H_α Emission in the Southern Milky Way, Mount Stromlo Obs.
Rodgers, A. W., Campbell, C. T., Whiteoak, J. B., Bailey, H. H., Hunt, V. O. 1960b, *Monthly Notices Roy. Astron. Soc.* **121**, 103
Ruprecht, J. 1964, *Trans.* IAU **12B**, 336
Sancisi, R. 1970, *Astron. & Astrophys.* **4**, 387
Sandage, A. 1961, The Hubble Atlas of Galaxies, Mount Wilson and Palomar Obs.
Simonson III, S. C. 1973, in J. M. Greenberg and H. C. Van de Hulst (Eds.), Interstellar Dust and Related Topics, IAU Symposium No. 52, p. 227

Sivan, J. P. 1974, *Astron. & Astrophys. Suppl.* **16**, 163
Spoelstra, Th. 1973, *Astron. & Astrophys.* **24**, 149
Tammann, G. A. 1974, in C. B. Cosmovici (Ed.), Supernovae and Supernovae Remnants, published Reidel, Dordrecht, Holland, p. 155
Verschuur, G. D. 1974, *Astrophys. J. Suppl.* **27**, 65
Vorontsov-Velyaminov, B. A. 1964, *Soviet Astron.* **7**, 63
Wade, C. 1958, *Rev. Mod. Phys.* **30**, 946
Westerlund, B. E. 1960, *Arkiv. Astron.* **2**, 467
Westerlund, B. E. 1966, Symposium of Radio and Optical Studies of the Galaxy, Mount Stromlo Obs. p. 78
Westerlund, B. E., Mathewson, D. S. 1966, *Monthly Notices Roy. Astron. Soc.* **131**, 371
Winzer, J. E., Roberts, J. A. 1971, *Roy. Astron. Soc. Canada J.* **65**, 100
Witt, A., Johnson, N. W. 1973, *Astrophys. J.* **181**, 363
Woltjer, L. 1972, *A. Rev. Astron. & Astrophys.* **10**, 155
Woltjer, L. 1974, ESO/SRC/CERN. Conference on Research Programmes for Large Telescopes

P. W. J. L. Brand
W. J. Zealey
Department of Astronomy
University of Edinburgh
Royal Observatory
Edinburgh EH9 3HJ
United Kingdom

Mon. Not. R. astr. Soc. (1982) 198, 545–562

Spiral arms and a supernova-dominated interstellar medium

P. W. J. L. Brand and S. R. Heathcote *University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

Received 1981 April 15; in original form 1980 December 17

Summary. Models of the interstellar medium (ISM) utilizing the large energy output of supernovae to determine the average kinematical properties of the gas, are subjected to an imposed (spiral) density wave.

The consequent appearance of the ISM is considered. In particular the McKee–Ostriker model with cloud evaporation is used, but it is shown that the overall appearance of the galaxy model does not change significantly if a modification of Cox's mechanism, with no cloud evaporation, is incorporated.

We find that a spiral density wave shock can only be self-sustaining if quite restrictive conditions are imposed on the values of the galactic supernova rate and the mean interstellar gas density.

1 Introduction

Since the work of Cox & Smith (1974) and its subsequent elaboration by McKee & Ostriker (1977, hereafter MO), it has been clear that supernova remnants (SNRs) play a major role in determining the state of the interstellar medium (ISM). Consequently some fraction of its volume must be occupied by a very hot, tenuous plasma. However, observational estimates of this fraction, mostly referring to the solar neighbourhood, vary from 10 per cent (Heiles 1980) to 80 per cent (Dwek & Scalo 1979; Higdon & Lingenfelter 1980). Furthermore, as stressed by Scott, Jensen & Roberts (1977), gas in such a hot state, occupying a large volume fraction throughout the galactic disc, might, by virtue of its high sound speed, prevent the formation of the strong shock waves, often invoked (Roberts 1969) to initiate star formation in spiral arms.

In response to this difficulty various attempts have been made recently, either to explain spiral structure independently of density waves (Seiden & Gerola 1979) or by only indirect appeal to density waves (Cowie 1980). Others have avoided the problem by suggesting that the hot component has a large volume filling factor only in spiral arms, the ISM between the arms remaining in a state capable of supporting density wave shocks (Scott *et al.* 1977; Reinhardt & Schmidt-Kaler 1979; Blitz & Shu 1980). Alternatively as suggested by Schmidt-Kaler & Weigandt (1980) the difficulty may not exist if the effective sound speed in the inhomogeneous medium is determined largely by the 'warm' component. This is discussed in Section 5.

In this paper we investigate, within the framework of the MO model, the consequences of assuming that (1) the pattern of the total mean density is similar to that predicted by two-armed spiral shocks, and (2) that the star formation rate and supernova rate are proportional to each other, and to some power of the local mean density. In Section 2 we shall briefly discuss modifications to the MO model necessary to permit its use in a time-dependent situation and describe the method we employ to calculate the variations of the supernova rate, ionizing photon emissivity and mean density. In Section 3 we describe the results for a spiral galaxy like the Milky Way. In Section 4, we discuss the importance of the cloud evaporation assumption, and by modifying Cox's (1979) discussion, show that similar results are predicted with a non-evaporative model. In the final section we consider the consequences of our results for galactic structure and examine what observational constraints might be imposed on the parameters of the model.

2 Assumptions

MO considered the consequences of the inhomogenous nature of the ISM for SNR evolution, within the framework of the Cox & Smith picture and found that, with the introduction of a number of simplifying assumptions, it was possible to develop a self-consistent model conforming with many of its observational properties. They picture the ISM as consisting of three more or less distinct components: (1) a hot low density plasma (HIM), filling 70–80 per cent of the interstellar volume, (2) cold, dense, neutral clouds (CNM) imbedded within the HIM and (3) warm, intermediate density, photo-ionized cloud coronae enveloping each cloud (WIM). The filling factor of the WIM is ~ 20 per cent while that of the CNM is only ~ 2 per cent although it contains most of the mass in the ISM. All three components are supposed to be in rough pressure equilibrium.

A fundamental ingredient in the MO scheme is the assumption that the thermal evaporation of clouds engulfed by a young SNR is sufficiently rapid for radiation to become important, permitting a transition to the 'snow-plough' phase, before either the external pressure becomes important, or the SNR radius becomes comparable with the scale height of the disc gas. For the time being we will assume that this is in fact the case, discussion of the opposite point of view advanced by Cox (1979) being postponed until Section 4. We will, however, include the effects of cloud geometry and of the dynamical response of the clouds to their surroundings in our model.

2.1 CLOUD DYNAMICS AND THE EVAPORATION RATE

From the work of Cowie & McKee (1977) and Cowie & Songaila (1977) the evaporation rate for a static spheroidal cloud imbedded in a medium of temperature T_h is given by,

$$\dot{M}_{ev} = 10^{4.44}\,\phi_m\,\phi_g(l)\,T_h^{5/2}\,a\,. \tag{1}$$

Here, a(pc) is the radius of a sphere having the same volume as the cloud and $\phi_g(l)$ is a geometrical factor which is unity for spherical clouds and increases with increasing major–minor axis ratio, l, for oblate or prolate spheroids. The constant ϕ_m is introduced to allow for partial quenching of thermal conduction by magnetic fields or turbulence. This equation ceases to be valid if the mean free path for electron energy exchange in the cloud becomes comparable to its dimensions ('saturation'), or if the rate of radiative cooling in the cloud envelope exceeds the rate of conductive heating so that the cloud grows by condensation. The saturated condition will always occur at a sufficiently early epoch in SNR evolution and its neglect leads to an overestimate of the net evaporation rate over the remnant's lifetime.

This effect is, however, relatively unimportant over the range of conditions considered here. Similarly, there are always some larger clouds in which radiative losses are important, however, with the steep cloud mass spectrum implied by observation (e.g. Hobbs 1974) and adopted here this is never important.

The surface pressure experienced by a cloud which is overrun by an SNR blast wave will increase sharply as the shock passes and then slowly decline as the remnant expands. The cloud's response to this changing surface pressure will be extremely complex (*cf.* Sgro 1975; McKee & Cowie 1975; Woodward 1976, 1979; Cox 1979) and large departures from pressure equilibrium are inevitable. Nevertheless, the cloud must relax back toward pressure equilibrium in a few times the cloud soundspeed crossing time τ_{cl}. Cowie & McKee (1977) found that in general the time-scale for evaporative flow in the cloud is very much shorter than τ_{cl} so that the cloud can be treated quasistatically and equation (1) will give the correct instantaneous evaporation rate.

If $\tau_{cl} \gg \tau_{SNR}$ (the mean interval between successive young SNRs overrunning a given cloud) *or* if evaporation plays an important role in SNR dynamics for a period $\tau_{ev} \ll \tau_{cl}$, then we can regard the cloud radius as a constant determined by the time averaged pressure it experiences. Conversely, if $\tau_{ev} \gg \tau_{cl}$ the cloud must remain close to pressure equilibrium with its immediate surroundings. Neglecting magnetic support and assuming the cloud remains isothermal we find in this latter case that the cloud radius a, varies with pressure P, according to

$$a(P) \propto P^{-\gamma}, \tag{2}$$

where $\gamma \simeq 1.8$ if recombination is sufficiently rapid to keep the cloud in ionization equilibrium or $\gamma = 1/3$ if the ionization state remains frozen. In practice we find that τ_{cl}, τ_{ev} and the recombination time-scale, τ_{rec}, are all of comparable magnitude so none of these cases is likely to be applicable in detail. Nevertheless, equation (2) with $\gamma = 0$, $1/3$ and 1.8 may provide some general indication of the effects of cloud dynamics on the evaporation rate.

With the evaporation rate given by (1) and (2), a straightforward extension of the work of MO (see Appendix) leads to a system of equations which can be solved iteratively to yield values of several quantities of interest. We have undertaken an extensive series of calculations employing different values for the various constants which enter into the MO model. In the present context we wish to focus attention primarily on those quantities which are expected to change from point to point as a consequence of galactic structure. Consequently we have set the remaining quantities at the values chosen by MO, (these values are tabulated in the Appendix) except that we use supernova energy $E_0 = 5 \times 10^{50}$ erg, half the value used by MO. Our calculations confirm, however, that our results are not unduly sensitive to the choice of these other parameters. In particular we find that the cloud geometry chosen is unimportant and hence consider only the case of spherical clouds. Further, in what follows we will take $\gamma = 1/3$ although our results are not too sensitive to its value.

2.2 THE MODEL OF THE GALAXY

The quantities we have chosen to vary in our model are the local mean total density $\bar{n}$, the emissivity of ionizing photons ϵ_{uv}, and the supernova rate per cubic parsec S. MO selected values for these parameters based on observational data pertaining to the solar neighbourhood. Here we calculate their values by means of a simple semi-empirical galactic model, constructed within the framework of density wave theory.

In the TASS model (Roberts 1969, Shu, Milione & Roberts 1973) the variation of the gas

density along a representative streamline can be approximated by

$$\bar{n}(R,t) = \bar{n}_0(R)\ \{1 + (c-1)\exp(-t/t_0)\}\ , \tag{3}$$

where t is the time since the last passage through the density wave shock and the decay time t_0 is of the order 3×10^7 yr. For simplicity we take c the density contrast or compression ratio, to be constant.

We obtain $\bar{n}_0(R)$ from the relation

$$\langle \bar{n}(R) \rangle = \int_0^{t_{\rm gal}(R)/2} \bar{n}(R,t)dt \simeq \bar{n}_0(R)\ \left\{1 + 2(1-c)\frac{t_0}{t_{\rm gal}(R)}\right\} \equiv \bar{n}(R,\mathrm{H}^0) + 2\bar{n}(R,\mathrm{H}_2),$$

where $\bar{n}(R,\mathrm{H}^0)$ and $\bar{n}(R,\mathrm{H}_2)$ are the observed azimuthally smoothed densities of neutral and molecular hydrogen respectively. Here $t_{\rm gal}$ is the galactic rotation period in a frame rotating at the pattern speed of the spiral potential $\Omega_{\rm p}$, i.e.

$$t_{\rm gal}(R) = 2\pi/[\Omega(R) - \Omega_{\rm p}]\ ,$$

where $\Omega(R)$ is the angular rotation speed of the galaxy. In what follows $\Omega_{\rm p}$ is taken to be 13.5 km s^{-1} kpc^{-1} and we adopt a flat rotation curve such that $\Omega(R) = 250/R$ km s^{-1} kpc^{-1}.

While the galactic distribution of H^0 is quite well established (Burton 1976) that of H_2 must be inferred indirectly from observations of CO. Both the local value of the ratio $\bar{n}(\mathrm{H}_2)/\bar{n}(\mathrm{CO})$ and the question of a possible radial variation of this quantity (due to a galactic abundance gradient) remain controversial (Solomon, Sanders & Scoville 1979 (SSS); Guibert, Lequeux & Viallefond 1978 (GLV); Blitz & Shu 1980). We therefore consider three different H_2 distributions based on (i) the data of SSS, (ii) the data of Gordon & Burton (1976; GB) and (iii) the values obtained by GLV from the data of GB by assuming a smaller value for $\bar{n}(\mathrm{H}_2)/\bar{n}(\mathrm{CO})$ in the solar neighbourhood and incorporating a strong radial gradient of the carbon abundance. For the vertical distribution of both H^0 and H_2 we adopt a Gaussian form with scale heights of 160 and 70 pc respectively (Cohen & Thaddeus 1977).

The relationship between the gas density and the rate of star formation has been a subject of considerable controversy since the classic work of Schmidt (1959) who proposed on empirical grounds that on a large scale the star formation rate varies as some power m of the total gas density. Recently, GLV have used a comparison of the radial and vertical distribution of several tracers of recent massive star formation with that of the interstellar gas and found some justification for adopting $m \simeq 1.3$–2.5. A similar study by Smith, Biermann & Mezger (1978) based primarily on the distribution of giant H II regions led to a somewhat smaller value in the range 0.5–1.0. Conversely from consideration of the past history of star formation in the solar neighbourhood assuming a time independent initial mass function Miller & Scalo (1979) have deduced $m \lesssim 0.5$. We here employ the Schmidt law as a convenient parameterization relating the *local* star formation rate to the *local* mean total density and construct models employing $m = \frac{1}{2}$, 1 and 2. It should, however, be stressed that the observational tests refer only to properties smoothed over large regions.

We thus take the supernova rate to be given by

$$S(R,t) = S_0 \bar{n}(R, t - t_*)^m \qquad m = \tfrac{1}{2}, 1 \text{ or } 2\ , \tag{4}$$

where t_* is the 'lifetime' of the stars which give rise to supernova events, which we take to be of order 10^7 yr, comparable to observational estimates of the delay between the formation of a dust lane and the appearance of the first massive OB stars in external galaxies (Mathewson, van der Kruit & Brouw 1972). To obtain the constant of proportionality, S_0, we equate the integral of $S_0 \bar{n}(R,t)^m$ over the whole disc to the value of the total galactic supernova rate, $S_{\rm gal}$. Estimates for this latter quantity have been obtained by a number of

independent methods, but all are subject to severe selection effects and/or ambiguities of interpretation (see, e.g. Clark & Stephenson 1977, Tammann 1978; Caswell & Lerche 1979a; Lozinskaya 1980; Higdon & Lingenfelter 1980; Taylor & Manchester 1977; Arnaud & Rothenflug 1980). Taken together all the available data appear to be consistent with $S_{gal} \sim 0.01-0.1\,yr^{-1}$. Similarly we obtain the emissivity of ionizing photons from the simple relationship

$$\epsilon_{uv}(R,t) = \epsilon_1 S(R,t) + \epsilon_2, \tag{5}$$

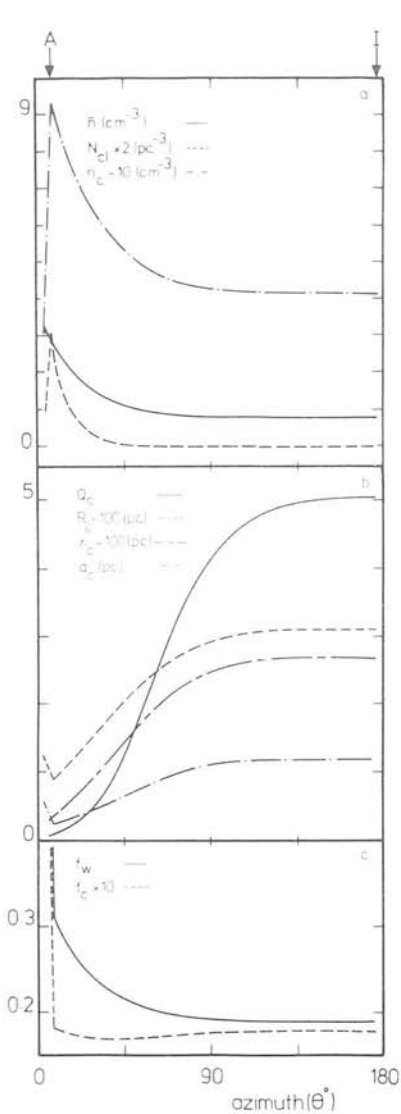
where the term proportional to S represents contribution from SNRs themselves and young hot stars, while the constant term describes the contribution from density bounded planetary nebulae. Observational estimates in the solar neighbourhood suggest that SNRs, unshielded B stars and planetary nebulae each provide $\epsilon \sim 1-1.5\times10^{-15}\,cm^{-3}\,s^{-1}$ (Salpeter 1979). Potentially the local contribution from unshielded O stars is a factor of 10–20 greater than this; however, Elmegreen (1976) has argued that most of this flux is absorbed within ~ 100 pc of its source. Neglecting this latter contribution and allowing for losses from the galactic disc we take $\epsilon_1 \sim 1.2\times10^{61}$ per SNR and $\epsilon_2 \sim 0.7\times10^{-15}\,cm^{-3}\,s^{-1}$; however, it must be recognized that a 'leak' of only 10 per cent of the O star flux would double this value.

Equations (3), (4) and (5) define our galactic model. In incorporating it into the three phase scheme we tacitly assume that the ISM remains in a quasi-steady state as S, $\bar{n}$ and ϵ_{uv} vary slowly with time and that the spatial variation of these parameters on length scales very much larger than the typical dimensions of an SNR do not drastically affect local conditions. Smith (1977) found that a quasi-equilibrium state was achieved in his Monte Carlo simulations based on the Cox & Smith (1974) model on a time-scale of order 10^7 yr. This relaxation time and the statistical distribution of supernovae will inevitably somewhat blur the parameter distributions we obtain. Nevertheless, solutions obtained assuming locally uniform values for S, $\bar{n}$ and ϵ_{uv} at each point in the disc, should not represent too severe an approximation except near the discontinuities at the passage of the density wave shock and the 'switch on' of the new generation of supernovae.

3 Results

For purposes of orientation we will begin by considering a specific model incorporating the BG density distribution and having spiral shock compression $c = 5.0$, exponent of star formation rate $m = 1.0$ and $S_{gal} = 0.02\,yr^{-1}$ (note that S_{gal} is the number of supernovae occurring per year in our Galaxy while S is the supernova rate per year per cubic parsec). For this case, Fig. 1 shows the variation of a number of parameters of the ISM with azimuth (θ) in the midplane of the galaxy, at the galactocentric distance of the Sun. The values of the same quantities at points on the loci of maximum supernova rate ('arm') and minimum mean density ('interarm') are shown as functions of galactocentric distance in Fig. 2.

In Table 1(a) we list the 'arm' and 'interarm' values both for $R = 10$ kpc and $R = 5$ kpc, the latter distance being chosen as broadly representative of points in the molecular ring. Column 1 gives Q_c, the porosity parameter, which measures the sum of the fractional volumes of all SNRs still in the pre-cooling phase of their evolution while column 2 gives R_c, the SNR radius at the end of this phase. These quantities, especially the former, increase very rapidly with *both* decreasing $\bar{n}$ *and* decreasing S. This behaviour is a direct consequence of the evaporative feedback mechanism in the MO scheme and is not dependent on the details of our galactic model. When $Q_c \gtrsim 1$, as occurs in the 'interarm' region interior to 4 kpc or exterior to 8 kpc, SNRs frequently overlap before cooling and the basic assumptions of the MO model break down. Similarly if R_c exceeds the scale height of the disc gas, energy losses to the halo may play a dominant role in SNR energy balance invalidating our results.



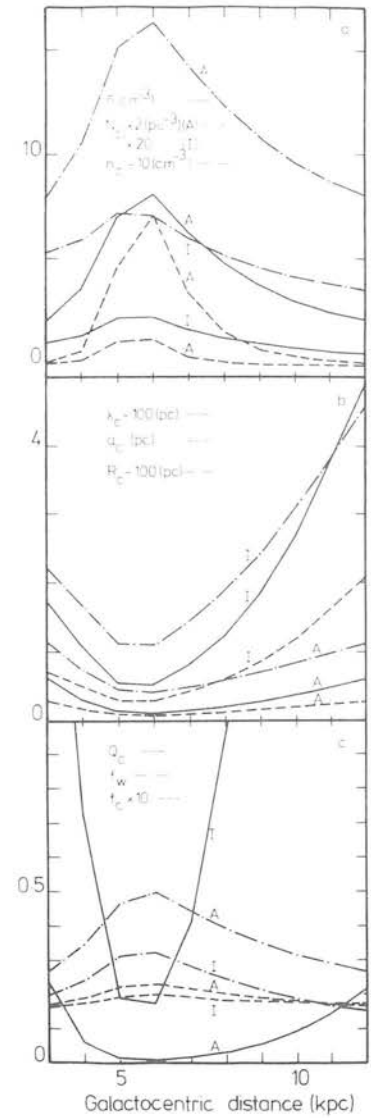


Figure 1. (Left) Calculated quantities as a function of azimuth θ°, in the midplane of the galaxy at a galactocentric distance of 10 kpc (see text). θ is zero at the position of the density wave shock. The positions at which the 'arm' (A) and 'interarm' (I) properties were evaluated are indicated. The properties of the warm envelopes may be estimated from the relationships in the text while $\langle P \rangle \approx 88 k n_c$.

Figure 2. (Right) The same quantities as in Fig. 1 but plotted as functions of galactocentric distance, R (kpc) along the 'arm' (A) and 'interarm' (I) loci.

The occurrence of either of these conditions is likely to lead to the violation of energy conservation in the disc gas. The excess energy will be transported to the halo where it may assist in driving a galactic wind or a galactic fountain (Shapiro & Field 1976; Chevalier & Oegerle 1979). Further, if dense shell formation is prevented under these circumstances, material will not be returned to the dense phase and the mass of the HIM will grow by evaporation at the expense of the clouds. The critical radius above which clouds begin to condense rather than evaporate decreases with increasing hot phase density (McKee & Cowie 1977). Thus ultimately a new equilibrium state may be achieved with evaporation of the

Table 1. Properties of the ISM for various values of supernova rate, arm compression ratio, density distribution, and star formation rate. See Section 3 for details.

	R (kpc)	$\bar{n}$ (cm^{-3})	S (10^{-13} pc^{-3} yr^{-1})	Q_c	R_c (pc)	$\langle P\rangle/k$ (10^4 K cm^{-3})	N_{cl} (pc^{-3})	$\langle a_c\rangle$ (pc)	λ_c (pc)	f_c (×100)	n_c (cm^{-3})	$\langle a_w\rangle$ (pc)	λ_w (pc)	f_w	n_w (cm^{-3})
(a) S_{gal} = 0.02 yr^{-1}, c = 5.0, m = 1.0, GB density law.															
A	5	7.01	10.69	0.01	45	1.32	2.34	0.09	13.5	4.66	150.5	0.28	1.69	0.22	1.04
I	5	2.25	2.86	0.19	112	0.63	0.06	0.30	54.6	3.12	72.1	0.93	6.36	0.19	0.45
A	10	3.04	4.64	0.09	85	0.84	0.18	0.20	38.2	3.17	96.0	0.62	4.52	0.18	0.61
I	10	0.79	0.93	5.06	310	0.37	7.4×10^{-4}	1.19	269.4	1.89	41.6	3.86	28.96	0.18	0.23
(b) S_{gal} = 0.02 yr^{-1}, c = 10.0, m = 1.0, GB density law.															
A	5	7.79	12.33	0.01	42	1.44	3.27	0.08	12.1	4.76	163.6	0.25	1.51	0.22	1 14
I	5	1.61	2.17	0.48	148	0.56	0.02	0.42	88.4	2.54	63.5	1.32	10.02	0.18	0.38
A	10	4.22	6.68	0.04	65	1.03	0.53	0.14	25.3	3.60	117.3	0.44	3.06	0.19	0.77
I	10	0.57	0.67	18.57	451	0.32	1.4×10^{-4}	2.05	479.6	1.56	36.3	6.70	50.48	0.18	0.19
(c) S_{gal} = 0.05 yr^{-1}, c = 5.0, m = 1.0, GB density law.															
A	5	7.01	26.73	0.02	39	2.54	4.89	0.06	16.6	2.43	288.3	0.18	2.00	0.12	1.89
I	5	2.25	7.15	0.26	98	1.23	0.11	0.19	69.8	1.62	139.4	0.62	7.89	0.10	0.83
A	10	3.04	11.59	0.13	74	1.63	0.33	0.13	48.9	1.64	185.6	0.41	5.60	0.10	1.12
I	10	0.79	2.32	6.56	266	0.70	1.5×10^{-3}	0.75	335.8	0.99	79.9	2.44	35.70	0.09	0.43
(d) S_{gal} = 0.02 yr^{-1}, c = 5.0, m = 1.0, SSS density law.															
A	5	16.57	13.38	0.01	28	1.41	16.38	0.06	4.3	10.60	159.5	0.19	0.58	0.43	1.17
I	5	5.32	3.58	0.02	65	0.64	0.49	0.19	16.2	7.32	72.8	0.56	2.04	0.37	0.51
A	10	6.45	5.21	0.02	55	0.81	1.03	0.14	13.3	7.02	91.9	0.43	1.69	0.34	0.65
I	10	1.67	1.04	0.40	168	0.33	0.01	0.63	73.6	4.43	37.7	1.96	8.47	0.31	0.23
(e) S_{gal} = 0.02 yr^{-1}, c = 5.0, m = 2.0, GB density law.															
A	5	7.01	44.38	0.02	36	3.65	6.90	0.05	18.9	1.69	414.9	0.14	2.24	0.09	2.64
I	5	2.25	3.18	0.20	110	0.68	0.06	0.28	55.9	2.90	77.7	0.88	6.49	0.18	0.48
A	10	3.04	8.34	0.11	77	1.29	0.28	0.15	44.4	2.08	146.3	0.47	5.15	0.12	0.90
I	10	0.79	0.33	6.90	421	0.18	1.7×10^{-4}	2.56	256.9	3.76	20.9	8.31	27.49	0.40	0.11
(f) S_{gal} = 0.02 yr^{-1}, c = 5.0, m = 1.0, GB density law.				y_{max}	R_{max}										
A	5	7.01	10.69	0.16	79	0.68	1.35	0.14	10.4	9.03	77.7	0.53	0.82	0.86	0.54
I	5	2.25	2.86	0.17	106	0.29	0.02	0.55	44.0	6.82	33.0	2.16	3.21	0.90	0.20
A	10	3.04	4.64	0.18	97	0.42	0.10	0.32	29.5	6.34	48.0	1.22	2.21	0.74	0.31
I	10	0.79	0.93	0.20	140	0.15	1.5×10^{-4}	2.80	234.1	4.54	17.3	11.47	15.67	0.98	0.09

A = Arm condition. I = Interarm condition.

small clouds balanced by condensation on to large clouds. Alternatively, matter may be returned to the cool dense phase by the operation of a galactic fountain.

Column 3 of Table 1(a) contains the volume average contribution to the ISM pressure provided by SNR, $\langle P \rangle$. The properties of the cloud population discussed below are calculated assuming the clouds to be in equilibrium at this pressure rather than at the 'typical' pressure employed by MO (see Appendix). Notice that the behaviour of this parameter is qualitatively similar to that of $\bar{n}$ in this case. In column 4 we give, $N_{\rm cl}$, the number of cloud centres per cubic parsec. This quantity is largest in the arm region of the molecular ring and falls dramatically with decreasing $\bar{n}$ and S. Notice that the distribution of $N_{\rm cl}$ in both R and θ is significantly more sharply peaked than for instance that of $\bar{n}$. Again this behaviour is a consequence of MOs evaporative feedback mechanism.

Columns 5–8 give respectively the mean radius, $\langle a_{\rm c} \rangle$, filling factor $f_{\rm c}$, intercloud distance, $\lambda_{\rm c}$ and density, $n_{\rm c}$, of the cold cores. Notice that the forms of the variation of $\langle a_{\rm c} \rangle$, and $\lambda_{\rm c}$ are very similar, while the assumption of pressure equilibrium dictates that $n_{\rm c}$ is simply proportional to $\langle P \rangle$. Columns 9–12 give the corresponding quantities; $\langle a_{\rm w} \rangle$, $f_{\rm w}$, $\lambda_{\rm w}$ and $n_{\rm w}$ for the warm envelopes. The ratios $\langle a_{\rm c} \rangle / \langle a_{\rm w} \rangle$, $\lambda_{\rm c}/\lambda_{\rm w}$, $f_{\rm c}/f_{\rm w}$ and $n_{\rm c}/n_{\rm w}$ are entirely fixed by the assumption of pressure equilibrium and MOs modelling of the cloud properties, specifically

$$\langle a_{\rm c} \rangle / \langle a_{\rm w} \rangle = 0.38(1 + x_{\rm w})^{-1/3}, \quad \lambda_{\rm c}/\lambda_{\rm w} = 6.25(1 + x_{\rm w})^{2/3},$$

$$n_{\rm w}/n_{\rm c} = 100(1 + x_{\rm w}), \quad f_{\rm w}/f_{\rm c} = 0.02(1 + x_{\rm w})^{-1} \ln(M_u/M_l),$$

where $x_{\rm w}$ is the fractional ionization in the warm envelope and M_u and M_l are the masses of the largest and smallest clouds permitted by the model. Since $x_{\rm w}$ varies only slightly the first three ratios are almost constant, having in the present case the values ~ 0.3, ~ 8.3 and ~ 0.06 respectively. Conversely M_u/M_l is quite sensitive to the local values of $\epsilon_{\rm uv}$ and $\langle P \rangle$ and hence $f_{\rm w}/f_{\rm c}$ varies rather more.

We may now consider the effect on our results of changing the constants in the galactic model from those used to derive Table 1(a). We have performed calculations employing each of the three density distributions discussed in Section 2 for $m = 0.5$, 1 and 2 and with c and $S_{\rm gal}$ having various values in the range 1–10 and 0.01–0.1 $\rm yr^{-1}$, respectively. These show that the gross *qualitative* features exhibited in Figs 1 and 2 are preserved, even though the absolute values of the derived parameters and the magnitudes of their variation with R and θ may change considerably. One important exception to this, however, is that the qualitative behaviour of the filling factors differs significantly according to the value of m. We will return to this point shortly.

3.1 EFFECT OF CHANGING THE SHOCK COMPRESSION RATIO c, AND THE SUPERNOVA RATE $S_{\rm gal}$

Understandably, the larger the value of c, the larger the contrast between the arm and interarm values of the derived parameters. This tendency is particularly pronounced in the cases of $N_{\rm cl}$ and $Q_{\rm c}$ and is least prominent for the filling factors. Rather more surprising is the relative insensitivity to the value of c of conditions in the arm region of the molecular ring. These trends are illustrated by the data in Table 1(b) which were calculated for $c = 10.0$, the other constants being the same as for Table 1(a). The arm–interarm contrasts in $N_{\rm cl}$ and $Q_{\rm c}$ shown in Table 1(b) are respectively 4–15 times and 3–8 times those shown in Table 1(a). Conversely the values of $f_{\rm w}$ and $f_{\rm c}$ at any point are virtually the same in both tables. In the arm region at $R = 5$ kpc the data in these two tables are the same to within ~ 30 per cent and excluding $N_{\rm cl}$ and $Q_{\rm c}$ brings this difference down to $\lesssim 10$ per cent. For

any individual parameter the difference between the values in the two tables is substantially greater at any other point.

Throughout the volume of parameter space examined here if the value of S_{gal} is *increased*, $\langle P \rangle$, N_{cl}, n_c and n_w all increase at any point, while R_c, the cloud radii and cloud filling factors decrease. To a fairly good approximation, the ratio of the values of any of these quantities obtained from models differing only in the value of S_{gal}, is both independent of position and of the values of the other constants. The values of Q_c, λ_c and λ_w are almost independent of S_{gal}. At points for which the local value of S exceeds $1 \times 10^{-13}\,pc^{-3}\,yr^{-1}$ the value of Q_c increases slightly with S_{gal} while below this the converse is true. Similarly the intercloud distances are slowly increasing functions of S_{gal} wherever $S \gtrsim 0.5 \times 10^{-13}\,pc^{-3}\,yr^{-1}$ and are slowly decreasing functions elsewhere. As a specific example Table 1(c) gives the values obtained for the same values for the constants as in Table 1(a) except that $S_{gal} = 0.05\,yr^{-1}$. Dividing the data in Table 1(c) by those in Table 1(a) we obtain the ratios ~ 2 for N_{cl}, ~ 0.65 for $\langle a_c \rangle$ and $\langle a_w \rangle$, ~ 0.53 for f_c and f_w, ~ 1.85 for $\langle P \rangle$, n_c and n_w and ~ 0.85 for R_c. These ratios are broadly typical of those obtained for any other pair of models having $S_{gal} = 0.02\,yr^{-1}$ and $S_{gal} = 0.05\,yr^{-1}$. In this case the corresponding ratio of the values of Q_c is ~ 1.3 while those of the intercloud distances are both ~ 1.2.

3.2 EFFECT OF CHANGING THE MOLECULAR HYDROGEN DENSITIES

Essentially the values of $\bar{n}(H_2)$ obtained by SSS only differ from those of GB in being everywhere ~ 2.5 times larger. Thus since in either case $\bar{n}(H^0) \ll \bar{n}(H_2)$, adopting the SSS values rather than those of GB affects the normalization constant, S_0, of equation (4), in essentially the same way as *decreasing* S_{gal}. The consequences of this for the values calculated at any position is of course complicated by the accompanying increase in the local value of $\bar{n}$. We find that the overall effect of this change is to *increase* the number of clouds, their densities and their filling factors while decreasing Q_c, R_c, the cloud radii and the intercloud distances. Again the ratio of the values obtained from two models differing only in the choice of the density distribution, are approximately independent of the values of the other constants and only change slowly with position. The data in Table 1(d) were obtained with the same constants as for Table 1(a) but using the SSS density distribution. The ratios here are ~ 0.1 for Q_c, ~ 10 for N_{cl}, ~ 0.6 for R_c, $\langle a_c \rangle$ and $\langle a_w \rangle$, ~ 0.34 for λ_w and λ_c and ~ 1.1 for n_c and n_w.

The density values derived by GLV are ~ 0.64 and ~ 0.2 times those of GB at 10 and 5 kpc respectively. The consequence of these considerably lower densities is that the cloud evaporation rate per unit volume is commonly insufficient for cooling to occur before the SNRs either overlap or burst out of the disc even in the arm region.

3.3 EFFECT OF CHANGING THE STAR FORMATION EXPONENT, m

For the case $m = 2$ our results are in general qualitatively similar to those for $m = 1$ except that the magnitudes of the arm–interarm and radial contrasts in the various quantities are substantially greater. However, the behaviour of the filling factors is exceptional in this respect. We find quite generally that if $m \simeq 1.46$ the value of f_w is independent of R and θ, while for f_c the same is true at $m \simeq 1.73$. As we have seen for the case $m = 1$, below these values the filling factors are higher in the arm region than in the interarm region and have a maximum value in the molecular ring. The converse is true for the case $m = 2$. Table 1(e) which gives the results for the same model as Table 1(a) but with $m = 2$ illustrates these trends and broadly typifies the results obtained for other cases. For $m = {}^1/_2$ the basic trends are consistent with those discussed above. However, generally in this case we find that the

value of f_w predicted approaches or exceeds unity everywhere. When f_w is comparable to unity the SNR evolution will be primarily controlled by the cloud component rather than the HIM and the MO scheme is inappropriate. The situation then becomes more reminiscent of the original Cox & Smith (1974) model.

4 A non-evaporative model

Thus far we have assumed that the evaporation of embedded clouds plays a dominant role in the early stages of SNR evolution. However, thermal evaporation can be quenched if the magnetic field in the cloud–HIM interface possesses a suitable geometry. Cox (1979) has examined the evolution of SNRs in a hot, tenuous intercloud matrix in the absence of evaporation. In particular his calculations show that for this case both the external pressure and interaction with the halo will become important prior to the onset of radiative cooling. He suggested that cloud compression might then be the dominant sink for SNR energy and demonstrated that this mechanism could represent an important heat source for the clouds. To assess the consequences of such a situation for our results we have constructed a model in which Cox's scheme for SNR evolution is incorporated into the representation of the cloud population employed in the MO model.

Cox characterized SNR evolution in terms of a normalized volume y defined so that the ambient pressure becomes important when $y = 1$. He then supposed that the remnant evolved until at $y = y_{max}$, all the explosion energy had been dissipated, and calculated the volume average pressure contributed by the ensemble of SNRs as a function of y_{max}. He also obtained an expression for the rate of nett compressional work done by an SNR on the clouds contained within it. Adopting the cloud geometry and size spectrum used in the MO model (*cf.* Appendix) the integral of this latter quantity from $y = 0.1$ to y' can be approximately represented by

$$w_{net}(y') \simeq \frac{0.955\,E_0(E_0/P_h)^{1/3}n_h^{1/2}}{n_w^{1/2}\langle a_w\rangle + 0.228\,(E_0/P_h)^{1/3}n_h^{1/2}}\; f_w(1-(10y')^{-1/4}), \tag{6}$$

where n_h is the density in the undisturbed matrix which we subsequently take to be $5\times 10^{-3}\,cm^{-3}$. E_0 is the supernova explosion energy in erg and P_h is the pressure in the hot matrix. Equation (6) has been corrected to allow for the complete compression of the smallest clouds. For simplicity we assume that the entire explosion energy is consumed by cloud compression and obtain y_{max} by setting $w_{net} = E_0$ in equation (6). This, together with Cox's expression for the pressure and the usual equations relating the cloud properties to $\bar{n}$ and ϵ_{uv} leads to a system of equations analogous to those obtained in the MO case.

Acceptable solutions to these equations can only be obtained over rather more limited ranges of values for the input parameters than in the case of the MO model. The asymptotic behaviour of the explicit y dependence in equation (6) together with the rapid decrease of f_w with increasing y, consequent upon the increasing pressure, results in $w_{net}(y)/E_0$ possessing a maximum near $y \sim 0.3$. Hence, our requirement that $w_{net}(y_{max})/E_0 = 1$ imposes as an absolute upper limit $S^{2/3}/\bar{n} \leqslant 3.6\times 10^{-9}$ at any point. Conversely if the ambient pressure is to be sufficient to confine the warm envelopes so that $f_w \leqslant 1$ then $S^{2/3}/\bar{n} \geqslant 1.4\times 10^{-9}$. The consequences of any failure to meet these conditions will be similar to those discussed for the MO model in the last section.

Wherever these conditions are satisfied for a particular galactic model we find, however, that the behaviour of the derived parameters with R and θ is qualitatively similar to those obtained in the MO model. Quantitatively the 'Cox' calculations lead to pressures typically $\sim \frac{1}{2}$ those obtained from the MO model for the corresponding case. Consequently, N_{cl},

the cloud densities and intercloud distances are smaller while the cloud radii and filling factors are rather larger in the Cox case. Table 1(f) lists the data for the GB density distribution with $c = 5.0, m = 1.0$ and $S_{gal} = 0.02\,yr^{-1}$ which can be compared directly with the data of Table 1(a). R_{max} is the SNR radius when $y = y_{max}$. The behaviour of f_w is noteworthy in that it is largest where $\bar{n}$ and S are smallest which differs from the behaviour of f_c. The trends exhibited in this case are typical of those for other models using the GB or SSS density laws and with $m = 1$ (bearing in mind the limits discussed above). For models having $m = 2$ one or both of the above limits are invariably violated at some point in the disc while as in the MO case models having $m = {}^1/_2$ generally predict $f_w > 1$ throughout the disc. With the GLV density law the density is too small to permit the complete dissipation of the explosion energy by cloud crushing for any reasonable value of S_{gal}.

5 Discussion

The essential trend exhibited in our results is that in regions where $\bar{n}$ and S are large, the energy deposited in an SNR is dissipated at a much earlier evolutionary epoch than in regions of low $\bar{n}$ and S, irrespective of whether this dissipation occurs predominantly via evaporation-enhanced radiation or via cloud crushing. Consequently, the ensemble of SNRs contribute to a larger volume average pressure in such regions. The natural concomitant of this (given the assumed cloud properties) is that the clouds are smaller, more dense and considerably more numerous under such conditions.

The dramatic decrease in the predicted number of clouds in passing from 'arm' to 'interarm' conditions, raises the important theoretical question of whether the cloud spectrum is capable of adjusting sufficiently rapidly to local conditions, as is assumed implicitly in our model. The bulk of this variation results from the destruction of the smallest and most numerous clouds. In the MO model the putative mechanisms for this are shock sweeping by SNRs in the 'snow-plough' phase and complete evaporation. The time-scale for the former process will be clearly comparable with that for SNR evolution, while we find that the time-scale for the latter process is typically of the same order of magnitude. Likewise the formation of the smallest clouds is supposed to occur by means of the break up of the dense shells of old SNRs and hence will take place on a similar time-scale. Conversely, it is very unlikely that larger clouds can be created and destroyed so rapidly. However, trial calculations employing modified cloud spectra and differing methods for fixing the upper limit to the cloud mass suggest that this will not unduly affect our results. For the Cox model the mechanism of cloud creation and destruction is unknown so we can make no such assurance.

5.1 MACH NUMBER OF THE SPIRAL SHOCK

Clearly the central question and that which initially motivated this study, concerns whether or not the ISM possesses properties consistent with the existence of a density wave shock. The condition for the existence of a strong shock in the gas is that the streaming velocity normal to the spiral pattern, $W_\perp \sim 10$–$20\,km\,s^{-1}$ should exceed the effective acoustic speed in the gas $\bar{c}$. Schmidt-Kaler & Weigandt (1980) have discussed the value of $\bar{c}$ appropriate to a multicomponent medium, in the limit of long wavelengths. Writing their result in a more convenient form we obtain,

$$\bar{c} = \frac{c_w c_h}{[(1-f_w)^2 c_w^2 + f_w^2 c_h^2 + f_w(1-f_w)(c_w^2 + c_h^2)]^{1/2}}, \tag{7}$$

where c_h is the sound speed in the HIM $\sim 120\,km\,s^{-1}$ and c_w is that in the WIM which lies

between 9.2 and 13.2 km s^{-1} depending on the fractional ionization. Even for quite small values of f_w the warm envelopes are remarkably effective in retarding shocks propagating through the medium.

In Fig. 3 we show the value of $\bar{c}$ calculated from equation (7) for the 'interarm' region as a function of R in a number of representative cases. This can be compared with the expected run of $w_\perp$ for the galaxy (Burton 1977; following Roberts, Roberts & Shu 1975). For the MO models if $m = 1$ or 2 only those employing the SSS density distribution satisfy the criterion $W_\perp > \bar{c}$ at any point and then only if $S_{gal} \lesssim 0.02\,yr^{-1}$. For $m = 1$ this only occurs in a rather limited region near $R = 5$ kpc; however, for $m = 2$ the form of the variation of f_w with $\bar{n}$ and S permits the condition to be satisfied over a much larger range in R for a given value of S_{gal}. Similarly, for the Cox models the requirement that all the explosion energy be dissipated, necessitates a sufficiently large value of f_w for the existence of a shock, but can only be satisfied when S_{gal} is small and $\bar{n}$ large. As already noted the models with $m = {}^1/_2$ predict large values of f_w everywhere and consequently will generally satisfy this criterion.

The point is often made (e.g. Scott, Jensen & Roberts 1977; Reinhardt & Schmidt-Kaler 1979; Blitz & Shu 1980) that supernovae of type I which yield a smaller ($\sim {}^1/_2$, Chevalier 1977) energy per explosion than supernovae of type II, probably predominate in the interarm region. To examine the consequences of this we have performed calculations using a reduced explosion energy for all the supernovae in the interarm region. We find that as would be

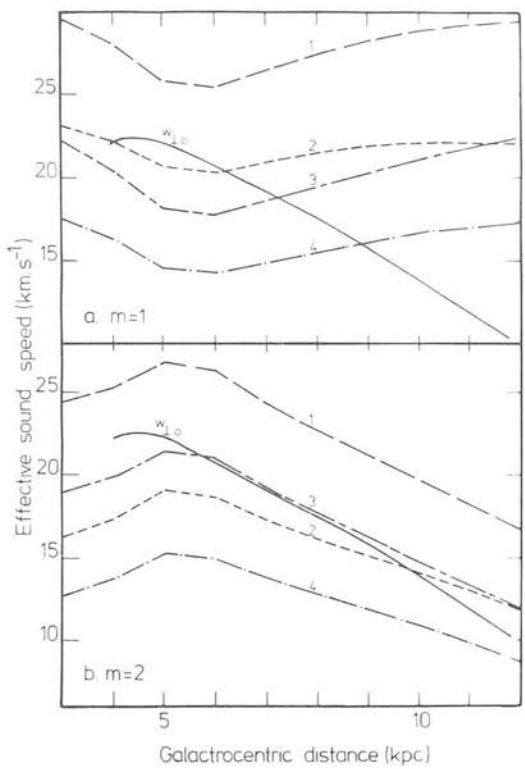


Figure 3. The effective sound speed $\bar{c}$ (km s^{-1}) of the interarm ISM as a function of galactocentric distance. For both panels a compression factor, c of 5.0 and galactic supernova rate, S_{gal} of 0.02 yr^{-1} was employed. For (a) the exponent m of star formation was 1, while for (b) it was 2. The numbered curves were calculated for (1) GB density law $E_0 = 5 \times 10^{50}$ erg, (2) SSS density law $E_0 = 5 \times 10^{50}$ erg, (3) GB density law $E_0 = 2.5 \times 10^{50}$ erg, and (4) SSS density law $E_0 = 2.5 \times 10^{50}$ erg. The solid line shows the streaming velocity perpendicular to the spiral arm $W_{\perp_0}$ (km s^{-1}) after Burton (1976).

expected this reduces $\langle P \rangle$ and hence leads to a modest increase in f_w (typically $\lesssim 70$ per cent if E_0 is halved to 2.5×10^{50} erg). Nevertheless, contrary to the suggestion of these authors the hot phase continues to be an important constituent of the interarm ISM. The increase in f_w does, however, lead to a decrease in $\bar{c}$ as indicated in Fig. 3 and a rather less restricted range of models are then able to satisfy the constraint on $W_\perp/\bar{c}$. In particular it becomes possible for models employing the GB density distribution with $m = 1$ and $S_{gal} \lesssim 0.02$ to satisfy this criterion in the molecular ring.

Even if the intercloud gas is incapable of supporting density wave induced shocks it remains possible that the 'cloud fluid' may be able to do so as has been proposed by Cowie (1980). The details of such a mechanism, however, remain to be worked out. In any case it should be borne in mind that observation both of our own and external galaxies do indicate the existence of a density enhancement associated with spiral arms (e.g. Visser 1980; Cohen *et al.* 1980) so our galactic model may not differ too much from the truth whatever mechanism is ultimately found to be responsible.

5.2 MOLECULAR CLOUDS IN THE MODEL

Another interesting question relates to the status of molecular clouds within our model. In modelling the cloud properties we have not included the presence of molecular cores although in calculating $\bar{n}$ we have included the molecular component. As a rough empirical criterion a cloud possesses sufficient self shielding to contain a significant H_2 content when its column density exceeds $\sim 6 \times 10^{20}\,\mathrm{cm}^{-2}$ (Savage *et al.* 1977), i.e. when its core radius exceeds $\sim 194/n_c$ pc. Consequently the numerous small clouds which primarily control our model never possess molecular cores, so neglecting their presence should not greatly affect our results. Overall, given the cloud spectrum employed ~ 40 per cent of the ISM mass is contained in clouds with sufficient self shielding to possess molecular cores. That this falls short of the molecular fraction actually observed and tacitly assumed is not, however, surprising.

Our cloud spectrum is assumed to terminate at the limiting mass given by Mouschovias & Spitzer (1976), corresponding to a radius typically ~ 10 pc in the spiral arms which as has often been noted falls considerably short of the observed scale length of Giant Molecular Cloud complexes (SSS; Stark & Blitz 1978). On the other hand, the very numerous small clouds found in the arms lead to typical cloud separations very much smaller than the average separation of the early type stars ~ 50 pc responsible for much of the ionization (deduced from Torres-Peimbert, Lazcano-Aruajo & Peimbert 1974). Thus mutual cloud–cloud shielding might become an important effect. Tentatively then we might envisage molecular cloud complexes as being agglomerates of many mutually shielding clouds spaced more closely than the average. The acceleration due to SNRs and stellar winds coupled with the gravitational attraction of the few largest members might provide a natural mechanism for the initial formation of such clumps. Certainly the observational evidence that molecular clouds are inhomogeneous (Silk 1980) provides some encouragement for such a point of view.

5.3 OBSERVATIONAL CONSTRAINTS

It is clearly of considerable interest to examine the observable consequences of our model and to consider what if any constraints can then be imposed upon it. MO compared many predictions of their model with the observed properties of the ISM in the solar neighbourhood, and found reasonable agreement in most respects bearing in mind the highly approximate nature of the model. Here we will concentrate on predictions relating to the larger scale

properties of spiral galaxies. In particular the question of whether the hot component of the ISM is truly ubiquitous or merely a local anomaly can clearly be most easily resolved by studying nearby galaxies.

Ultimately the most unequivocal evidence for a pervasive hot phase in any external galaxy would come from observations of its soft X-ray emission. Following MO the effective soft X-ray opacity of a multiphase medium is approximately given by,

$$k_x \sim (\lambda_c + 1/n_c\sigma)^{-1} + (\lambda_w + 1/n_w\sigma)^{-1},$$

where σ is the effective cross-section per hydrogen atom (Brown & Gould 1970; the more detailed analysis of Ride & Walker 1977, requires a different cross-section for the CNM and WIM: however, we will neglect this in view of the approximate nature of the calculation). For the MO model we find that the volume average emissivity of the HIM is proportional to S and almost independent of Q_c. It is then straightforward to calculate the relative intensity emitted at a particular energy at each point in our models. In general the S dependence of the emissivity dominates over the increased opacity resulting from larger $\bar{n}$ so that regions of high $\bar{n}$ and S would be expected to be 'bright' at soft X-ray energies. For instance in the Wisconsin groups B band (0.17 keV, Burstein *et al.* 1977) we find from the results of Table 1 that the ratios of the arm to interarm intensities are ~ 1.25 and 1.42 at 5 and 10 kpc respectively. At the higher energy of the M band (0.4–0.85 keV) the corresponding ratios are ~ 2.23 and 1.85. Because most of the soft X-ray flux is emitted by the younger, smaller remnants the actual distribution of flux will inevitably be very irregular. Furthermore, the interpretation of any observations will be complicated by the presence of emission from the galaxies, halo as well as 'contamination' by the emission from our own Galaxy.

Another potential tracer of the hot gas and one for which observations already exist is its non-thermal radio emission. Unfortunately the details of the emission mechanism in SNRs remain unclear and in particular depend critically on the geometry of the magnetic field which is as yet unknown (Caswell & Lerche 1979b). It is therefore not possible to make any useful predictions of its distribution on the basis of our model.

The statistical properties of resolved SNRs may represent a useful probe of the conditions in their vicinity (Higdon & Lingenfelter 1980). Note in particular that our model predicts an increase in the typical remnant size scale with decreasing $\bar{n}$ and S which shows up in our model as an increase in remnant size with increasing galactic radius. Such a trend has already been noted for our galaxy both for radio remnants (Jones 1975) and for the giant H I shells reported by Heiles (1979), (Bruhweiler *et al.* 1980). However, other interpretations of this trend are possible and selection effects are undoubtedly important.

It might seem that observations of the properties of the CNM and WIM in external galaxies might provide a test of our model and hence by inference provide information on the hot phase. Photographic surface photometry can provide information on the distribution of dust and hence gas in nearby galaxies with a linear resolution of tens of parsecs while 21-cm aperture synthesis measurements yield comparable data for H I alone at rather lower resolution. Unfortunately such observations ultimately can only provide a measurement of the local value of $\bar{n}$ at any point and hence while certainly useful for testing our galactic model they provide only limited information on the properties of the cloud population. In principle the statistical fluctuations in the observed column density provide a measure of N_{cl}; however, because of the large volume elements involved we do not expect the effect to be measurable. A statistical study of individual large (resolved) clouds (Elmegreen 1980) can help to elucidate some features of the model but the insensitivity of our results to the details of the cloud spectrum probably render such data of limited value in testing our model. Similar arguments can also be made with regard to Hα observations of the warm envelopes,

in addition to which the emission measure predicted from the WIM is small compared to that from a single normal H II region rendering interpretation difficult.

6 Conclusions

Our initial assumptions were: (1) that SNRs interact with the inhomogenous ISM in the manner proposed by MO, (2) that the pattern of mean total density $\bar{n}$ is similar to that predicted by TASS models; (3) that the local star formation rate and supernova rate are proportional to each other and to some power m, of $\bar{n}$. We have also examined the consequences of relaxing the first of these assumptions to the extent of replacing evaporation enhanced radiation by cloud crushing as the dominant sink of SNR energy as proposed by Cox (1979).

We conclude:

(1) Our results are not strongly dependent on whether we employ evaporation enhanced radiation or cloud crushing, nor are they grossly affected by changes of detail in our model. Consequently we expect similar results to follow from any three phase model.

(2) The theoretical case for a ubiquitous hot phase remains strong. Contrary to previous expectations if the hot phase is important in spiral arms it remains so in the interarm region. This continues to be the case even when due allowance is made for the lower energy of the type I supernovae which predominate in the interarm region.

(3) Our model imposes strong constraints on the condition which must prevail in spiral galaxies if a density wave shock is to be self-sustaining. For our own Galaxy if $m \geq 1$ then we require that both the lower estimates of the galactic supernova rate (i.e. $S_{\text{gal}} < 0.02$) and the higher estimates of $\bar{n}$ (i.e. that of SSS) are both correct. If either requirement is not met then the density wave shock can only be supported if some other mechanism such as that suggested by Cowie (1980) acts to produce a strong density enhancement in the arms. For $m \leq {}^1/_2$ our model fails because the warm component becomes predominant everywhere. Such a situation would favour density waves. However, in this instance the star formation rate has such a low arm/interarm contrast that blue spiral arms would be imperceptible, and type II supernovae would occur frequently between the arms. We believe that this provides an observational argument implying that $m > {}^1/_2$. Since observations suggest the existence of a spiral density wave in the galaxy, we must conclude that our requirements are in fact met.

(4) Observationally, it seems unlikely that studies of external galaxies can supply information on the vital question of the importance of the hot phase in the near future.

(5) The high pressure we predict within spiral arms must make the support of giant molecular clouds against self-gravity even more difficult to understand. Conversely, we expect a great many small clouds in the spiral arms. Consequently we tentatively propose a clumpy cloud model similar to that of Norman & Silk (1980). However, instead of T Tauri winds we propose that hot phase material separates, pressurizes and possibly drives the motions of the densely packed small clouds. This will be the subject of a future paper.

(6) The mechanisms we have described in this paper are perhaps subject to feedback, in the sense that for instance, supernova overproduction will drive much of the material into the hot phase and eventually reduce the star formation rate. The fact that our results (a) describe the state of our own Galaxy (and possibly others) fairly well and (b) indicate a narrow range of viable conditions in which the model sustains itself, may be taken as tentative evidence of such feedback. We are currently investigating this possibility.

Theoretical development is urgently required in the dynamics of multicomponent media

both in the large scale (effective sound speed, shocks etc.) and the small scale (cloud collision and collapse etc.)

Acknowledgment

SRH acknowledges the support of an SRC studentship.

References

Arnaud, M. & Rothenflug, R., 1980. *Astr. Astrophys.*, **87**, 196.
Blitz, L. & Shu, F. H., 1980. *Astrophys. J.*, **238**, 148.
Brown, R. H. & Gould, R. J., 1970. *Phys. Rev. D*, **1**, 2252.
Bruhweiler, F. C., Gull, T. R., Kafatos, M. & Sofia, S., 1980. *Astrophys. J.*, **238**, L27.
Burstein, P., Borken, R. J., Kraushaar, W. L. & Sanders, W. T., 1977. *Astrophys. J.*, **213**, 405.
Burton, W. B., 1976. *A. Rev. Astr. Astrophys.*, **14**, 275.
Caswell, J. L. & Lerche, I., 1979a. *Mon. Not. R. astr. Soc.*, **187**, 201.
Caswell, J. L. & Lerche, I., 1979b. *Proc. astr. Soc. Austr.*, **3**, 343.
Chevalier, R. A., 1977. *A. Rev. Astr. Astrophys.*, **15**, 175.
Chevalier, R. A. & Oegerle, W. R., 1979. *Astrophys. J.*, **227**, 398.
Clark, D. H. & Stephenson, F. R., 1977. *Mon. Not. R. astr. Soc.*, **179**, 87P.
Cohen, R. S. & Thaddeus, P., 1977. *Astrophys. J.*, **217**, L155.
Cohen, R. S., Cong, H., Dame, T. M. & Thaddeus, P., 1980. *Astrophys. J.*, **239**, L53.
Cowie, L. L., 1980. *Astrophys. J.*, **236**, 868.
Cowie, L. L. & McKee, C. F., 1977. *Astrophys. J.*, **211**, 135.
Cowie, L. L. & Songaila, A., 1977. *Nature*, **266**, 501.
Cox, D. P., 1979. *Astrophys. J.*, **234**, 863.
Cox, D. P. & Smith, B. W., 1974. *Astrophys. J.*, **189**, L105.
Dwek, E. & Scalo, J. M., 1979. *Astrophys. J.*, **233**, L81.
Elmegreen, B. G., 1976. *Astrophys. J.*, **205**, 405.
Elmegreen, D. M., 1980. *Astrophys. J. Suppl.*, **43**, 37.
Gordon, M. A. & Burton, W. B., 1976. *Astrophys. J.*, **208**, 346 (GB).
Guibert, J., Lequeux, J. & Viallefond, 1978. *Astr. Astrophys.*, **68**, 1 (GLV).
Heiles, C., 1979. *Astrophys. J.*, **229**, 533.
Heiles, C., 1980. *Astrophys. J.*, **235**, 833.
Higdon, J. C. & Lingenfelter, R. E., 1980. *Astrophys. J.*, **239**, 867.
Hobbs, L. M., 1974. *Astrophys. J.*, **191**, 395.
Jones, E. M., 1975. *Astrophys. J.*, **201**, 377.
Lozinskaya, T. A., 1980. *Astr. Astrophys.*, **84**, 26.
Mathewson, D. S., van der Kruit, P. C. & Brouw, W. N., 1972. *Astr. Astrophys.*, **17**, 468.
McKee, C. F. & Cowie, L. L., 1975. *Astrophys. J.*, **195**, 715.
McKee, C. F. & Cowie, L. L., 1977. *Astrophys. J.*, **215**, 213.
McKee, C. F. & Ostriker, J. P., 1977. *Astrophys. J.*, **218**, 148 (MO).
Miller, G. E. & Scalo, J. M., 1979. *Astrophys. J. Suppl.*, **41**, 513.
Mouschovias, T. Ch. & Spitzer, L., 1976. *Astrophys. J.*, **210**, 326.
Norman, C. & Silk, J., 1980. *Astrophys. J.*, **238**, 158.
Reinhardt, M. & Schmidt-Kaler, Th., 1979. *Astrophys. Space Sci.*, **66**, 121.
Ride, S. K. & Walker, A. B. C. Jr, 1977. *Astr. Astrophys.*, **61**, 339.
Roberts, W. W., 1969. *Astrophys. J.*, **158**, 123.
Roberts, W. W., Roberts, M. S. & Shu, F. H., 1975. *Astrophys. J.*, **196**, 381.
Salpeter, E. E., 1979. In *IAU Symp. 84, The Large Scale Characteristics of the Galaxy*, p. 245, ed. Burton, W. B., Reidel, Dordrecht.
Savage, B. D., Bohlin, R. C., Drake, J. F. & Budich, W., 1977. *Astrophys. J.*, **216**, 291.
Schmidt, M., 1959. *Astrophys. J.*, **129**, 243.
Schmidt-Kaler, Th. & Wiegandt, R., 1980. *Astr. Astrophys.*, **89**, 67.
Scott, J. S., Jensen, E. B. & Roberts, W. W., 1977. *Nature*, **265**, 123.
Seiden, P. E. & Gerola, H., 1979. *Astrophys. J.*, **233**, 56.
Sgro, A. G., 1975. *Astrophys. J.*, **197**, 621.
Shapiro, P. R. & Field, G. B., 1976. *Astrophys. J.*, **205**, 762.

Shu, F. H., Milione, V. & Roberts, W. W., 1973. *Astrophys. J.*, **183**, 819.
Silk, J., 1980. In *Star formation*, p. 133, eds Maeder, A. & Martinet, L., Saas Fee.
Smith, B. W., 1977. *Astrophys. J.*, **211**, 404.
Smith, L. F., Biermann, P. & Mezger, P. G., 1978. *Astr. Astrophys.*, **66**, 65.
Solomon, P. M., Sanders, D. B. & Scoville, N. Z., 1979. In *IAU Symp. 84, Large Scale Characteristics of the Galaxy*, p. 35, ed. Burton, W. B., Reidel, Dordrecht.
Stark, A. A. & Blitz, L., 1978. *Astrophys. J.*, **225**, L15.
Tammann, G. A., 1978. *Mem. Soc. astr. ital.*, **49**, 315.
Taylor, J. H. & Manchester, R. N., 1977. *Astrophys. J.*, **215**, 885.
Torres-Peimbert, S., Lazcano-Araujo, A. & Peimbert, M., 1974. *Astrophys. J.*, **191**, 401.
Visser, H. C. D., 1980. *Astr. Astrophys.*, **88**, 149.
Woodward, P. R., 1976. *Astrophys. J.*, **207**, 484.
Woodward, P. R., 1979. In *IAU Symp. 84, The Large Scale Characteristics of the Galaxy*, p. 159, ed. Burton, W. B., Reidel, Dordrecht.

Appendix

Following MO we introduce the evaporation parameter Σ, which measures the efficiency of evaporation and is defined by

$$\Sigma(r) = \alpha/[4\pi N_{\rm cl} < a(r_{\rm c}) > \phi_{\rm g}(l)\,\phi_{\rm m}]\ [P(r)/P(r_{\rm c})]^{-\gamma}\ . \tag{A1}$$

Unlike the MO case Σ is here a function of the SNR radius r because of the pressure dependence of the cloud radius (equation 2). In this expression $r_{\rm c}$ is the remnant radius at the cooling point and $\langle a(r_{\rm c})\rangle$ is the average over the cloud spectrum of the cloud radius evaluated at $r_{\rm c}$. α denotes the constant ratio of the blast wave velocity to the isothermal sound speed in the remnant interior. $N_{\rm cl}$ is the number of clouds per unit volume. The remaining terms are defined in equation (2).

Hence in the evaporative limit as discussed by MO we find for the dependence of the remnant age t upon r

$$t({\rm yr}) = 7.17\times10^{-2}(\eta/\alpha)\ [\eta^{-1}(1-1.625\,\eta)E_0\ \Sigma(r)]^{1/6}R^{5/3}({\rm pc}), \tag{A2}$$

where E_0(erg) is the initial explosion energy and $\eta = 6/(10+3\gamma)$. (The overall time dependence of r is given by $r \propto t^{\eta}$. For $\gamma = {}^1/_3$ as adopted in the text $\eta = 6/11$.) The determination of the cooling radius $r_{\rm c}$, the calculation of the dependence of the remnant pressure, temperature, etc. upon r and the computation of their global average values, then proceeds exactly as in MO. In particular we obtain

$$\Sigma(r_{\rm c}) = 5.37\times10^{3}\eta^{1.57}(1+3\eta)^{0.48}(1-1.625\eta)\alpha^{-0.57}\beta^{1.04}E_0^{0.04}(Q_{\rm c}/S)^{0.48} \tag{A3}$$

and

$$P(r_{\rm c})/k = 7.6\times10^{-51}(1+3\eta)^{0.69}(\eta/\alpha)^{0.58}\beta^{-0.11}E_0^{0.89}(Q_{\rm c}/S)^{-0.69}\ , \tag{A4}$$

where $Q_{\rm c}$ is the remnant porosity parameter, $S({\rm pc^{-3}\,yr^{-1}})$ is the volumetric supernova rate and β is the radiative enhancement factor.

Our treatment of the cloud properties follows that of MO except that we assume the clouds to be in pressure equilibrium at the average pressure $\langle P\rangle$ rather than the 'typical' pressure employed by MO, since $Q_{\rm c}$ may differ substantially from 0.5 in our models. Like MO we write $\langle P\rangle = P(r_{\rm c})(Q/Q_{\rm c})^{\chi}$, where χ is determined by power law fits to the computed volume average pressure and depends on both γ and the range of values of $Q_{\rm c}$ under consideration.

Using the mass and ionization balance relations together with the definition of the radius $a_{\rm wb}$ of the smallest cloud possessing a cold core, and of the radius a_{0l} of the smallest cloud,

(equations 38, 45, 41 and 37 of MO respectively) and equation (A1) leads directly to

$$\Sigma(r_c) = 0.912\,\phi_g(l)\,\phi_m\,\alpha^{-1}\,a_{wb}^2\,K_l(1+x_w)^{-1}\ln(a_{0u}/a_{0l})\,(\langle P\rangle/\bar{n}k\,T_w)\,[P(r_c)/\langle P\rangle]^{\gamma} \tag{A5}$$

$$(1-x_w)^2/(1+x_w) = 2.96\times10^{-2}\,\phi_g(l)\,\phi_m\,\alpha^{-1}\,K_l\ln(a_{0u}/a_{0l})\,(kT_w/\bar{n}\langle P\rangle)\,[P(r_c)/\langle P\rangle]^{-\gamma}\,\Sigma(r_c), \tag{A6}$$

$$\ln(a_{0u}/a_{0l}) = 4.9\times10^{-17}\langle P\rangle\,\bar{n}\,x_w^2/[K_l(1+x_w)\,k\,T_w^{1.8}\,\epsilon_{uv}] \tag{A7}$$

and

$$a_{wb} = [(1+x_w)(T_w/T_c)\,K_l^{-1}]^{1/3}\,a_{0l}\,. \tag{A8}$$

These correspond to equations (46)–(48) and (37) of MO respectively. Here $\bar{n}$ and ϵ_{uv} are respectively the mean gas density and mean volume emissivity of ionizing radiation, T_w and T_c are the temperatures of the warm and cold cloud components and x_w is the fractional ionization in the warm envelopes. K_l is the ratio of the smallest cloud mass to that of the largest fully ionizable cloud. We take a_{0u} the radius of the largest cloud to be that of the most massive magnetically supported cloud as given by Mouschovias & Spitzer (1976). Given values of S, E_0, ϵ_{uv} and $\bar{n}$ and adopting values for T_w, T_c, α, β, γ, $\phi_g(l)$, ϕ_m and K_l the equations (A3)–(A8) can be solved iteratively to obtain values of Q_c, a_{wb}, a_{0u}, a_{0l}, $\langle P\rangle$ and x_w. Thence, we may obtain values for the cloud filling factors and other parameters discussed in the text by means of the equations given by MO.

The corresponding problem for the non-evaporative solution is solved in an analagous manner but with equation (6) and Cox's (1979) expression for $\langle P\rangle$ replacing equations (A3) and (A4).

For convenience of reference we list below the values of the various parameters adopted from MO together with the ranges of the galactic model parameters discussed in the text.

c	Density jump across density wave shock = 5, 10
E_0	Explosion energy per supernova = 2.5×10^{50} erg, 5×10^{50} erg
K_l	Ratio of smallest cloud mass to mass of largest wholly ionizable cloud = 2
m	Exponent of star formation = $^1/_2$, 1, 2
$\bar{n}$	Total mean density of interstellar gas (cm^{-3})
S_{gal}	Galactic supernova rate = 0.02, 0.05 galaxy^{-1} yr^{-1}
T_w	Temperature of warm envelopes = 8000 K
T_c	Temperature of cold cores = 80 K
t_0	Full width at $1/e$ height of spiral shock in units of time = 3×10^7 yr
t_*	Time delay between star formation and first supernova = 1×10^7 yr
α	Blast wave velocity divided by isothermal sound speed in remnant interior = 2.5
β	Radiative cooling enhancement factor = 10
γ	Cloud compression exponent = $^1/_3$
ϵ_{uv}	Mean volume emissivity of ionizing photons ($cm^{-3}\,s^{-1}$)
ϕ_m, ϕ_g	Magnetic and geometrical efficiency constants for evaporation = $^1/_3$, 1
Ω_p	Density wave pattern speed = 13.5 km s^{-1} kpc^{-1}
$\Omega(R)$	Galactic rotation curve = $250/R$ km s^{-1} kpc^{-1}

Mon. Not. R. astr. Soc. (1983) **203**, 67–86

The state of clouds in a violent interstellar medium

S. R. Heathcote and P. W. J. L. Brand *University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ*

Received 1982 May 17; in original form 1982 January 14

Summary. A highly approximate but simple model of the interaction of a supernova blast wave with an interstellar cloud is described. This is used to explore the behaviour of a cloud when exposed to conditions prevalent in a violent interstellar medium. We find that having been shocked a cloud is rarely allowed sufficient time to return to pressure equilibrium with its surroundings before encountering a second shock. Significant departures from pressure equilibrium are therefore inevitable. The disruption of a cloud by its passage through a blast wave is quite effective and the half life of clouds cannot greatly exceed the mean interval between shocks striking a given cloud. In particular, we find that composite core-envelope clouds are not viable under typical conditions. Results for the net rates of thermal evaporation and PV work are also presented.

1 Introduction

Interaction between supernova remnants (SNR) and interstellar clouds play a central role in our present conception of the interstellar medium (ISM). In the three-phase model (hot, warm and cool material, McKee & Ostriker 1977; Cox 1981; Brand & Heathcote 1982) SNRs generate a pervasive hot but tenuous medium, whose pressure is responsible for the confinement of cooler and denser cloud components. In addition SNRs are a major source of cloud energy both in the form of bulk motions and thermal energy (cosmic-ray, UV and soft X-ray heating, 'cloud crushing'). Conversely, interstellar clouds dissipate SNR energy both in their direct role as energy sinks and by providing a source of new material for the hot phase (evaporation, mechanical ablation) thereby increasing the efficiency of direct radiative losses from the hot gas. This 'symbiotic' relationship between clouds and SNRs determines the equilibrium properties of the ISM. The nature of the mechanisms involved guarantee significant departures from thermal and pressure equilibrium at the level of individual clouds. For conditions in the solar neighbourhood the time-scales governing the essential physical processes (SNR shock repetition time, cloud dynamical response time, radiative cooling time, ionization and recombination times, evaporation time-scale etc) are all similar, and close to $\sim 10^5$–10^6 yr. The state of any given cloud will then be a complex function of its past history as well as the nature of its present surroundings.

In the present paper we attempt to follow the response of individual clouds to their changing environment. In particular we will be concerned with the determination of some sort of average cloud state. This is important both for the construction of global theoretical models and for comparison with observation. In the next section we assemble a highly simplified model for the interaction of an isolated cloud with a single SNR blast wave, generalized from the numerical example by Woodward (1976).

In Section 3 we enumerate a number of the complicating factors associated with the 'real world' of a three-phase ISM and establish the domain of validity of this simple model. In Section 4 we discuss a number of the consequences of our model for theories of the ISM while in the final section we summarize our principal results.

2 Basic cloud model

A number of authors have addressed the problem of the dynamical response of an interstellar cloud to being overrun by a shock wave (e.g. Bychcov & Pikel'ner 1975; McKee & Cowie 1975; Sgro 1975; Dyson & Gulliford 1975). The most 'realistic' calculations available to date are those of Woodward (1976, 1979, hereinafter W76 and W79) who employed a two-dimensional hydrodynamic code to follow the cloud evolution in considerable detail. These computations were carried out for essentially only one set of model parameters chosen for their relevance to the problem of star formation in massive clouds, imploded by density wave shocks. The formidable scale of the computing problem involved precluded any attempt at a parameter study. For our purposes it is necessary to extend the applicability of these calculations to encompass the large volume of parameter space appropriate to the interaction of clouds with SNR blast waves. To accomplish this we employ Woodward's results to identify the important physical processes which can be expected to occur and we then attempt to model these with simple highly approximate analytic calculations. In this way it is possible to assemble a crude model of the interaction dynamics capable of reproducing the gross features of the detailed computations which then forms the basis for a plausible extrapolation to other conditions. A few points about this procedure are worth noting at the outset.

(i) We will not hesitate to sacrifice numerical accuracy in favour of analytical simplicity. The uncertainties in the initial conditions inherent in our current understanding of the three-phase ISM are sufficient that a factor of 2 or so is unlikely to prove of vital importance.

(ii) Although throughout we will be guided by physical principles, the justification for our selection of the important processes and our approximate treatment of them rests heavily on the numerical models.

(iii) We are aware that processes not apparent in the computed models may become important in other regions of parameter space. Some of these will be discussed in Section 3, and their consequences in Section 4.

We begin the discussion of our approximate model with a brief summary of the important results of Woodward's computations.

2.1 WOODWARD'S RESULTS

In W76 the collision of a plane parallel shock of mach number $M_b \sim 2.6$ with an initially spherical cloud having a density ~ 80 times that of the pre-shock intercloud medium (ICM) was studied. The cloud was initially at rest in, and in pressure equilibrium with, the ICM.

The equation of state of the ICM was a $\gamma = 5/3$ adiabat while that of the cloud was effectively isothermal. In W79 the parameters of the cloud were unchanged, but an isothermal equation of state was used for the ICM. This had the important dynamical effect that the post-shock flow was supersonic with respect to the cloud and so a bow shock formed upstream of the cloud. The evolution was also followed to a considerably later stage in this case. The main features of the interaction are, however, similar in both studies. These are (see also Fig. 1):

(i) As the external shock (S_0) overruns the cloud its ram pressure drives a strong shock (S_1) into the front face of the cloud (Sgro 1975; Silk & Solinger 1973). Subsequently, when steady flow around the cloud is established, this shock is driven by the high pressure in the vicinity of the front stagnation point.

(ii) Having swept over the cloud, the various segments of the shock meet almost head-on at the back of the cloud and are reflected leading to a transiently high pressure there. This drives a somewhat weaker shock (S_2) into the back of the cloud. The pressure behind the cloud falls rapidly as the intersection point of the external shock segments moves downstream and the angle of incidence increases. However, the flow around the cloud conspires with the sideways flow of the reflected shocks to set up a vortex system the pressure of which, combined with the thermal pressure of the ICM, continues to drive S_2 (Fig. 1a).

(iii) The shock driven into the sides of the cloud (S_3) is quite weak and the lateral compression of the cloud is only moderate. Together with points (i) and (ii) this implies that the cloud collapse is essentially one-dimensional and for much of the time it can, in fact, be treated as approximately plane parallel.

(iv) In the subsonic case, the sides of the cloud are Kelvin–Helmholtz unstable. The instability grows rapidly and is later amplified by the Rayleigh–Taylor instability. When the flow over the cloud is strongly supersonic the Kelvin–Helmholtz instability is suppressed (Gerwin 1968).

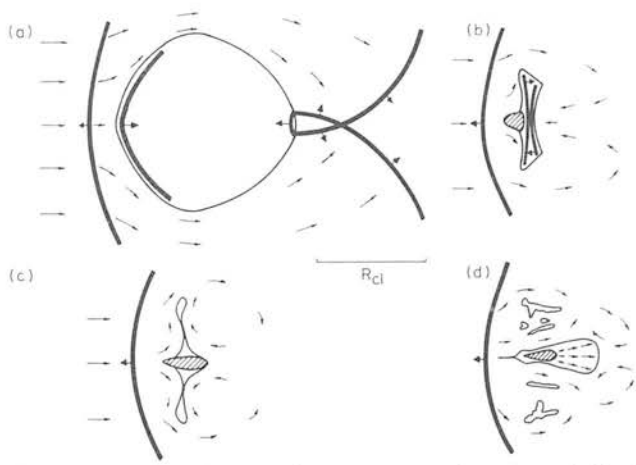


Figure 1. A schematic representation of the clouds structure at various stages during its collapse and subsequent re-expansion. Heavy lines represent shocks; broad tipped arrows attached to these indicate the downstream side of the shock. The cross hatched areas indicate the zone of enhanced density associated with the axial clump. Narrow tipped arrows depict the flow of the gas around the cloud. (a) $\bar{t} \simeq 3A/M_b$ ($t \sim 0.01$ Myr for the typical cloud of Table 1). The various segments of the external shock have just met at the back of the cloud. (b) $\bar{t} \simeq 0.6/M_b$ ($t \sim 0.5$ Myr). The shocks driven into the front and back of the cloud are about to collide. (c) $\bar{t} > 0.6/M_b$ the re-expansion of the cloud in the downstream direction has commenced. (d) $\bar{t} \sim 0.6/M_b + 0.5$ ($t \sim 1.6$ Myr) a typical stage during the re-expansion.

(v) The convergence of the shock S_1 causes it to accelerate (Zel'dovich & Raiser 1968) and leads to the front face of the cloud being Rayleigh–Taylor unstable. The resulting flow of cloud gas towards the symmetry axis leads to the formation of a dense clump of material there. In the subsonic case this collection process is limited to the zone between the Kelvin–Helmholtz tongues closest to the symmetry axis (Fig. 1b).

(vi) Eventually S_1 and S_2 collide. The shock reflected towards the back of the cloud is much stronger than that reflected towards the front. The cloud at this stage is strongly flattened so this shock reaches the back of the cloud very soon after the interaction (Fig. 1c).

(vii) At the back of the cloud a weak shock is propagated into the ICM and a strong rarefaction wave is reflected back into the cloud. This 'engulfs' most of the material near the symmetry axis including the dense clump and a fan of quite dense material expands rapidly in the downstream direction. Any expansion upstream is prevented by the pressure at the forward stagnation point.

(viii) The remaining cloud material is confined to an extremely thin sheet. This also begins to re-expand but is quite rapidly fragmented by Kelvin–Helmholtz instability before this process has progressed very far (Fig. 1d).

As noted above, the cloud equation of state employed by Woodward was essentially isothermal. Fortunately, a calculation for the same case as W76 but with an adiabatic equation of state has been performed by Thorogood (1979). The main dynamical features discussed above in (i)–(iii) and (vi) are reproduced in this calculation except that the cloud is substantially less flattened. This latter effect can be entirely understood in terms of plane parallel calculations discussed below. Unfortunately, the resolution of the computation was insufficient to show the instability in (iv) and (v) although Thorogood estimates that they should have occurred. Because of the important role played by the dense axial clump in the re-expansion phase, we cannot be confident of the details of the re-expansion phase in adiabatic clouds.

The phases outlined in (i)–(viii) above seem to be inevitable parts of a shocked cloud's evolution. Phases (iv) and (v) cause the greatest difficulty, since they represent highly non-linear processes, namely the interactions between surface instabilities and the complex flow regime outside the cloud. In the following subsections we shall approximate each of the phases either by an analytical solution, or by a fit to the data derived by reasonable extrapolation from Woodward's model.

2.2 THE COLLAPSE PHASE

To the extent that we may neglect self-gravity, and provided cooling and heating of the gas occur on time-scales either very long or very short compared to the dynamical time-scale then we may scale results obtained for a cloud of given size to yield those appropriate to a cloud of different size by simply multiplying all lengths and times by a common factor. It is thus convenient to introduce the dimensionless lengths times and velocities defined by

$$\tilde{R} = R/R_{\mathrm{cl}} \qquad \tilde{t} = C_{\mathrm{cl}}\, t/R_{\mathrm{cl}} \qquad \tilde{u} = u/C_{\mathrm{cl}}. \tag{1}$$

Here R_{cl} and C_{cl} are respectively the initial scale length and sound speed of the undisturbed cloud. The problem then reduces to one containing essentially only two parameters, namely the Mach number, M_{b} of the incident shock and the ratio A of C_{cl} to the sound speed of the undisturbed ICM, C_{ic}. We must also specify the polytropic indices describing the equation of state of the cloud and the ICM, γ_{cl} and γ_{ic} respectively. In the case of infinite cooling times we take $\gamma = 5/3$ while in the opposite case of infinitesimal cooling times we take $\gamma = 1$.

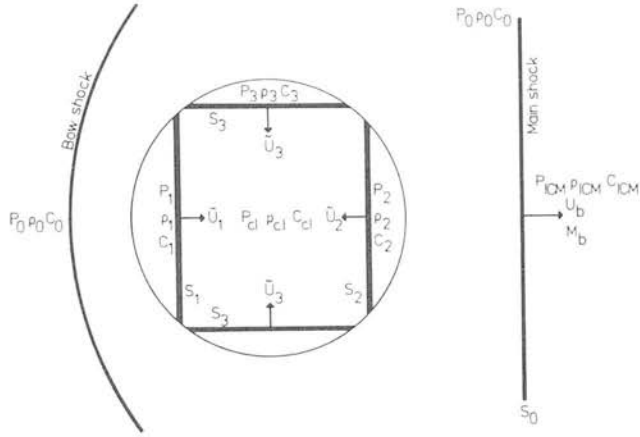


Figure 2. A schematic representation of the cloud during the collapse phase showing our notation for the several shocks and their velocities and for the pressures P, densities ρ and sound speeds C in the various zones.

Our approximation is to assume that the front cloud shock S_1 and the back cloud shock S_2 are plane parallel and are driven by the front stagnation point pressure, P_1 and the pressure at the back of the cloud, P_2 (see Fig. 2). We also ignore the influence of the lateral cloud shock, S_3 upon the motion of these shocks. In reality, although as noted above S_1 and S_2 are almost plane, their strengths decrease with distance from the symmetry axis (the distinction between S_1, S_3 and S_2 is of course purely artificial) so our results will only be approximately true on the symmetry axis. However, comparison with Woodward's model suggests that this approximation is a good one.

The calculation of P_1 is straightforward. Following the impact of the external shock, S_0, with the front of the cloud, a strong shock is transmitted into the cloud and a weak shock is reflected back into the ICM. There then follows a transient phase (Silk & Solinger 1973; Sgro 1975) during which steady flow becomes established around the cloud. The reflected shock either disappears upstream, if the flow behind S_0 is subsonic with respect to the cloud, or else forms a standing bow shock. This phase only lasts for a time comparable with the cloud crossing time for the external shock $\sim A/M_b$ and can be shown to have only minor importance for the cloud dynamics.

Subsequently, the situation is that discussed by McKee & Cowie (1975). They found that for $\gamma_{ic} = 5/3$,

$$\begin{aligned} P_1/P_0 &= (1 + M^2/3)^{5/2} && M \leqslant 1, \\ P_1/P_0 &= 1.5\, M^2\,[1 - 1/(5M^2)]^{-3/2} && M \geqslant 1. \end{aligned} \tag{2a}$$

Similarly, we find for $\gamma_{ICM} = 1$,

$$\begin{aligned} P_1/P_0 &= \exp(M^2/2) && M \leqslant 1, \\ P_1/P_0 &= M^2 \exp[1/(2M^2)] && M \geqslant 1. \end{aligned} \tag{2b}$$

Here, P_0 is the pressure behind S_0, while M is the local Mach number of the flow behind S relative to the cloud boundary. Applying the jump conditions for normal shocks to obtain the velocities of the gas behind S_0 and S_1 we find that in either case (remembering M_b is the Mach number of the initial shock and $A = C_{cl}/C_{ic}$)

$$M = (M_b^2 - 1)\ \{[\gamma_{ic} M_b^2 - {}^1\!/_2(\gamma_{ic} - 1)]\ [{}^1\!/_2(\gamma_{ic} - 1)M_b^2 + 1]\}^{1/2}\ \{1 - [A(\gamma_{ic} + 1)/(\gamma_{cl} + 1)(\tilde{u}_1^2 - 1)(M_b^2 - 1)(M_b/\tilde{u}_1)]\}\ . \tag{3}$$

Finally, since the cloud and ICM are supposed initially to be in pressure equilibrium the normal shock jump conditions imply that the velocity $\tilde{u}_1$ behind S_1 is given by

$$\tilde{u}_1^2 = \{(P_1/P_0)(\gamma_{cl}+1)/[(\gamma_{ic}+1)\gamma_{cl}]\ [\gamma_{ic}M_b^2 - {}^1/_2(\gamma_{ic}-1)]\} + {}^1/_2(\gamma_{cl}-1)/\gamma_{cl}. \quad (4)$$

Equations (2), (3) and (4) together reduce to a transcendental equation for $\tilde{u}_1$ as a function of M_b and A. McKee & Cowie (1975) obtained an approximate solution for the case $\gamma_{ic} = 5/3$ under the simplifying assumption that both S_0 and S_1 are strong. From this an expression can be derived for $\tilde{u}_1$ in terms of M_b and A which errs from the exact solution by less than 50 per cent even near $M_b \sim 1$. Similarly, for $\gamma_{ic} = 1$ we have derived an expression for $\tilde{u}_1$, which is everywhere accurate to ~ 20 per cent. When A is very small as it will be in all cases of interest here we may further simplify by taking the limit $A \to 0$ which leads to,

$$\tilde{u}_1 \sim 1.4\,[(\gamma_{cl}+1)/\gamma_{cl}]^{1/2} M_b \qquad \gamma_{ic} = {}^5/_3 \quad (5a)$$

$$\tilde{u}_1 \sim 0.71\,[(\gamma_{cl}+1)/\gamma_{cl}]^{1/2} M_b^2 \qquad \gamma_{ic} = 1\,. \quad (5b)$$

The pressure at the back of the cloud just after the collision of the external shock segments can be estimated from the standard results for the head-on collision of two plane shocks. Again this pressure is only transient and we may neglect its effects on the motion of S_2. Because of the contribution from the ram pressure of the vortex system behind the cloud an exact calculation of the value of P_2, during the subsequent evolution of the cloud would be extremely complex. However, comparison with Woodward's results suggests that an adequate approximation is $P_2 = \xi P_0$ with ξ of order unity. Thus, in the strong shock limit we may write

$$\tilde{u}_2^2 \simeq \xi(\gamma_{ic}/\gamma_{cl})\,[(1+\gamma_{cl})/(1+\gamma_{ic})]\,M_b^2\,. \quad (6)$$

In the limit of very low M we may estimate the pressure at the sides of the cloud from the solution for the steady flow of an incompressible fluid around a rigid sphere (e.g. Currie 1974). The results may then be crudely extended for finite, but still subsonic incident flow speeds by means of the Prandtl–Glauert rule (Currie 1974). For our purposes a sufficient approximation to the value of $\tilde{u}_3$ for arbitrary M, which is in agreement with both these approximate subsonic calculation and Woodward's results, is obtained from equation (6) with $\xi \sim 0.25$.

The collapse phase ends when S_1 and S_2 collide after a time

$$\tilde{t}_{col} = (2 + \tilde{u}_2 \tilde{t}_{inj})/(\tilde{u}_1 + \tilde{u}_2) \quad (7)$$

where $\tilde{t}_{inj}$ is the time of injection of S_2 into the back of the cloud which we estimate to be $\sim 3A/M_b$. Employing the strong shock and small A approximations we obtain

$$\tilde{t}_{col} \sim [\gamma_{cl}/(\gamma_{cl}+1)]^{1/2} M_b^{-1} \qquad \gamma_{ic} = {}^5/_3\,, \quad (8a)$$

$$\tilde{t}_{col} \sim 3\,[\gamma_{cl}/(\gamma_{cl}+1)]^{1/2}\,[M_b\,(M_b+1)]^{-1} \qquad \gamma_{ic} = 1\,. \quad (8b)$$

Comparison with W76 and W79 suggests that these estimates are ~ 40 per cent too small. The more exact calculation using equation (4) for $\tilde{u}_1$ gives a result ~ 11 per cent too small.

2.3 THE DEVELOPMENT OF THE DENSE AXIAL CLUMP

What fraction of the clouds mass can be accumulated into the dense clump created by the Rayleigh–Taylor initiated lateral collapse of material behind the front shock S_1? This depends on:

(i) The effective area of the front surface of the cloud from which material may be collected. If the incident flow is subsonic this is determined by the lateral distance to the Kelvin–Helmholtz tongues closest to the symmetry axis.

(ii) The rate at which the Rayleigh–Taylor instability can collect the available material.

Following W76 we obtain for the ratio of the growth time of a given Kelvin–Helmholtz mode to the dynamical time-scale of the cloud,

$$t_{\mathrm{KH}}(\tilde{K})/t_{\mathrm{dyn}} \sim (\gamma_{\mathrm{ic}}/\gamma_{\mathrm{cl}})^{1/2}\tilde{K}^{-1}(\tilde{u}_1/M_{\mathrm{slip}})(C_{\mathrm{cl}}/C_1)$$

where $\tilde{K}$ is the wave number of the mode in units of R_{cl}^{-1}, M_{slip} is the local Mach number of the flow at the cloud boundary and C_1 is the sound speed in the gas behind S_1. If we again employ the solution for incompressible flow about a rigid sphere, we find $M_{\mathrm{slip}} \sim M \sin\phi$ where ϕ is the polar angle of a given point on the cloud surface measured from the symmetry axis. Using this together with the exact solutions to equations (2)–(4) we find that the ratio of the time-scales for fixed ϕ becomes almost independent of M_{b} for $M_{\mathrm{b}} \gtrsim 2$ ($\gamma_{\mathrm{ic}} = 1$ or $^5/_3$). Thus for incident shocks stronger than this but weak enough that the post-shock flow is still subsonic with respect to the cloud, the general form of the Kelvin–Helmholtz ripples will be very similar to those found in W76. Since we will generally be interested in $M_{\mathrm{b}} \gtrsim 2$ we can therefore assume a roughly constant collecting area for the axial clump in the subsonic regime.

For $A = 0$ ('rigid' cloud) the flow behind S_0 is subsonic with respect to the cloud when $M_{\mathrm{b}} < 2.7$ for $\gamma_{\mathrm{ic}} = {}^5/_3$ and $M_{\mathrm{b}} < 1.62$ for $\gamma_{\mathrm{ic}} = 1$. However, for non-zero A this limit can be appreciably raised. Once the flow around the cloud is supersonic, as noted above, the Kelvin–Helmholtz instability is suppressed (Gerwin 1968) and effectively the whole area of the front face of the cloud is available for the collection of clump material.

As discussed in W76 the ratio of the growth rate of the Rayleigh–Taylor instability to the dynamical time-scale is only weakly dependent on M_{b} in an indirect manner through changes in the degree of flattening of the front face of the cloud. There is thus no simple way of determining the relationship between M_{b} and the efficiency with which material is collected to form the clump, and we estimate the fraction collected by inspection of Woodward's models.

At the end of the computation for the subsonic case (W76) just before the collision of S_1 and S_2 the clump contained ~ 5 per cent of the cloud's mass but was still growing. If all the available material was incorporated into the clump before the re-expansion of the cloud commenced then we estimate that the fraction of the cloud mass contained in the clump would be $F \simeq 0.1$. Careful 'planimetry' of the diagrams for the supersonic case (W79) suggests that there $F \simeq 0.5$. We will take these two values as representative of the subsonic and supersonic regimes, although possible weak dependence of F on M_{b} through the collection efficiency should be borne in mind.

2.4 THE COLLISION OF SHOCKS S_1 AND S_2

To the extent that S_1 and S_2 are very nearly parallel at the time of their collision we may employ simple shock tube theory (Bradley 1962) to calculate the effect of the collision. When $\gamma_{\mathrm{cl}} = 1$ (the isothermal case) the consequences of the collision are straightforward. The shocks S_1 and S_2 'pass through' each other, their Mach numbers being unaffected, although of course they now travel into a medium of higher density which has already passed through the other shock (the sound speed in this gas is still C_{cl}, however). The density of the twice shocked material between S_1 and S_2 becomes ~ $\tilde{u}_1^2\tilde{u}_2^2\rho_{\mathrm{cl}}$.

Conversely, when $\gamma_{cl} = {}^5/_3$ the properties of the resultant shocks can only be found from a complete analysis. Nevertheless, a procedure which is sufficient to determine the resulting density, given that S_1 and S_2 are not of drastically different strengths is to employ the result for the collision of two identical shocks of strength $(\tilde{u}_1 + \tilde{u}_2)/2$. Comparison with the exact calculation shows that this procedure gives a value for the resulting density which errs by $\lesssim 30$ per cent from the true mean density in the cases of interest.

Since as we have seen the cloud becomes considerably flattened for quite moderate shock strengths, the time between the collision of S_1 and S_2 and the arrival of the rearward moving shock at the back of the cloud is quite short compared to t_{coll}. In what follows we will therefore neglect it and take the re-expansion phase as starting immediately after the collision of S_1 and S_2.

2.5 THE RE-EXPANSION PHASE: TAIL GROWTH

Although the outflow behind the rarefaction wave is initially plane parallel, as the cloud 'tail' develops it expands laterally and the plane parallelism quickly disappears. Comparison with W79 suggests that the ensuing motion is more nearly spherically symmetric. In our approximate treatment of the tail's evolution we therefore treat it as forming a segment of a gas sphere expanding freely into vacuum.

For $\gamma_{cl} = 1$ the density profile and equation of motion are (Kahn 1972; Imshenik 1960).

$$\rho(r, t) = \rho_c(t) \exp\left[-r^2/2R(t)\right] \tag{9a}$$

$$P(r, t) = C_{cl}^2 \rho(r, t) \tag{9b}$$

$$v(r, t) = \sqrt{2}\, C_{cl} \{\ln [R(t)/R(0)]\}^{1/2} r/R(t) \tag{9c}$$

$$C_{cl}\, t/R(0) = \sqrt{2}\, R(t)/R(0) D \langle \{\ln [R(t)/R(0)]\}^{1/2} \rangle. \tag{9d}$$

Here,

$$D(x) = \exp(-x^2) \int_0^x \exp(-y^2) dy\,,$$

and the boundary conditions $R = R(0)$, $\dot{R}(t) = 0$ at $t = 0$ have been imposed. Obviously, the validity of this solution requires that the above density profile should hold at all times including the initial instant. Although this is patently not true it is to be expected that at times sufficiently great that the initial conditions have been forgotten, the evolution for the case of an arbitrary initial density distribution will not differ greatly from equation (9). Furthermore, comparison with W79 shows that the actual density profile, velocity profile and evolutionary time-scale do not differ grossly from that given by this solution even at quite early times.

For $\gamma_{cl} = {}^5/_3$ the corresponding density profile and equation of motion are (Zel'dovich & Raizer 1968)

$$\rho(r, t) = \rho_c(t) \{1 - [r/R(t)]^2\}^{3/2} \tag{10a}$$

$$P(r, t) = P_c(t) \{1 - [r/R(t)]^2\}^{5/2} \tag{10b}$$

$$v(r, t) = \sqrt{3}\, C(o) \{1 - [R(o)/R(t)]^2\}^{1/2} r/R(t) \tag{10c}$$

$$C(0) t/R(0) = \langle {}^1/_3 \{[R(t)/R(0)]^2 - 1\} \rangle^{1/2} \tag{10d}$$

here $C(0)$ is the sound speed of the tail gas at the beginning of the expansion and the same boundary conditions as above have been employed.

To complete the solution in either case we must obtain the value for the initial scale length, $R(0)$. To do this we require that the mass of the conical segment representing the tail should equal the mass of the dense clump from which we obtain

$$R(0) = R_{cl}\,\{[\rho_{cl}/\rho_c(0)]\,F\Theta(\theta)\}^{1/3} \qquad \gamma_{cl} = 1 \tag{11a}$$

$$R(0) = R_{cl}\,\{[\rho_{cl}/\rho_c(0)]\,F\Theta(\theta)\}^{1/3} \qquad \gamma_{cl} = {}^5/_3 \tag{11b}$$

where

$$\Theta(\theta) = [2 + \sin^2\theta\cos\theta + \cos^3\theta - 3\cos\theta]^{-1}$$

and θ is the semi-angle of the conical segment. From W79 we estimate $\theta \sim 20^\circ$ so that $\Theta \sim 8$.

The expansion phase ultimately ends when the tail reaches pressure equilibrium with its surroundings. We take $P_c(t)/2.7$ as the typical pressure in the tail (this is obtained at $r = 2R$ and $r \sim {}^1/_3 R$ for the isothermal and adiabatic cases respectively). We also take the pressure of the surroundings to be the pre-shock pressure P_{ic} rather than the post-shock pressure P_0. Hence, we obtain for the expansion time-scale

$$\tilde{t}_{exp} \approx 2F^{1/3}D\langle{}^1/_3 \ln\{[0.4P_c(0)/P_{ic}]\}\rangle^{1/2} \qquad \gamma_{cl} = 1 \tag{12a}$$

$$\tilde{t}_{exp} \approx [C_{cl}/C(0)]\,F^{1/3}\,\{1-1.5\,[P_{ic}/P_c(0)]^{2/5}\}^{1/2} \qquad \gamma_{cl} = {}^5/_3\,. \tag{12b}$$

Table 1. Values for the critical phases of evolution of a 'representative cloud', showing the scaling with the controlling parameters A and M_b. $R_{cl} = 2$ pc $C_{cl} = 0.8\,\mathrm{km\,s^{-1}}$, $t_{cl} = R_{cl}/C_{cl} = 2.5$ Myr, $A = C_{cl}/C_{ic} = 6.67\times10^{-3}$, $M_b = 3.5$, $\gamma_{cl} = 1$, $\gamma_{ic} = {}^5/_3$.

Phase	Epoch		Extent				Comments
			Lateral		Axial		
	t/t_{cl}	Myr	Y/R_{cl}	pc	Z/R_{cl}	pc	
Shock reaches front of cloud	0		2.0	4.0	2.0	4.0	
External shock reaches back of cloud	$3A/M_b$	0.014	2.0	4.0	2.0	4.0	Shock is slowed by cloud boundary. Fig. 1(a)
'Typical' stage during collapse		0.23		3.0		1.0	Time is $t_{col}/2$
Shocks S_1 and S_2 collide	$0.64/M_b$	0.46	0.98	2.0	$0.9M_b^{-2}$	0.15	Time is t_{col} equation (8a). Extents calculated from $\tilde{U}_1$ equation (6a). and $\tilde{U}_2$ equation (7). Fig. 1(b).
'Typical' stage during re-expansion		1.6		0.55		1.2	Time is $t_{col} + t_{exp}/2$ Fig. 1(d).
End of re-expansion pressure equilibrium	$0.72 + 0.64/M_b$	2.3	0.55	1.1	1.2	2.4	Time is $t_{col} + t_{exp}$ equation (12a). t_{exp} is nearly independent of M_b. Length is scale length of tail $R(t_{exp})$ equation (9c): Width from 20° opening angle.

2.6 SUMMARY OF CLOUD MODEL: A REPRESENTATIVE CASE

To provide a better feel for the results of this section we give values for the cloud dimensions and time-scales for a 'representative' cloud at various phases during its evolution in Table 1. The parameters chosen for this 'representative' cloud are $R_{cl} = 2$ pc, $C_{cl} = 0.8\,\mathrm{km\,s^{-1}}$, $C_{ic} = 120\,\mathrm{km\,s^{-1}}$, $M_b = 3.5$, $\gamma_{cl} = 1$ and $\gamma_{ic} = 5/3$. The term 'typical' in the table is used to describe times half way through the collapse and half way through the re-expansion.

3 The 'real world' of a three-phase ISM

So far we have considered a highly idealized model for a cloud shock wave interaction. Our model is an over-simplification in a number of respects:

(i) Our initial conditions were that the cloud was in pressure equilibrium with, and at rest in, the ICM.

(ii) Our cloud was initially spherically symmetric and structurally homogeneous.

(iii) We employed a simple polytropic equation of state for the cloud throughout its evolution.

In addition, a number of potentially important physical processes have so far been neglected. In this section we will concentrate on establishing the time-scales of these various processes to compare with the clouds evolution time-scale. The effects that the processes have on the

Table 2. Adopted values of various ISM parameters for the solar neighbourhood. After McKee & Ostriker (1977) and Brand & Heathcote (1982).

Parameter	Value
Intercloud medium	
Sound speed	$C_{ic} \simeq 95\sqrt{\gamma_{IC}}\,\mathrm{km\,s^{-1}}$
Temperature	$T_{ic} \simeq 7\times10^{5}\,\mathrm{K}$
Density	$n_{ic} \simeq 2\times10^{-3}\,\mathrm{cm^{-3}}$
'Cold clouds'	
Sound speed	$C_{cl} \simeq 0.8\sqrt{\gamma_{cl}}\,\mathrm{km\,s^{-1}}$
Temperature	$T_{cl} \simeq 10^{2}\,\mathrm{K}$
Density	$n_{cl} \simeq 30\,\mathrm{cm^{-3}}$
$A = C_{cl}/C_{ic}$	$A \simeq 8.4\times10^{-3}\sqrt{\gamma_{cl}/\gamma_{IC}}$
'Warm clouds'	
Sound speed	$C_{cl} \simeq 10\sqrt{\gamma_{cl}}\,\mathrm{km\,s^{-1}}$
Temperature	$T_{cl} \simeq 10^{4}\,\mathrm{K}$
Density	$n_{cl} \simeq 0.2\,\mathrm{cm^{-3}}$
Ionization fraction	$x_{cl} \simeq 0.5$
$A = C_{cl}/C_{ic}$	$A \simeq 0.11\sqrt{\gamma_{cl}/\gamma_{IC}}$
Supernova remnants	
Volumetric supernova rate	$S \simeq 1\times10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}}$
Energy per supernova	$E_0 \simeq 1\times10^{51}\,\mathrm{erg}$
Porosity parameter	$Q_c \simeq 0.5$
SNR cooling radius	$r_c \simeq 182\,\mathrm{pc}$
Interior density at r_c	$n_c \simeq 5\times10^{-3}\,\mathrm{cm^{-3}}$
Interior temperature	$T_h \simeq 4.4\times10^{5}(M_b/M_{bc})^2\,\mathrm{K}$
Blast wave mach number at r_c	$M_{bc} \simeq 3.5$
Shock repetition time	$t_{sn} \simeq 3\times10^{5}\,\mathrm{yr}$
Miscellaneous	
Ambient photoionizing UV flux	$J \simeq 0.2\times10^{5}\,\mathrm{cm^{-2}\,s^{-1}}$
Cloud number density	$N_{cl} \simeq 6\times10^{-3}\,\mathrm{pc^{-3}}$
Cloud velocity parameter	$\langle R_{cl}^2 V_{cl}\rangle \simeq 9.5\,\mathrm{pc^2\,km\,s^{-1}}$

cloud evolution will be discussed in Section 4. The processes considered are: multiple shocks, cloud evaporation, ionization, H_2 formation, cloud–cloud collisions and self-gravity.

As discussed above the dynamics of the cloud are specified by essentially three parameters; the cloud radius, R_{cl}; the ratio of its sound speed to that of the ICM, A, and the Mach number of the incident shock, M_b. Furthermore, when A is small, as it often is in the cases of interest, the problem ceases to be sensitive to its value. However, many of the other processes are dependent on additional parameters of the cloud and its surroundings. For definiteness of presentation we will adopt a single set of values for these various quantities which will be employed in all numerical examples and then indicate the scaling (usually simple) to other cases. The values, selected on the basis of both observation and theoretical models of the three-phase ISM (McKee & Ostriker 1977; Brand & Heathcote 1982) as being appropriate to the solar neighbourhood, are listed for ease of reference in Table 2. A general feature of any model of the ISM is the existence of two preferred equilibrium cloud temperatures viz $\lesssim 100$ K ('cold clouds') and $\sim 10^4$ K ('warm clouds'). We will treat these two cases separately.

The boundary of the region of validity in the (M_b, R_{cl}) plane appropriate for cold clouds is shown in Figs 3–5 for the following cases;

Fig. 3. Isothermal cloud, adiabatic ICM ($\gamma_{cl} = 1, \gamma_{ic} = {}^5/_3$).
Fig. 4. Isothermal cloud, isothermal ICM ($\gamma_{cl} = 1, \gamma_{ic} = 1$).
Fig. 5. Adiabatic cloud, adiabatic ICM ($\gamma_{cl} = {}^5/_3, \gamma_{ic} = {}^5/_3$).

The calculation of the position of this boundary will now be described and the generalization to other cases will be discussed. At the end of the next section the question of which γ_{cl} to use is also settled.

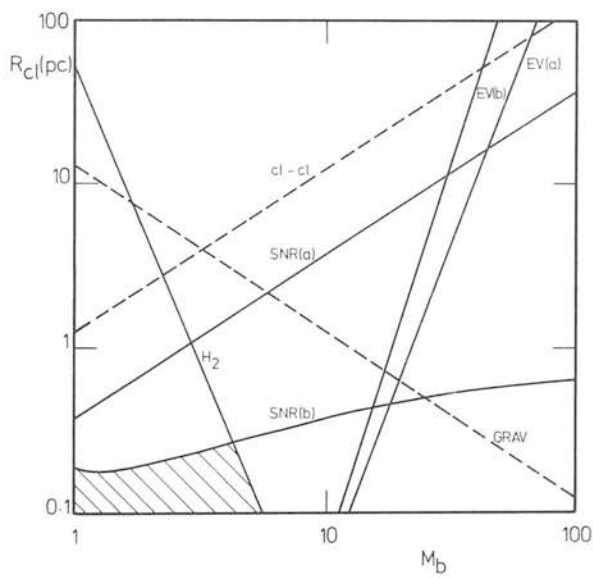


Figure 3. The domain of validity of our simplified model in the initial cloud radius (R_{cl}), incident shock Mach number (M_b) plane for 'cold' clouds when $\gamma_{ic} = {}^5/_3$ and $\gamma_{cl} = 1$. Each labelled line represents one of the various constraints enumerated in Section 3. The label is always on the side of the line where the assumptions of the model are violated. For instance, above the line labelled SNR(a) clouds do not have sufficient time to complete their collapse before being overrun by a second shock. For details and the scaling to other conditions, see text. Shading indicates the region within which none of our assumptions is violated.

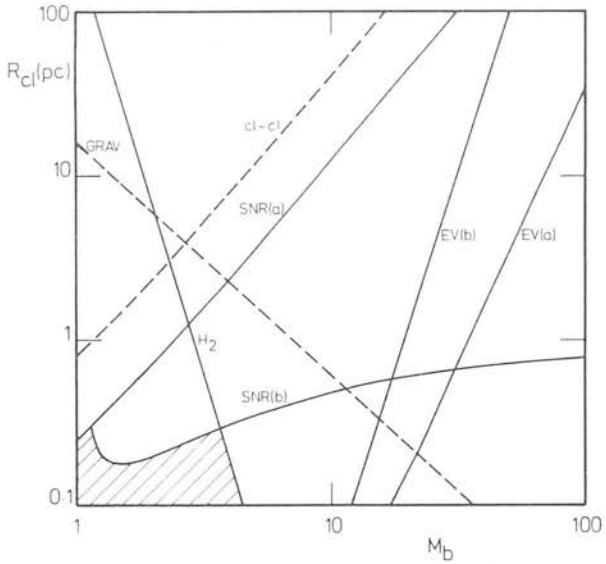


Figure 4. As Fig. 3 but $\gamma_{ic} = 1$ and $\gamma_{cl} = 1$.

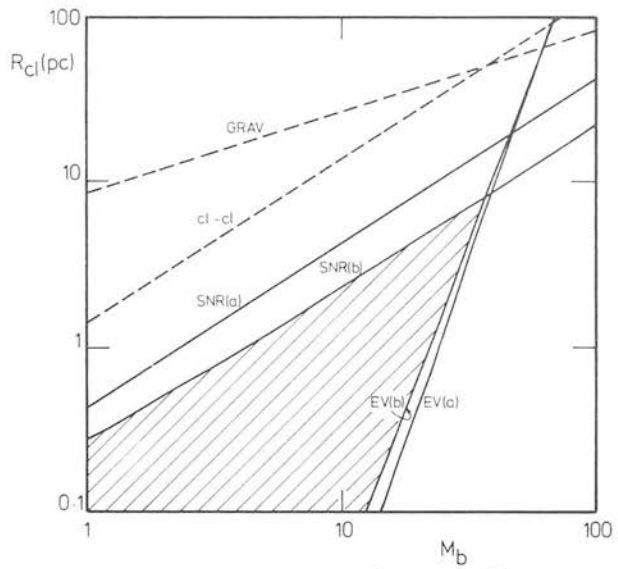


Figure 5. As Fig. 3 but $\gamma_{ic} = {}^5/_3$ and $\gamma_{cl} = {}^5/_3$.

3.1 SHOCK REPETITION

Quite generally the time between successive SNR having radii r smaller than some value r' is

$$t_{sn}(r \leqslant r') \simeq 2.4 \times 10^6 (100\,\mathrm{pc}/r)^3 (10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}}/S)\,\mathrm{yr}$$

where S is the volumetric supernova rate. By adopting a specific model of SNR evolution we may re-express this in a more useful form involving the shock Mach number. For instance if the standard evaporative model of McKee & Ostriker (1977, see also Brand & Heathcote 1982) is appropriate then

$$t_{sn}(M_b > M_b') \simeq 2.4 \times 10^6 (100\,\mathrm{pc}/r_{bc})(10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}}/S)(M_b'/M_{bc})^{9/2}\,\mathrm{yr} \quad M_b' > M_{bc}$$

$$\simeq 5.5 \times 10^6 (100\,\mathrm{pc}/r_{bc})(10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}}/S)(M_b'/M_{bc})^{1.92}\,\mathrm{yr} \quad M_b' < M_{bc}.$$

Here M_{bc} is the Mach number of the blast wave just before the SNR cools to form a dense shell of radius r_{bc},

$$M_{bc} \simeq 2.1\,(E_0/10^{51}\,\mathrm{erg})\,Q_c^{-0.75}\,(S/10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}})^{0.75}$$

$$r_{bc} \simeq 200\,(E_0/10^{51}\,\mathrm{erg})\,Q_c^{0.23}\,(S/10^{-13}\,\mathrm{pc^{-3}\,yr^{-1}})^{0.23}$$

where E_0 is the initial explosion energy and Q_c is the porosity parameter, $0 \leqslant Q_c \leqslant 1$. Notice that because of the very strong dependence of t_{sn} on M_{bc}, the weaker shocks are by far the most common. Clouds have an almost negligible probability of interacting with SNR having $M_b \gtrsim 3M_{bc}$.

The line labelled SNR (a) in Figs 3–5 shows the locus of points for which $t_{sn} = t_{col}$. Similarly, the line labelled SNR(b) shows the locus of points for which $t_{sn} = (t_{col} + t_{exp})$ when F, the axial clump mass fraction equals 0.5. The curves for other values of F are displaced parallel to the R_{cl} axis by a factor $(F/0.5)^{1/3}$. The value of t_{sn} used in these figures has been calculated assuming all SNR, up to the maximum radius at which the remnants overlap and come into pressure equilibrium, contribute to the compression of the clouds. Changing the value of t_{sn} simply displaces these curves parallel to the R_{cl} axis such that $t_{sn}\,C_{cl}/R_{cl}$ = constant. So long as the cloud dynamical time-scales are insensitive to A ($A \lesssim 0.1$, $\gamma_{ic} = 5/3$; $A \lesssim 0.05$, $\gamma_{ic} = 1$) the same is true if C_{cl} is changed. In particular, the scaling to the case of warm clouds, may be made to sufficient accuracy by making an appropriate vertical displacement.

3.2 CLOUD EVAPORATION

Following Cowie & McKee (1977) the evaporation time-scale for a static spherical cloud is

$$t_{ev} \sim 3 \times 10^5\,n_{cl}\,(R_{cl}/\mathrm{pc})^2/\phi_m\,(T_h/10^6\,\mathrm{K})^{5/2}\,\mathrm{yr}$$

where T_h is the temperature of the surrounding medium and ϕ_m is an uncertain correction factor $\simeq 1/3$. To apply this to the case of our collapsing and (except at the initial instant) non-spherical cloud we multiply by a term $(\tilde{a}\,\phi_g)^{-1}$ evaluated at a typical stage in the appropriate phase of the clouds evolution. The term $\tilde{a}$ is the radius of a sphere having the same volume as the collapsed cloud (in units of R_{cl}) while ϕ_g is a correction for the flattening of the cloud estimated from the results of Cowie & Songaila (1977) for spheroidal clouds. During the collapse phase we evaluate $(\tilde{a}\,\phi_g)^{-1}$ when $\tilde{t} = \tilde{t}_{col}/2$ and similarly during the expansion phase when $\tilde{t} = \tilde{t}_{col} + \tilde{t}_{exp}/2$. The values of $(\tilde{a}\,\phi_g)$ derived in the strong shock small A limit are listed in Table 3.

The line labelled EV(a) in Figs 3–5 is the locus of points for which t_{ev} evaluated for the cloud as a whole equals t_{coll} while that labelled EV(b) is evaluated for the axial clump

Table 3. The evaporative reduction factor $(\tilde{a}\,\phi_g)$ (Section 3.2) at a typical point during the collapse, at the end of the collapse and at a typical point during re-expansion as functions of M_b.

γ_{cl}	γ_{ICM}	$(\tilde{a}\,\phi_g)$		
		$\tilde{t}_{col}/2$	$\tilde{t}_{col}$	$\tilde{t}_{exp}/2$
1	1	$[1 - 0.7/(M_b + 1)]^{2/3}$	$M_b^{-1}\,[1 - 1.4/(M_b + 1)]^{2/3}$	$0.2\,F^{1/3}$
1	5/3	0.7	$0.5\,M_b^{-2/3}$	$0.2\,F^{1/3}$
5/3	1	$0.9\,\{1 - [0.5/(M_b + 1)]\}^{2/3}$	$0.6\,[1 - (M_b + 1)^{-1}]^{2/3}$	$0.2\,F^{1/3}\,[C(0)/C_{cl}]^{1/3}$
5/3	5/3	0.7	0.4	$0.2\,F^{1/3}\,[C(0)/C_{cl}]^{1/3}$

when $F = 0.5$. The curves for other values of F are displaced parallel to the R_{cl} axis by a factor $(F/0.5)^{4/3}$. The scaling to other values of T_h is exactly as discussed in the case of t_{SN} (Section 3.1). Since $n_{cl} \propto C_{cl}^2$ it follows that $t_{ev} \propto (R_{cl}/C_{cl})^2$. Thus we may also scale to other values of A (in particular that appropriate to warm clouds) by displacing the plotted curves parallel to the R_{cl} axis such that $t_{ev} C_{cl}/R_{cl} = \text{constant}$.

3.3 IONIZATION

We estimate the time required for the ionization of an initially neutral static spherical cloud exposed to a steady isotropic flux J of ionizing photons to be (assuming a weak R ionization front and neglecting recombination)

$$t_{ion} \simeq 3 \times 10^5 n_{cl} R_{cl} (10^5 \text{ cm}^{-2} \text{s}^{-1}/J) \text{ yr}.$$

As in Section 3.2 above we correct to the case of a collapsing cloud by multiplying by $(\tilde{a}^2 \psi)^{-1}$ where ψ is again a flattening correction. Notice that since t_{ion} and the cloud dynamical time-scales are both proportional to R_{cl} the ratio $t_{ion}/(t_{col} + t_{exp})$ is simply an increasing function of M_b. For the value of J given in Table 2, t_{ion} is always much large than the dynamical time-scales for either warm or cold clouds. Typically, for t_{ion} to be comparable with $(t_{col} + t_{exp})$ when M_b is 1 requires $J \gtrsim 10^6$ photons cm^{-2} s^{-1} for cold clouds and $J \gtrsim 5 \times 10^4$ photons cm^{-2} s^{-1} for warm clouds.

A closely related time-scale which will be of interest in Section 4 is the time required for the formation of an equilibrium warm envelope around a cold core as envisaged by McKee & Ostriker (1977). Under the same assumptions we estimate this to be

$$t_{env} \simeq t_{ion} (n_{env}/n_{cl}) (R_{env}/R_{cl})^3$$

where R_{env} is the envelope radius.

3.4 H_2 FORMATION

The formation time-scale for H_2 molecules is (Spitzer 1978)

$$t_{H_2} \sim 1.6 \times 10^7 (100 \text{ cm}^{-3}/n) \text{ yr}$$

where a cloud temperature ~ 80 K is assumed. The value for n is derived from equations (9) or (10). The existence of a significant steady-state abundance of H_2 requires that the column density through the cloud should be great enough to provide adequate self-shielding in the Werner bands. An appropriate empirical criterion (Savage *et al.* 1977) is $R(t) \gtrsim (200 \text{ cm}^{-3}/n)$pc.

There exists a region in Figs 3–5 within which the tails of cold clouds are sufficiently shielded and t_{H_2} is shorter than the expansion time. The boundary of this region which is only weakly dependent on the clump mass fraction is labelled H_2.

3.5 CLOUD–CLOUD COLLISIONS

We estimate

$$t_{cl-cl} \sim 0.6 \times 10^5 / [(N_{cl}/\text{pc}^{-3}) \langle (R_{cl}/\text{pc})^2 (V_{cl}/\text{km s}^{-1}) \rangle] \text{ yr}$$

where N_{cl} is the number of clouds per unit volume and V_{cl} is the rms cloud velocity. The average denoted by $\langle\rangle$ is taken locally over the entire ensemble of interstellar clouds. Because clouds overrun by the blast wave are accelerated to differing extents depending essentially

on their column density (McKee, Cowie & Ostriker 1978) clouds will tend to be struck from downstream by clouds of lower column density overrun at an earlier stage in the remnants expansion. Similarly, they will tend to run into clouds of higher column density overrun at a later stage. The shapes of the locii along which t_{cl-cl} is equal to the cloud dynamical time-scales are of course identical to those found for t_{sn} above, and obey the same scaling relations. The line for which $t_{cl-cl} = t_{col}$ is labelled cl–cl in Figs 3–5.

3.6 SELF-GRAVITY

The boundary of the region within which the self-gravity of the axial clump exceeds its thermal energy at the end of the collapse phase is indicated by GRAV in Figs 3–5. The clump is idealized as a sphere of mass F (clump fraction) times the mass of the cloud and radius equal to the thickness of the cloud along the symmetry axis at the end of the collapse phase. F has been taken as 0.5 scaling to other values can be obtained by multiplying by a factor $(F/0.5)^{-1/2}$.

3.7 RADIATIVE COOLING

The most useful cooling characteristic for treating the shocks is the cooling column density which can be written entirely in terms of the shock velocity, u_s. An important complication in determining the cooling column density is the ionization of the pre-shock material by radiation emitted by the shock itself. We employ the following relations from McKee & Hollenbach (1980).

$$N_{cool} \simeq 10^{19}\text{–}10^{20}\,\mathrm{cm^{-2}} \qquad 1\,\mathrm{km\,s^{-1}} \lesssim u_s \lesssim 20\,\mathrm{km\,s^{-1}}$$

$$N_{cool} \simeq 6.4 \times 10^{23}(\mathrm{km\,s^{-1}}/u_s)\,\mathrm{cm^{-2}} \qquad 20\,\mathrm{km\,s^{-1}} \lesssim u_s \lesssim 60\,\mathrm{km\,s^{-1}}$$

$$N_{cool} \simeq 8.0 \times 10^{8}(u_s/\mathrm{km\,s^{-1}})^{4.2}\,\mathrm{cm^{-2}} \qquad 60\,\mathrm{km\,s^{-1}} \lesssim u_s \lesssim 1000\,\mathrm{km\,s^{-1}}.$$

Adopting the standard cold cloud parameters we find that for $M_b \lesssim 11$ clouds having $R_{cl} < 0.1$–1 pc have cooling lengths N_{cool}/n_{cl} greater than R_{cl} and hence are 'adiabatic' while outside this region all cold clouds are more nearly isothermal. Conversely, for warm clouds having the standard parameters the cooling length is longer than R_{cl} when $M_b \gtrsim 4.7$ so that warm clouds struck by shocks stronger than this are effectively adiabatic.

4 Discussion

As a consequence of the phenomena discussed in Section 3, the typical cloud evolution will not be as straightforward as that described in the Woodward model. However, it is possible to use the simple model as a basis for the discussion of these phenomena and their effect on clouds, and to suggest what as a consequence is happening to cloud-like material in the supernova-dominated interstellar medium. Specifically, we discuss the mean pressure and shape of clouds and compressional heating of clouds. We also discuss mass-loss rates by shock ablation and evaporation and by warm envelope stripping.

4.1 DEPARTURES FROM PRESSURE EQUILIBRIUM

Perhaps the most striking result apparent in Figs 3–5 is the length of the cloud dynamical time-scales especially that for the re-expansion of isothermal clouds ($\gamma_{cl} = 1$) by comparison with the SNR shock repetition time. This has the important consequence that interstellar

clouds (with the exception of the very smallest) struck by a shock will rarely have time to relax back to pressure equilibrium with their surroundings before encountering a second shock. Furthermore, for clouds of still only modest size, the interaction with this second shock is quite likely to occur during the collapse phase. Clearly then the details of the interaction of any given shock with a cloud will depend in a most complex way on the past history of the cloud. In particular, they will be strongly dependent on the strength and direction of the previous shock(s) which interacted with the cloud, and how long has passed since that last interaction. We may, nevertheless, speculate about the general features of such an interaction on the basis of our simplified model.

First, consider the smaller clouds for which the shock repetition time is only slightly shorter than the expansion time-scale, and longer than the collapse time-scale. We will neglect the presence of the highly flattened sheet like outer region of the cloud and just consider the interaction of an SNR with an entity similar to the rather elongated, over-pressured and asymmetrically expanding cloud tail produced by a previous collision. If this new shock wave is very much stronger than the last shock to strike the cloud, then the expansion can be halted and the subsequent evolution of the cloud will be dominated by this new shock and will be similar to that computed. In the more probable case that both shocks are of similar strength (the majority of shocks have M_b comparable with M_{bc} Section 3.1), the expansion of the cloud is likely to continue. Nevertheless, shocks will be driven into the cloud in a manner similar to that envisaged. If the cloud tail is over-pressured with respect to the pre-shock ICM by a factor η, the 'velocities' of the cloud shocks will be reduced by a factor of $\eta^{1/2}$ in the strong shock, small A (large cloud–intercloud contrast) limit. Obviously, in this case the collapse and re-expansion of the cloud will be rather less ordered than that considered here but our results will still give a rough approximation.

In the extreme opposite limit of clouds for which the collapse time is very much longer than the shock repetition time we expect an entirely different pattern of events. Such clouds will experience a succession of external shocks of similar strength arriving from (by assumption) random directions. Consequently, they will 'see' an effective external pressure comparable to the sum of the post-shock thermal plus ram pressure for a typical shock. Any shocks driven into them will be relatively weak. Clouds somewhat smaller than this limiting case will most likely resemble the 'irregularly vibrating or writhing clouds' envisioned by Cox (1979) which at all times contain several shocks of moderate strength.

4.2 CLOUD MASS-LOSS RATES

As discussed in Sections 2.1 and 2.3 when the flow of the gas behind the external shock is supersonic with respect to the cloud ($M_b \gtrsim 3$, $\gamma_{ic} = {}^5/_3$. $M_b \gtrsim 1.6$, $\gamma_{ic} = 1$ for the 'cold' clouds considered here and correspondingly $M_b \gtrsim 6.0$ or $M_b \gtrsim 2.0$ for the 'warm' clouds) a fraction, $F \sim 0.5$ of the clouds mass is collected into the cloud 'tail' while the remainder is left in a thin 'disc' which is rapidly fragmented by Rayleigh–Taylor instability. This process of shock ablation is a very effective mass-loss mechanism; the mass halving time is roughly the SNR shock repetition time (Section 4.1) $t_{sn} \sim 3 \times 10^5$ yr for the parameters adopted here, making it at least comparable in importance with the related process of cloud fission in cloud–cloud collisions (Chièze & Lazereff 1980).

If the incident flow is subsonic only ~ 0.1 of the cloud mass is confined to the tail. If the flattened disc is still disrupted as seems likely this effectively leads to the complete destruction of the cloud after only one shock passage. Although as shown in Section 3.2 thermal evaporation does not play a major role in the dynamics of the cloud-shock interaction under most circumstances, the evaporation rate is an important parameter in theoretical models

of the ISM as discussed in Section 1. As discussed in Section 3 the compression and flattening of the cloud reduce the evaporation rate relative to that of an undisturbed cloud by a factor $\sim \tilde{a}\phi_g$, which was tabulated at typical stages during the cloud's evolution (Table 3). An equally important effect in fixing the net evaporation rate at late times in the cloud's evolution is the disruption of the 'disc'. If this process leads to the formation of N_{fr} fragments of approximately equal size and possessing similar geometry to that of the original sheet the net evaporation rate for the 'sheet' is increased by a factor $N_{fr}^{2/3}$. As a rough estimate of the average evaporation rate over the cloud's evolution, we write

$$\langle \dot{M}_{ev} \rangle \simeq \dot{M}_{ev}(Sp)\,[(\tilde{a}\phi_g)_{col}\,(t_{col} + N_{fr}^{2/3}\,t_{exp}) + (\tilde{a}\phi_g)_{exp}\,t_{exp}]\,/(t_{exp} + t_{col})$$

where $\dot{M}_{ev}(sp)$ is the evaporation rate of the original undisturbed cloud and $(\tilde{a}\phi_g)_{col}$ and $(\tilde{a}\phi_g)_{exp}$ are respectively the values of the reduction factors from Table 3 evaluated during the collapse phase and the expansion phase. It is readily apparent that for $N_{fr} \sim 5\langle \dot{M}_{ev} \rangle$ is little different from $\dot{M}_{ev}(Sp)$ so that the cloud dynamics may not be of major importance in determining the net evaporation rate.

4.3 COMPOSITE CLOUDS; ENVELOPE STRIPPING

McKee & Ostriker (1977) proposed that interstellar clouds were composite consisting of a warm partially ionized envelope maintained by the interstellar UV and soft X-ray flux, surrounding a cold neutral core. Subsequently, Chièze & Lazereff (1980) have demonstrated that (a) such a cloud envelope is efficiently removed by passing SNR blast waves and (b) that the time-scale for the reformation of such an envelope by photoionization is long compared to the shock repetition time. Our work confirms both these latter findings.

The first conclusion, that of efficient envelope removal, rests on the very different velocities imparted to the low density material of the envelope and the dense material of the core on passing through an SNR blast wave. An approximate estimate of the velocity imparted to a single density cloud on passage through an adiabatic blast wave is obtained simply by application of momentum conservation,

$$V_{cl} \sim 0.9\,(n_c/n_{cl})\,(r_c/R_{cl})\,[(r/r_c)^{-4/3} - M_{bc}^{-2}]\,M_{bc}\,C_{cl} \qquad r \leqslant r_c\,.$$

This expression in which n_c is the mean density in the SNR interior just before the remnant cools makes allowance for the enhancement of the interior density by evaporation (McKee & Ostriker 1977). In the post-cooling ($r > r_c$) case the remnant evolves with approximately constant momentum so this equation with $r = r_c$ still applies. Comparison with more careful calculation (McKee *et al.* 1978) suggests that although this certainly gives a slight overestimate it does give the correct dependence on n_{cl} and R_{cl}. For the standard ISM parameters of Table 1 we obtain $V_{cl} \simeq (80\,\mathrm{pc}/R_{cl})\,\mathrm{km\,s^{-1}}$ for warm clouds and $V_{cl} \simeq (0.5\,\mathrm{pc}/R_{cl})\,\mathrm{km\,s^{-1}}$ for cold clouds.

These values exceed those obtained by Chièze & Lazareff from a similar analysis because they used the expression for M_{bc} given by McKee & Ostriker (1977) which contains a numerical error, leading to an underestimate of M_{bc} by a factor of 2.

We expect this to give a reasonable estimate of the differential velocity imparted to a cloud envelope relative to the core. This clearly indicates that any warm envelope can be removed completely from the average cold cloud of McKee & Ostriker (1977) in a time shorter than the shock repetition time unless some agency such as a magnetic field acts to retain it. Some of the warm material might be trapped at the front of the cloud but this material is likely to be compressed and collected into the dense clump.

We have already given an expression for the growth time of an equilibrium warm envelope in Section 3.3. This differs slightly from that given by Chièze & Lazareff (1980) since they assumed a plane parallel envelope. For the value of the ambient UV flux given in Table 1 this expression gives a value for the envelope growth time a factor of ~ 100 longer than the SNR shock repetition time for the smallest clouds of McKee & Ostriker (1977) which have $R_{\rm env} \sim 2$ pc and $R_{\rm cl} \sim 0.4$ pc. For McKee & Ostriker's cloud spectrum (see their Appendix B) the quantity $R_{\rm env}^3/R_{\rm cl}^2$ has a minimum for clouds ~ 1.6 times as massive as the smallest clouds. Even there, however, assuming the average UV field leads to an envelope growth time a factor of 10 longer than the repetition time. We have compared the values of the shock repetition time and the envelope growth time for all our galactic three-phase ISM Models (Brand & Heathcote 1982) and find that this conclusion holds everywhere. We have also considered the effects of the local enhancement of the UV flux in an SNR cavity during the period of shell formation. Even then the envelope growth time remains long compared to the shock repetition time.

It thus seems that cold clouds cannot retain an equilibrium warm envelope if current estimates of the shock repetition rate are correct. If, as we believe, a significant filling factor for warm material is essential, both to account for observations and for a proper theoretical understanding of the global properties of the ISM we must look for another means of providing it. Cox (1979) has already pointed out that compressional heating of clouds can convert some cold clouds into warm clouds. He has also (Cox 1981) suggested that if a galactic fountain operates, material returned to the disc will generally be warm. We further note that the cooling time from ~ 10^4 to 10^2 K for the quite tenuous material which forms the 'cool dense' shells of SNRs which evolve in a three-phase ISM is very long. Hence, we tentatively suggest these shells as an alternative site for warm cloud formation. This will be discussed further in a future paper.

4.4 CLOUD COMPRESSIONAL HEATING

Cox (1979) has suggested that the PV work done upon clouds when overrun by SNR blast waves may represent an important source of thermal energy for clouds and a significant energy sink for SNR. The compressional work done by the external medium during the collapse phase is particularly simple to work out for our model since the pressures driving the cloud shocks are assumed constant. If the cloud re-expands the work done back on the external medium by the cloud should be subtracted from this (the cloud re-expands at a pressure near $P_{\rm ic}$). The results of such calculation are summarized in Table 4. As is to be expected these results are very similar to those obtained by Cox (1979) for the corresponding case. Our values are slightly higher because Cox assumed the cloud shocks to be driven by the thermal pressure of the external medium while we have included a ram pressure contribution for the main cloud shock which does most of the work.

Table 4. The PV work done on the cloud during the collapse phase $\delta W_{\rm col}$ and by the cloud tail during the re-expansion phase $\delta W_{\rm exp}$ (Section 4.4).

$\gamma_{\rm cl}$	$\gamma_{\rm ic}$	$\delta W_{\rm col}/(4\pi/3 R_{\rm cl}^3 P_o)$	$\delta W_{\rm exp}/(4\pi/3 R_{\rm cl}^3 P_o)$
1	1	$M_b^2\,[M_b^4\,(M_b+3)+3M_b+1]/(M_b+1)^3$	$0.1\,F$
1	$5/3$	$3M_b^2$	$0.1\,F$
$5/3$	1	$2M_b^2\,[M_b^4\,(M_b+2)+2M_b+1]/(M_b+1)^3$	$0.1\,F[C(0)/C_{\rm cl}]$
$5/3$	$5/3$	$2M_b^2$	$0.1\,F[C(0)/C_{\rm cl}]$

5 Conclusion

We have constructed a highly simplified analytical model of the interaction of an SNR blast wave with an interstellar cloud based on the detailed numerical calculations of Woodward (1976, 1979). This model permits extrapolation of Woodward's results to other values of the parameters of the shock and the cloud. Under the assumptions made the dynamics of the interaction are determined completely by three parameters; the Mach number of the incident shock, M_b, the cloud radius, R_{cl}, and the ratio of the cloud to intercloud sound speeds A. Further, when A is small as is often the case the dependence on this parameter is quite weak. The scaling with R_{cl} and M_b is found to be quite simple.

Any theoretical advance in this area is likely to require a fully three-dimensional computation with at least as great a resolution as that employed by Woodward in his two-dimensional calculations (the utility of calculations at lower resolution is put in question by the importance of small-scale instabilities of the cloud boundary).

We have imposed the conditions believed to prevail in a 'realistic violent' three-phase interstellar medium upon our model in order to assess the importance of processes not included in the model (e.g. a high shock repetition rate, thermal evaporation etc) and to determine the effect of the cloud dynamics on processes important to the global properties of the ISM (evaporation, cloud crushing etc). We conclude that:

(i) If current estimates of the rate at which a given interstellar cloud encounters SNR shocks are correct, then a shocked cloud is rarely allowed sufficient time to relax back to pressure equilibrium with its surroundings before being overrun by another shock. Large departures from pressure equilibrium are expected, and the dynamics and properties of a given cloud will generally depend in a complex way on its past history. Nevertheless, we expect our model to be qualitatively correct under a wide variety of conditions.

(ii) The final product of the interaction is a highly flattened disc of material which is easily disrupted by Rayleigh–Taylor instabilities, and a rather extended cloud 'tail' expanding in the downstream direction. Each component is generally expected to contain ~ ½ the mass of the original cloud. This process of shock ablation may represent an important mechanism for cloud disruption. If it is then the half life of the cloud is comparable to the shock repetition time.

(iii) Although the flattening and compression of clouds overrun by SNR shocks reduces the evaporation rate below that for a static spherical cloud this is probably compensated by enhancement of the evaporation rate due to the fragmentation of the disc.

(iv) Warm envelopes are efficiently removed from cold cores by SNR shocks and require a time considerably longer than the shock repetition time for their recreation by ionization. Nevertheless, a warm low density component of the ISM is observed, and is necessary for proper theoretical understanding of its global properties. Three-phase models of the ISM are almost certainly sufficiently 'robust' to overcome this difficulty (we have suggested some tentative mechanisms for the production of warm material, and will pursue this matter further in a future paper).

(v) The expected rate of PV work done by SNR on clouds for our model is similar to that found by Cox (1979).

While the assumption of pressure equilibrium is useful in simplifying the study of interstellar conditions, the prevalence of shock waves in the ISM requires that considerable caution be exercised in employing it. Many, or most, clouds should be observed to have structures rendered complex by being associated with 'tails', with shock ablated material, and with partially reformed envelopes, and by having one or more shock waves traversing

their interiors. It appears that the 'warm envelope' of the three-phase model does not exist, but that if warm material is created elsewhere – as it must be – then despite these complications a three-phase model can adequately explain the current state of the ISM.

References

Bradley, J. N., 1962. *Shock Waves in Chemistry and Physics*, Methuen, London.

Brand, P. W. J. L. & Heathcote, S. R., 1982. *Mon. Not. R. astr. Soc.*, **198**, 545.

Bychcov, K. V. & Pikel'ner, S. B., 1975. *Sov. Astr. Lett.*, **1**, 14.

Chièze, J. P. & Lazereff, B., 1980. *Astr. Astrophys.*, **91**, 290.

Cowie, L. L. & McKee, C. F., 1977. *Astrophys. J.*, **211**, 135.

Cowie, L. L. & Songaila, A., 1977. *Nature*, **266**, 501.

Cox, D. P., 1979. *Astrophys. J.*, **234**, 863.

Cox, D. P., 1981. *Astrophys. J.*, **245**, 534.

Currie, I. G., 1974. *Fundamental Mechanics of Fluids*. McGraw-Hill.

Dyson, J. E. & Gulliford, P., 1975. *Astrophys. Space Sci.*, **37**, 477.

Gerwin, R. A., 1968. *Rev. Mod. Phys.*, **40**, 652.

Imshenik, V. S., 1960. *Soviet Phys. Dokl.*, **5**, 253.

Kahn, F. D., 1972. In *Interstellar Matter, Second Advanced Course of the Swiss Society of Astrophysics and Astronomy*, p. 343, Geneva Observatory.

McKee, C. F. & Cowie, L. L., 1975. *Astrophys. J.*, **195**, 715.

McKee, C. F., Cowie, L. L. & Ostriker, J. P., 1978. *Astrophys. J.*, **219**, L23.

McKee, C. F. & Hollenbach, D. J., 1980. *A. Rev. Astr. Astrophys.*, **18**, 219.

McKee, C. F. & Ostriker, J. P., 1977. *Astrophys. J.*, **218**, 148.

Savage, B. D., Bohlin, R. C., Drake, J. F. & Budich, W., 1977. *Astrophys. J.*, **216**, 291.

Sgro, A. G., 1975. *Astrophys. J.*, **197**, 621.

Silk, J. & Solinger, A., 1973. *Nature Phys. Sci.*, **244**, 101.

Spitzer, L. Jr., 1978. *Physical Processes in the Interstellar Medium*, Wiley, New York.

Thorogood, P. J., 1979. *PhD thesis*, University of Leeds.

Woodward, P. R., 1976. *Astrophys. J.*, **207**, 484 (W76).

Woodward, P. R., 1979. *The Large Scale Characteristics of the Galaxy, IAU Symp. 84*, p. 159, ed. Burton, W. B., Reidel, Dordrecht, Holland (W79).

Zel'dovich, Ya. B. & Raizer, Yu. P., 1968. *Elements of Gasdynamics and the Classical Theory of Shock Waves*, Academic Press, New York.

FORCED DIFFUSION: A MECHANISM TO CREATE DUST CLOUDS

PETER W. J. L. BRAND
Dept. of Astronomy, University of Edinburgh, Scotland

(Received 29 July, 1974)

Abstract. Photodesorption from dust grains by an anisotropic radiation field will lead to much greater momentum transfer than that associated with radiation pressure. This can cause the segregation of dust to the centres of moderately dense interstellar clouds.

1. Introduction

Reddish (1971) has proposed a mechanism which will cause gas/dust segregation in clouds round OB stars, whereby photodesorption imparts a net outward momentum to the grains with respect to the gas surrounding the star.

The idea developed here is that this process will operate in the general interstellar medium causing dust cloud structure to be enhanced with respect to the gas clouds.

The mechanism is as follows. A photon desorbs a particle (atom, molecule) from the surface of a grain which by consideration of time-scales for grain rotation and for desorption will depart from the grain into the hemisphere containing the direction from which the photon arrived. A large part of the photon energy is converted into particle energy, and since for a given energy, momentum increases with mass, the net momentum imparted to the grain is much greater than that carried by the photon.

2. Momentum due to Desorption

The case will be considered of a grain at the edge of a large cloud of optical depth τ in an isotropic field of radiation I_0. The unit outwards normal to the cloud boundary, $\mathbf{k}$, is taken as the axis, and θ is an angle measured with respect to this with $\mu = \cos\theta$. Thus the radiation field seen by the grain is

$$I = I_0 \qquad \mu \geqslant 0$$
$$I = I_0 e^{-\tau} \qquad \text{otherwise.}$$

(This obviously requires slightly peculiar cloud geometry for the simplification to be true, but it is a reasonable approximation.)

Consider an element of grain surface which receives N photons, each of which releases p ($\leqslant 1$) particles of mass m with velocity V isotropically. The mean momentum ejected will be normal to the surface, and will be $\frac{1}{2}mV$ per particle, so $\frac{1}{2}pmVN$ is the mean momentum ejected normal to the surface.

Astrophysics and Space Science **33** (1975) 231–234.

If radiation $I\,d\Omega$ in direction **n** falls on a sphere of radius a then a surface area $2\pi a^2 \sin\phi\,d\phi$ with projected area $2\pi a^2 \sin\phi\cos\phi\,d\phi$ receives $N = 2\pi a^2 \cdot \sin\phi\cos\phi\,d\phi \times (I/h\bar{\nu})\,d\Omega$ photons per second, where $h\bar{\nu}$ is the mean photon energy, and ϕ the angle with respect to $-\mathbf{n}$.

The momentum released in desorbed particles can be resolved into components perpendicular to **n** which cancel by symmetry, and antiparallel to **n**, proportional to $\cos\phi$. Thus the back momentum released by the sphere due to $I\,d\Omega$ is

$$\tfrac{1}{2}pmV \cdot 2\pi a^2 \int_0^{\pi/2} \sin\phi\cos^2\phi\,d\phi(I/h\bar{\nu})\,d\Omega.$$

More rigorously, we should use $Q\pi a^2$ instead of πa^2 where Q is the absorption factor.

Hence the total momentum ejected by a spherical particle due to an axially symmetric radiation field with axis of symmetry **k** is the **k**-component of the above expression integrated over 4π. In the case considered, the total momentum ejected per second due to the radiation field is

$$\dot{m}V = \tfrac{1}{3}pmVQ\pi a^2(\pi I_0/h\bar{\nu})(1 - e^{-\tau}).$$

Away from the cloud, the radiation field I_0 is isotropic, and $\pi I_0/h\bar{\nu} = cn_{\rm ph}/4$ where $n_{\rm ph}$ is the interstellar photon density.

From the figures for the interstellar radiation measured by Witt and Johnson (1973) in the wavelength range 1200 Å–3000 Å a figure for $n_{\rm ph}$ of approximately 10^{-2} photons $\rm cm^{-3}$ may be deduced. For $h\bar{\nu} \simeq 6$ eV, $V \simeq 20\ \rm km\ s^{-1}$ for hydrogen molecules.

The total desorption rate is

$$\dot{m}_d = mQ\pi a^2 pJ\ \rm g\ s^{-1},$$

where m is the mean mass of desorbed species, and

$$J = \int (I/h\bar{\nu})\,d\Omega = 0.5cn_{\rm ph}(1 + e^{-\tau}).$$

Since the particles leave with mean velocity V, the ratio of net momentum ejected by the grain per second to $\dot{m}_d V$ is

$$\dot{m}V/\dot{m}_d V = 1/6(1 - e^{-\tau})/(1 + e^{-\tau}) = f(\tau).$$

The adsorption rate of hydrogen is $\dot{m}_a = m_H S\pi a^2 n_H c_{\rm ad} p_s$, where $S\pi a^2$ is the collision cross section and p_s is the sticking probability. Taking $m \simeq m_H$, $S \simeq 1$, $p/p_s \simeq 0.1$, $c_{\rm ad} = 3 \times 10^5\ \rm cm\ s^{-1}$, $Q = 1$, $\dot{m}_d/\dot{m}_a = 1$ at $n_H \simeq 100\ \rm cm^{-3}$, and is greater than this at lower densities, and much greater than this for atoms other than hydrogen. Thus one may envisage a grain consisting of a hard non-desorbable substrate (silicate?) on which particles adsorb, and after a short time, possibly allowing for surface reaction between adsorbed species, are desorbed at a rate which is therefore regulated by the adsorption rate.

3. Motion of the Dust Grain

The forces on the grain are (i) subsonic drag force $F_D = K\pi a^2 \varrho c_{ad} v$, where v is the velocity of the grain with respect to the gas of density ϱ, and $K = 4/3$ for neutral collisions, or $K = 4.5 \times 10^{-5} (q/aT)^2$ for Coulombic collisions where q is the grain charge (Baines *et al.*, 1965; Spitzer, 1968); (ii) radiation pressure force $F_R = 0.5 Q\pi a^2 h\bar{\nu} n_{ph}(1 - e^{-\tau})$; and (iii) the desorption momentum force $\dot{m}V$.

The dust grain is assumed to have constant mass m_0, and the desorption rate to be equal to the adsorption rate. The equation of motion is $m_0 \dot{v} = \dot{m}V - F_D + F_R$.

Now since

$$\dot{m}_d = \dot{m}_a = m_H p_s S\pi a^2 n_H c_{ad} = 0.5 m Q\pi a^2 p c n_{ph}(1 + e^{-\tau}),$$

$$F_D = \dot{m}_d(\mu K/Sp_s)v,$$

and

$$F_R = \dot{m}_d f(6h\bar{\nu}/mc),$$

$$m_0 \dot{v} = \dot{m}_d[f(V + 6h\bar{\nu}/mc) - (\mu K/Sp_s)v].$$

Since V is the desorption velocity and $h\bar{\nu}/mc$ is the velocity the desorbed particle would have if it possessed the incident photon's momentum, this latter term is negligible. Putting $t_d = m_0/(\dot{m}_d \mu K/Sp_s)$ and $V_0 = fVSp_s/\mu K$

$$v = V_0[1 - \exp(-t/t_d)].$$

If we take the density of grain material as 1 g cm^{-3}, $p_s = 1$, $s = 1$, and if also we consider Coulombic drag as small, so that $\mu K/Sp_s \doteqdot 2$, and suppose that m (the desorbed particle mass) $= 2m_H$, then $t_d = 5000\, a$ yr, where a, the grain radius, is measured in units of 0.1 μm, a typical grain size. If Coulombic drag dominates, this time should be multiplied by the corresponding K, approximately. V_0 is 2 km s^{-1}, reduced correspondingly if Coulombic drag dominates. Thus in 10^6 yr, dust grains at the edge of moderately dense dust clouds and subject to the near UV radiation field in the galaxy could migrate into the cloud some 2 pc.

4. Conclusions

The time-scale for achieving the final velocity of the dust grain due to this mechanism increases, and the final velocity decreases, both by large factors if the grain is highly charged. However, inside a reasonably cool cloud, it appears that the forced diffusion will significantly segregate grains in the central regions of gas clouds, whose outer regions consequently are less shielded by the grains, and can heat and dissipate. Since the optical depth in dust will increase as the grains crowd together, the rate of the effect is enhanced (the figures in the last section assumed a fairly large optical depth, $\tau > 2.5$). Moreover, there is no natural limit to the process, which ought therefore to create continuously more and more compact dust regions in suitable parts of the inter-

stellar medium. The process will of course be reversed by cloud collisions or other dramatic events (such as supernova explosions) which can disrupt cloud structures.

It is suggested elsewhere (Brand and Zealey, 1975) that this is the origin of the extensive filamentary structure seen in absorption in the galactic plane.

Acknowledgement

I thank Dr T. J. Lee for useful discussions.

References

Baines, M. J., Williams, I. P., and Asebiomo, A. S.: 1965, *Monthly Notices Roy. Astron. Soc.* **130**, 63.

Brand, P. W. and Zealey, W. J.: 1975, *Astron. Astrophys.*, in press.

Reddish, V. C.: 1971, *Nature* **232**, 40.

Spitzer, L.: 1968, *Diffuse Matter in Space*, Interscience.

Witt, A. N. and Johnson, M. W.: 1973, *Astrophys. J.* **181**, 363.

THE ASTROPHYSICAL JOURNAL, 223:161–167, 1978 July 1

THE PROPERTIES OF WEAK SHOCKS IN H II REGIONS

P. W. J. L. BRAND* AND JOHN S. MATHIS
Washburn Observatory, University of Wisconsin–Madison
Received 1977 November 14; accepted 1978 January 17

ABSTRACT

The structure of the cooling region behind weak shocks (speeds ≤ 30 km s^{-1}, isothermal Mach number, $M_0 \leqslant 3$) in H II regions is determined by numerical integration. The cooling region behind the shock is too thin to be spatially resolved even in nearby H II regions. The observable quantities are, therefore, the various line strengths in the equilibrium regions ahead of and behind the shock. The line strengths can be approximately understood in terms of the effects of the compression upon their parent ionic abundances. The increase in strength of lines of the principal stage of ionization of an element is about the same as the increase in the hydrogen line strengths. For lines from stages lower than the principal stage, the increase is greater than for the hydrogen lines, and vice versa for higher stages. Qualitative criteria are given for distinguishing shocks from changes in path length or discontinuities in extinction from point to point in the nebula. The determination of M_0 from line ratios is discussed; the problem is the separation of the contribution of unshocked material superposed upon the shock. The best procedure is the simplest: obtain a lower limit to M_0 from the increase in a hydrogen line across the shock. Other, more elaborate procedures are not more accurate.

Subject headings: nebulae: general — shock waves

I. INTRODUCTION

Several pieces of evidence suggest that shock waves should be present in H II regions. The classical paper of Wilson *et al.* (1959) demonstrated extensive line splitting in the Orion Nebula, indicating that there were many regions in that nebula where gas flows moved with sonic or supersonic velocity relative to one another in the same line of sight. Similar situations have been observed in many other bright nebulae (e.g., Elliott and Meaburn 1975*a*, *b*, *c*). If these gas streams are in the same neighborhood along the line of sight, shocks must form where they collide.

Dyson (1968) showed that neutral globules embedded in an H II region explain many properties of nebulae. In Dyson's model, the gas flow becomes supersonic some distance from the globule and the interaction between flows from adjacent globules will cause shocks to appear.

Marsh (1970) and Goldsworthy (private communication) have discussed the interaction of an ionization front and a contact discontinuity, and in particular the situation in which a fast-moving ionization front overruns a region of greatly enhanced density. The resulting flow includes a shock reflected back into the direction of approach of the ionization front, traveling at a relatively low Mach number.

We have calculated numerical models of the structure behind rather weak shocks in H II regions. Since a shock is geometrically thin in comparison with any other relevant lengths in an H II region, plane-parallel geometry is assumed without loss of generality. We will show that the shock is always optically thin, so the radiation field in the unshocked nebula applies throughout the region immediately behind the shock as well.

We discuss the models in § II. Results will be discussed in § III, initially in terms of a rather simple treatment which is useful in interpreting the results of more detailed calculations. Criteria are given for distinguishing a shock from other possible causes of a discontinuity in nebular surface brightness. Finally, we discuss the difficulties of estimating the Mach number from observational line strengths.

II. THE CALCULATIONS

The shock is characterized by the isothermal Mach number $M_0 = v_0/c$, where $c^2 = P/\rho$ in the unshocked material, and v_0 is the speed of the shock with respect to the undisturbed material ahead of it. The value of c in a typical H II region ($T = 8500$ K) is 10.7 km s^{-1}. We consider only shocks with $M_0 \leq 3$, implying shock speeds of at most 30 km s^{-1}. Stronger shocks, in which the material is heated so much that collisional ionization is important, have been considered by several authors (e.g., Dopita 1976 and references therein).

Shocks in H II regions are often referred to as "isothermal" because the strong cooling in H II regions reduces the postshock temperature to near the preshock temperature in a relatively short distance. The cooling length behind the shock may be estimated from the cooling rates given in Osterbrock (1974). Our detailed calculations give more accurate values, discussed later. One finds that the optical depth through the shock cooling region is at most a few hundredths, independent of the density.

* On leave from the Department of Astronomy, University of Edinburgh.

We choose a coordinate frame in which the shock is stationary, so that explicit time derivatives do not occur. The equations to be satisfied are mass conservation, momentum conservation, and the first law of thermodynamics. In addition, the ionization-rate equations for the principal species are solved point by point. The radiation field is taken to be that in the undisturbed nebula, since the shock is optically thin. For the same reason, the direction of the shock propagation is not important. The radiation field consists of F_ν^*, the attenuated flux from the central star, and $4\pi J_{\nu d}$, the diffuse radiation produced by the nebula. The radiation fields are taken from models of Mathis (1978), in which the nebular density (assumed uniform), chemical composition, and central star's spectrum are given. Radiation transport, using an iterative procedure, was used to calculate the mean intensity of radiation at each point. The chemical elements included, and their abundances by number, relative to H, where He, 0.1; N, 7.6×10^{-5}; O, 5.25×10^{-4}; Ne, 3.2×10^{-4}; and S, 1.5×10^{-4}. Stellar energy distributions, all with $\log g = 4$, were taken from the $T = 35{,}000$ K, plane-parallel model of Kurucz, Peytremann, and Avrett (1974), the 40,000 K model of Balick and Sneden (1976), kindly furnished by Dr. Balick, and the 45,000 K model of Hummer and Mihalas (1970). Our models have the same chemical composition as the Balick and Sneden model atmospheres. A fraction ϕ of the nebular volume was assumed filled with material. For most models, ϕ was taken to be unity, but values of ϕ down to 0.01 were sometimes used.

The radiation field was taken from a nebular model of assumed uniform density, which was identified with the density of the undisturbed gas ahead of the shock. Densities ranging from 10 to 10^4 H atoms cm^{-3} were considered.

Naturally, the shock causes an increase in density (approximately by a factor of M_0^2) behind it. Perhaps it seems inappropriate to apply the radiation field predicted by a uniform model to a calculation involving density variations caused by the shock, but such is not the case. The spectral distribution of the ionizing diffuse radiation depends very weakly on the temperature and even more weakly on the density. Similarly, as we discuss next, the intensity of the diffuse radiation coming from the dense shocked gas to the region of the shock is approximately the same as the intensity of radiation coming to it from the unshocked gas. The emission is higher ($\propto N^2$) in the dense shocked gas, but also the absorption is greater in just the same proportion. The absorptions are determined mainly by neutral H or He, so the absorption coefficient is proportional to XN, where X is the fraction of neutral H or He. For a given total radiation field, $X \propto N$ from the ionization balance, so $k_\nu \propto N^2$, as is the emissivity. Hence the model without the shock has the same local radiation field, in the region of the shock, as would a model containing a shock.

Let us consider a shock of given Mach number in a nebula with a given radiation field. In order to calculate the radiation emitted by the shocked material, we must know the temperature, density, and ionization fraction at each point. These quantities are obtained by integration of an appropriate set of simultaneous differential equations. The starting conditions are known because the shock initially heats and compresses the unshocked material adiabatically. For the modest Mach numbers considered in this paper, collisional ionizations are negligible in the shocked material, so the shock does not instantaneously change the ionic abundances. We also assume that the shock does not change the chemical composition of the gas appreciably through destruction of interstellar grains, although there is evidence that grains are destroyed in shocks (Shull, York, and Hobbs 1977; Jura 1976). However, the lack of the 3.1 μm feature of H_2O in the spectrum of interstellar dust (Harris 1976) suggests that there is little water in the grains until there is a very large visual extinction. A negligible amount (less than 10%) of oxygen would be in the silicates, if the Si/O abundance is anywhere near solar. Hence, the destruction of grains in a shock might appreciably increase the amounts of gaseous heavy elements (Na, Ca, Fe, Mg, etc.), but not of oxygen, the important coolant. Sulfur is hardly depleted in the interstellar medium (Spitzer and Jenkins 1975) and is therefore not likely to be in grains. If H_2O is not frozen into grains, it is unlikely that NH_3 would be, and hence N would be unaffected by grain destruction. We therefore neglect the effects of grain destruction on the gas, although the evaporation of extensive coatings of H_2O and NH_3, if present, could have important effects.

The shock does not initially affect the specific stagnation enthalpy (see, e.g., Landau and Lifshitz 1959). The enthalpy, $g(r)$, is thereafter determined from the net cooling rate per unit volume, $\Gamma(r)$, and the mass flux, j, a constant, by

$$dg/dr = \Gamma j^{-1}, \tag{1}$$

where Γ is a function of the ionic abundances through collisional excitation of forbidden lines, but also includes free-free and recombination cooling (Osterbrock 1974). For a perfect gas with a ratio of specific heats $c_p/c_v = \gamma$,

$$g = \gamma(\gamma - 1)^{-1}P/\rho + v^2/2, \tag{2}$$

where v is the local speed of the gas in the frame in which the shock is stationary.

Just behind the shock, the temperature, T_S, and density, N_S, are given in terms of the unshocked values, T_0 and N_0, by the usual adiabatic conservation relations (e.g., Landau and Lifshitz 1959, p. 331):

$$T_S = (\gamma + 1)^{-2}[2(\gamma - 1)M_0^2 + 6\gamma - \gamma^2 - 1 - 2\gamma(\gamma - 1)M_0^{-2}]T_0 \tag{3}$$

$$N_S = N_0(\gamma + 1)M_0^2[(\gamma - 1)M_0^2 + 2\gamma]^{-1}, \tag{4}$$

provided $M_0 \geq \gamma^{1/2}$ (or, the adiabatic Mach number exceeds unity). The behavior for various values of M_0 can be most easily understood from Axford's (1961)

discussion of plane-parallel ionization fronts. If $\gamma^{1/2} \geq M_0 > 1$, no true shock (i.e., discontinuity in T and N) exists. Instead, a slow compression occurs, with a length scale equal to the cooling length rather than the collision mean free path. In this region the density increases, the flow speed drops toward the sonic value, and temperature first *increases* as the gas loses energy. The speed reaches the isothermal sonic point, and the temperature finally relaxes to equilibrium. The temperature rise depends only on M_0 and the adiabatic exponent γ. The temperature increase above T_S is $T_S(\gamma - 1)^2(4\gamma)^{-1}$ for $M_0^2 = \gamma$, and is less for $M_0^2 < \gamma$.

For $M_0^2 > \gamma$, a true adiabatic shock exists, but if $M_0^2 < (3\gamma - 1)(3 - \gamma)$ a temperature rise after the shock will occur.

For our numerical models, we used $\gamma = 5/3$. The spectrum emitted by the shocked material is calculated by numerically integrating equation (1) and 12 ionization-rate equations for H^0, He^0, N^0, N^+, N^{+2}, O^0, O^+, Ne^0, Ne^+, S^+, S^{+2}, and S^{+3} by a Hamming-type procedure (Ralston and Wilf 1960), using the continuity of enthalpy and ionic abundances across the shock as initial conditions. The stiffness of the equations introduced by the short time constant in some of the ion species was dealt with by setting these species to equilibrium at any integration step at which the time constant was sufficiently short. Variation of the integration parameter indicated that the relative ion abundances (except for extremely small abundances), temperature, and density were accurate to 1% or 2%. In extreme cases, the cooling length changed by 10% as the integration parameter was varied. In view of the smallness of this length in observable terms, this was regarded as satisfactory accuracy.

III. DISCUSSION AND RESULTS

There are two possibilities for observationally detecting shocks in nebulae. If a shock is seen edge-on, one might directly observe the cooling region behind the shock. Alternatively, even with poor spatial resolution, one should observe the strong density (and hence brightness) contrast across the shock. Although the shocked material relaxes into a normal (equilibrium) H II region at a higher density than the unshocked, the continuity of the radiation field across the shock leads to a correlation of the line intensities emitted in the shocked material with those emitted in the unshocked, thus providing diagnostics for the presence of the shock.

Let the cooling length be defined as the length within which the temperature drops from the peak value to halfway between the peak and final values. For a wide variety of conditions in density, stellar temperature, and filling factor, the hydrogen column density through this cooling length depends only on M_0 and the ionization stage of oxygen. Our integrations show that this column density is about 1.3×10^{17} atoms cm^{-2} for $M_0 = 2.5$ providing that O^{++} is the main coolant. For $M_0 = 1.5$, the column density is about 5×10^{17} atoms cm^{-2}. If O^+ cooling dominates, the H column density is about 3 times as large. For different values of $A(O)$, the oxygen abundance, by number, relative to hydrogen, these column densities should be scaled by $3 \times 10^{-4} A(O)^{-1}$. Unfortunately, these column densities imply quite small distances; only $5 \times 10^{16} N_0^{-1}$ cm for $M_0 = 2.5$, and about 5 times larger for $M_0 = 1.5$. At a distance of 1 kpc, the cooling zone subtends only 0.17″ for $M_0 = 1.5$, $N_0 = 100$ cm^{-3}. Greater densities, found in all bright nebulae, make the cooling region proportionally smaller. Thus we conclude that it will *not* be possible to resolve the cooling region behind the shock directly, for nearby nebulae, even with satellite resolution (say, 0.1″). For this reason, we do not give plots of the details of our calculations as a function of distance behind the shock.

For the rest of this paper, we will discuss the comparison of line intensities across an observed intensity discontinuity which may be a shock seen edge-on. Alternatively, such an intensity discontinuity can be caused by a sudden change in line-of-sight path length through the emitting material while the line emissivity remains relatively constant. The discontinuity can also be caused by our seeing the sharp edge of an obscuring cloud projected against uniformly bright nebular material. In § III*a*, we will discuss how various line emissivities change across a shock, on a simplified picture which is found by direct calculation to be fairly accurate. In § III*b*, we give results of our numerical calculations and discuss how to tell, under some circumstances, a shock from a path length or extinction discontinuity.

a) General Considerations

Let a subscript 1 denote conditions far downstream, after cooling has occurred behind the shock and the material has reached thermal equilibrium, and a subscript 0 refer to material ahead of the shock. As is well known, in a truly isothermal shock, the conservation conditions require $N_1 = M_0^2 N_0$ (as can be seen from eq. [4], with $\gamma = 1$). Since j_H, the volume emissivity of a hydrogen line, is proportional to N^2, we will have a ratio $j_{H1}/j_{H0} = M_0^4$, which is a large number for even mild shocks. However, collisional de-excitation of forbidden lines tends to quench the cooling, and usually T_1/T_0 is greater than unity. The difference $T_1 - T_0$ is greatest for the shocks which span a density range where quenching changes from a small effect to a large one. In nearly every case calculated, the dimensionless change t, defined as $(T_1 - T_0)/T_0$, was less than 0.1. A figure of 0.05 is a reasonable average figure. By considering the conservation of mass and momentum, one can show that, to first order in t, the density ratio, η, becomes

$$\eta = N_1/N_0 \approx M_0^2[1 - t(1 - M_0^{-2})^{-1}] . \quad (5)$$

As the shocked material cools, the density ratio increases from a value given by equation (4), just behind the shock, to a value given by equation (5) far downstream.

Most of the line emissivities are dominated by this increase, although some strongly temperature-sensitive

lines, such as the [O III] line λ4363, may peak in the cooling region.

Suppose one observes a sudden increase, by a factor of R_H, in the intensity of a hydrogen line between two adjacent areas in a nebula. It might be caused by a shock being seen edge-on, so that the material on one side of the shock is compressed by a factor η. If the geometry is favorable, the change in intensity is entirely caused by an increase in emissivity with the shocked material. Any uniform foreground reddening across the shock will not affect R_H.

Let the emissivity of the hydrogen line be j_H, and $d(\log j_H)/d(\log T) = -s_H$. Then straightforward manipulation yields, to first order in the logarithmic temperature difference, t,

$$R_H = j_{H1}(j_{H0})^{-1} = \eta^2(1 - s_H t) \quad (6a)$$

$$= M_0^4\{1 - [s_H + 2(1 - M_0^{-2})]t\}. \quad (6b)$$

Similarly, for an unquenched forbidden line, the ratio of emissivities after the shock to before is

$$R_L = j(L)_1/j(L)_0 = \eta^2(X_1 X_0^{-1})[1 + (K - \tfrac{1}{2})t], \quad (7)$$

where the X's are the ionization fractions of the ion, and $K = T_{ex}/T_0$, where kT_{ex} is the excitation of the upper level of the lines. The ion fraction (X_1/X_0) also depends on the density ratio η. Let s_L be the temperature exponent of the recombination coefficient into the stage of ionization of the line. It is numerically about 0.7 in cases of interest. If one stage of ionization (not necessarily that of the line) dominates, a simple expression is possible. For instance, if $X_0 \approx 1$, then $X_1 \approx 1$ also; if $X_0 \ll 1$, then

$$X_1 X_0^{-1} = \eta^{\pm 1}(1 \pm s_L t), \quad (8)$$

where the plus sign is taken if the line arises from a stage one higher than the dominant one, and the minus sign is taken if the line is from the stage lower than the dominant. Combining equations (7) and (8), we obtain

$$R_L = \eta^{2+n}[1 + (K - \tfrac{1}{2} - n s_L)t] \quad (9a)$$

$$= M_0^{2(2+n)}\{1 + [K - \tfrac{1}{2} - n s_L + (2 + n)(1 - M_0^{-2})^{-1}]t\}, \quad (9b)$$

where $n = 0$ if $X_0 = 1$, $n = -1$ if the element is mostly in the next lower stage of ionization, and $n = +1$ if the element is mostly in the next higher stage of ionization.

For a line which is fully quenched even before the shock, the factor of $(2 + n)$ in equations (9) should be replaced by $(1 + n)$, and the $\frac{1}{2}$ should be removed from the coefficient of t.

The form of equations (9) is not changed if the ionization equilibrium of the ion in question is affected by charge exchange with hydrogen or helium rather than by photoionization. In particular, the density dependence is the same. In the limit that charge exchange dominates, it can be shown (see Osterbrock 1974) that the recombination exponent s_L should be replaced by $(s_B - \chi/kT)$. Here

$$s_B = d \log \alpha_B / d \log T = 0.80,$$

where α_B is the recombination coefficient of hydrogen to the second and higher levels, and χ is the energy difference between the lower stage of ionization plus a proton and the higher stage and a hydrogen atom.

Equations (6) and (9) predict quite large increases in line strengths across the shock, mostly from the high powers of M_0 involved. For example, in the typical nebular situation where most of oxygen is O^{++} and the remainder is almost all O^+, the [O III] lines increase as M_0^4, and [O II] increases as M_0^6 across the shock. If O^+ is instead dominant, the increases for [O III] and [O II] are M_0^2 and M_0^4, respectively. An H line increases as M_0^4 in both cases. Hence the dominant stage of ionization of an element in the unshocked material can be inferred from the behavior of lines of that element relative to hydrogen lines.

We next consider the numerical results in view of this simplified analysis.

b) How Do We Identify a Shock?

In the preceding subsection we showed that R_H, the ratio of H line emissivity after the shock to that before, should be approximately M_0^4 (eq. [6b]), and similarly for other ions. Let us express the actual results, averaged for calculations with $M_0 = 1.5$ and $M_0 = 2.5$, by defining an exponent q_L for each line: $q_L = \log(R_L)/\log(M_0)$, so that the line ratio across the shock is $(M_0 ** q_L)$. Table 1 gives values of q_L for various lines, for various regions of the nebula differing in ionic abundances. The first column gives a label for each region. This label is used for brevity later (in Table 2). The other columns are self-explanatory.

One interesting feature of Table 1 is what does *not* need to be specified; namely, the stellar temperature, filling factor, or position in the nebula. What *is* important is the ionization state of the preshocked material, which is indirectly specified by the above parameters.

We note from Table 1 the following results.

1. The predictions of the simple theory are fairly accurate quantitatively. The lines of dominant ions indeed increase as M_0^4, approximately, while stages lower than the dominant have larger exponents (but about 5, instead of 6 as predicted). Stages higher than the dominant have exponents smaller than 4; about 2.2 for O^{++} instead of 2 as predicted. This allows one to give a good estimate for the behavior of ions not explicitly discussed here, if one can decide how the stage of ionization of the lines lies with respect to the dominant stage.

2. H lines have $q_H = 3.7$ for a wide variety of physical conditions, instead of the 4.0 predicted by equation (6b) (with $t = 0$). The reason is that the equilibrium temperature is higher in the shocked region, inhibiting recombinations. Since hydrogen line emissivities have almost the same temperature dependence, the Balmer and Paschen lines have practically

TABLE 1

THE EXPONENT q, WHERE $R_L = M_0^q$

LABEL OF REGION	IONIZATION OF REGION					q_L			
	$x(O^{++})$	$x(O^+) \sim x(N^+)$	$x(S^+)$	$\log N_0$ (cm^{-3})	q_H	[O III] λ5007	[O II] λ3727	[N II] λ6583	[S II] λ6725
A	~1	≪1	≪1	≤4	3.7	4.0	*	*	*
B	~1	<1	≪1	≤3	3.7	3.7	5.4	5.5	*
C	<1	~1	≪1	≤3	3.7	3.4	5.0	5.4	*
				4	3.4	2.8	3.4	4.5	*
D	<1	~1	<1	≤3	3.8	2.2	3.9	4.2	5.5

* Line strengths so weak that intensity is probably unobservable [$I(3727) < 0.05$ Hβ; $I(6583) < 0.01$ Hα; $I(6725) < 0.001$ Hα].

the same values of q. By contrast, the values of q_L for other lines *do* vary appreciably. In fact, different models which have similar ionization structure give different values of q_L. For case A in Table 1, q_L for [O III] was always within 0.1 of 4. For cases B, C, and D, variations among the various models we computed for a given instance were much larger. The simple theory is expected to break down in the situation where more than one stage of ionization is appreciable, as must be the case if both [O III] and [O II] lines have measurable strengths. For q_L near 3, most models gave q_L within ± 0.3 of the quoted value; for q_L near 4, about ± 0.3, and for $q_L \geqslant 5$, about ± 0.4. The extreme values for the [S II] lines, for instance, were 5.0 and 6.1. Hence the values for q_L in Table 1 are only semiquantitative.

3. Under a wide range of conditions, we find q([N II] λ6583) ≈ q([O II] λ3727), implying that the ratio $I(3727)/I(6583)$ does not change appreciably across the shock. The physical reason for this behavior is the similarity of the ionization potentials. For our assumed abundances, this ratio was about 5 in the O^{++} zone and 3 in the O^+ zone, to within about 20%. Thus λ6583 can be used in lieu of λ3727 in determining the behavior of O^+ across a shock, which is important because λ3727 is frequently not measured in the same spectrographic range as Hβ, λ5007, Hα, and λ6583. The $I(3737)/I(6584)$ ratio changes across the shock only when the density is high enough (more than 10^3 cm^{-3}) to severely quench the λ3727.

4. Table 1 shows that the predictions of the discussion in § III*a*, above, are basically correct as regards λ5007. Consider cases A and B, when O^{++} is the dominant stage of ionization. Then q(λ5007) is about the same as for H lines. Hence the Hβ/5007 ratio should be constant across the shock. For cases C and D, q(λ5007) is less than q(Hβ) if oxygen is not mainly O^{++}, as in cases A and B. Hence j(λ5007)/j(Hβ) can only be constant or decrease across a shock. If there is an observed brightness discontinuity in which λ5007 increases more than Hβ, we must be concerned with a discontinuity in path length, with the longer path having stronger λ5007 emission. A discontinuity in reddening can hardly be an explanation for lines of such similar wavelengths. In some, but by no means all, models with O^{++} dominant, [N II] λ6583 and λ6548 are emitted with strengths of more than 0.1 j(Hβ). They have $q \approx 5.2 \pm 0.3$ instead of 6.0 as predicted by the simple theory above. Hence a shock in material which initially has O^{++} dominant [i.e., $I(\lambda 5007)/I(H\beta) \approx$ constant across the shock] will have $j(\lambda 6583)/j(H\alpha)$ and $j(\lambda 3727)/j(H\beta)$ *increasing* across the shock. If these observed intensity ratios do not increase, the λ6583 and λ3727 are contributed by projected unshocked gas, and there is merely a path length difference which causes the intensity discontinuity.

For any region with O^{++} initially dominant, S^+ is very low in abundance and [S II] λ6725 is negligible ($< 10^{-3}$ Hα). Hence if [S II] increases across an intensity discontinuity which has constant (λ5007/Hβ), we are involved with a path length discontinuity rather than a shock: a shock with constant λ5007/Hβ has oxygen mostly O^{++} and therefore practically no [S II].

5. Table 1 shows that if the preshocked material has oxygen mostly as O^+ (i.e., cases C and D, in which the shock is near the edge of the H^+ zone), then $q(\lambda 5007) \approx 2.5$. Hence λ5007 increases, but not as much as Hβ, and λ5007/Hβ decreases across the brightness discontinuity. A discontinuity in reddening, in contrast, would have a constant λ5007/Hβ ratio.

Table 1 also shows that $q(\lambda 6583) = q(H\alpha)$ across the shock, and $q(\lambda 3727) = q(\lambda 6583)$ for postshock densities of at most 10^3 cm^{-3}. Therefore λ6583/Hα remains constant across the shock. Similarly, λ3727/Hβ remains constant for low densities and decreases for high densities. It also turns out that S^{++}, and not S^+, is the dominant stage of ionization of sulfur, so the [S II] λ6726 lines are appreciable in strength (~0.1 Hβ), but have q[S II] λ6726 lines are appreciable in strength but have q[S II] = 5.5. Hence λ6725/Hα will increase as $M_0^{5.5-3.7} = M_0^{1.8}$, an appreciable amount. If there is no increase observed in [S II]/Hα, one is observing [S II] emitted by unshocked material, and so any quantitative attempts to find M_0 by comparing line ratios are likely to be very uncertain.

6. Combining results (3) and (4) above, we see that there is no shock for which the λ5007/λ6583 ratio increases in the shocked material. Different path lengths in the nebula can produce an increase in this ratio.

Table 2 summarizes the qualitative behavior of the various line ratios across a shock. If the geometry is fairly favorable, all line strengths should increase together. Any line which is continuous is surely contributed by unshocked material. If the observations

TABLE 2

BEHAVIOR OF VARIOUS LINE RATIOS ACROSS A SHOCK

TYPE OF SHOCK*	[O III] λ5007 / Hβ	[O II] λ3727 / Hβ	[N II] λ6583 / Hα	[S II] λ6725 / Hα
A.....	Constant	Negligible	Negligible	Negligible
B.....	Constant	Increases	Increases	Negligible
C.....	Decreases slightly	Increases	Increases	Negligible
D.....	Decreases	Constant	Constant	Increases

* Same notation as in Table 1. "Negligible" means not strong enough to observe.

cannot be fitted into one of the columns of Table 2, one is concerned not with a shock, but rather with a discontinuity in either extinction or in path length.

A discontinuity in extinction can most easily be detected by the change in the Balmer (or some similar) decrement across the discontinuity. The discontinuity in path length remains as the most general possibility, in which line ratios can vary quite arbitrarily.

c) *How Strong Is the Shock?*

If we observe a brightness discontinuity which meets the above criteria for a shock seen edge-on, we would like to determine the Mach number of the shock from the observed increase of the brightnesses of the various emission lines. We assume that there is a reasonably large value of R_H, the enhancement in the hydrogen lines across the shock; say, a factor of 4 or so. A major complication is that it is likely that some of the H radiation is contributed by uniform, unshocked material seen projected upon the shock. In this section, we investigate quantitatively how accurately one can estimate M_0 from observations of line intensity ratios.

We consider a simplistic analysis of the effects of unshocked material seen projected on the shocked. We suppose that we view the shock edge-on, so that we see a discontinuity in brightness of a certain line. We further suppose that only a fraction x of the emission measure of the brighter side is contributed by shocked material, with a fraction $(1 - x)$ contributed by unshocked material either in front of or behind the shocked material. We neglect internal absorption within the shocked material, so the position of the projected unshocked material is not important.

The analysis is straightforward. Let M_0 be the true Mach number of the shock. The emissivity in the shocked material is, as above, $M_0{}^q$. Let α be the fraction of the path length in the bright side which is filled by shocked material. Let R_{obs} be the observed ratio of intensities across the discontinuity. We then find, after writing R_{obs} in terms of the emissivities in the shocked and unshocked material, that

$$\alpha = x(x - xM_0{}^q + M_0{}^q)^{-1}, \tag{10}$$

$$R_{obs} = 1 - \alpha + M_0{}^q\alpha \tag{11a}$$

$$= M_0{}^q(x - xM_0{}^q + M_0{}^q)^{-1}. \tag{11b}$$

For the simplest estimate of the Mach number, we consider a hydrogen line, because it has the best determined value of q (=3.7; see Table 1). We then assume that there is only shocked material seen on the brighter side, so that we erroneously set $x = \alpha = 1$. Equation (11b) then gives $M = R_H{}^{1/q}$. Let this value of M be M_H.

It is also possible to use lines beside hydrogen to estimate M. However, when a stage of ionization is dominant, so that the lines are strong, the exponent is about 4, or close to that of hydrogen. Hence one gains little additional information by considering that stage of ionization. To have information different from the H lines, one must consider lines from minor stages of ionization, such as O^+ or N^+ for an O^{++} zone, or S^+ for an O^+, N^+ zone. These minor stages of ionization can have rather strong lines (even stronger than Hβ in some circumstances), but have $q \approx 5$–6.

Consider a line from a minor species, say, [S II] λ6725. The simplest procedure for estimating M from [S II] lines is to proceed as with H: assume that all of the material on the bright side of the discontinuity has been shocked, so that we calculate M_L, our estimate for M_0, by assuming the q for the line (i.e., $q = 5.5$ for [S II]), and then taking the qth root of the brightness ratio, R_L.

Table 3 gives values of M_H and M_L for shocks of $M_0 = 1.5$ and 2.5, for various values of x, the fraction of the H emission on the bright side contributed by

TABLE 3

ESTIMATES OF MACH NUMBER

x	$M_0 = 1.5$		$M_0 = 2.5$	
	M_H	M_L	M_H	M_L
1.0.....	1.50	1.50	2.50	2.50
0.95.....	1.44	1.45	1.97	2.11
0.9.....	1.38	1.41	1.73	1.93
0.7.....	1.24	1.28	1.36	1.58
0.5.....	1.14	1.18	1.20	1.39

NOTES.—The quantity x = fraction of intensity of Hβ of bright material contributed by shocked material; M_0 = true Mach number; $M_H = R_H{}^{1/3.7}$; R_H from eq. (11b) with $q = 3.7$; $M_L = R_L{}^{1/5.5}$; R_L from eq. (11b) with $q = 5.5$.

the shock. The value of q_L was taken to be 5.5. We see from Table 3 that both M_H and M_L fall rapidly as x declines from unity, especially for $M_0 = 2.5$. Thus one cannot be very accurate in estimating M_0 from a single line ratio, especially for a strong shock. However, such estimates do provide a lower limit for M_0.

Note that Table 3 was computed assuming that the shock really had exactly the value of $q_L = 5.5$ assumed in obtaining M_L. Hence the only error involved in the table is the assumption that the bright material is completely shocked. However, the extreme values of q we found for models with strong [S II] were 5.0 and 6.1. The 10% uncertainty in the true value of q introduces about a 10% error in the estimate of M, in addition to the error arising from the projected unshocked material. Thus the estimate of M from the hydrogen lines is at least as accurate as that from lines of nondominant stages of ionization, in spite of having a slightly worse estimate for a given value in x in Table 3.

Table 3 also shows that, for a given value of x, $M_H \approx M_L$, so that the two estimates, both quite wrong, give consistent results. The agreement of M_H and M_L does not mean that either is correct!

One attractive possibility is to use R_H and R_L together, assuming R_L refers to a line with $q > 4$, to correct quantitatively for unshocked material superposed upon the shocked. Assuming that the shocked path length α is the same for H and the line, one eliminates α from the two equations (11) for hydrogen and the line. One obtains a simple transcendental equation for the estimated Mach number, M_{HL}, in terms of R_H, R_L, q_H, and q_L. Unfortunately, this procedure turns out to be impractical. If one knew the values of q_L, R_H, and R_L exactly, then indeed $M_{HL} = M_0$. Even with an exact value of q_L, photometric errors could still be a problem; it turns out that, roughly, $M_{HL} \propto (R_L/R_H)$, so that reasonable errors (20%, say) in the line ratios can make M_{HL} rather uncertain. A much more serious problem, though, is the uncertainty in q_L. For instance, if one assumes $q_L = 5.5$ when the shock actually produces the line with $q_L = 6$, then M_{HL} is 1.16 for $M_0 = 2.5$, a worse estimate of M_0 than either M_H or M_L. Also, M_{HL} is not a lower limit, but can be either larger or smaller than M_0.

Our recommendation for estimating M_0 is simple: choose a region in the nebula where the brightness contrast in the H lines, R_H, is large, test other line ratios to see if the brightness change is likely to be caused by a shock seen edge-on, and then take $R_H^{1/3.7}$ as a lower limit to the Mach number of the shock. Any projected unshocked material will cause the estimate to be lower than the true value, but using a line ratio other than H does not increase the accuracy. Unfortunately, it is necessary to observe the brightness of the lines in the faint (unshocked) material, for only by comparison of brightness ratios across a discontinuity can a shock be separated from other causes of brightness changes.

We note that this paper has been concerned entirely with predictions rather than observations. The reason is that we are unaware of the kinds of observations (brightness of several line ratios, on both the bright and faint sides of intensity discontinuities) which are needed. However, we see no reason why such observations are impractical, and hope that observers will provide them in the near future.

This research has been partially supported by the US National Science Foundation under grant AST 76-15384, for which we are grateful. P. W. J. L. B. would like to acknowledge partial support by the Science Research Council of Great Britain, and would like to thank Washburn Observatory, where this work was carried out, for its magnificent hospitality and stimulating atmosphere. We are grateful to Drs. B. D. Savage and J. P. Cassinelli for their comments on the manuscript.

REFERENCES

Axford, W. I. 1961, *Phil. Trans. Roy. Soc. London, A*, **253**, 301.
Balick, B., and Sneden, C. 1976, *Ap. J.*, **208**, 336.
Dopita, M. 1976, *Ap. J.*, **209**, 395.
Dyson, J. E. 1968, *Ap. Space Sci.*, **2**, 461.
Elliott, K. H., and Meaburn, J. 1975*a*, *Ap. Space Sci.*, **35**, 81.
———. 1975*b*, *M.N.R.A.S.*, **170**, 237.
———. 1975*c*, *M.N.R.A.S.*, **172**, 427.
Harris, D. 1975, Ph.D. thesis, University of Arizona.
Hummer, D., and Mihalas, D. 1970, JILA Rept., No. 101.
Jura, M. 1976, *Ap. J.*, **206**, 691.
Kurucz, R. L., Peytremann, E., and Avrett, E. H. 1974, *Blanketed Model Atmospheres for Early Type Stars* (Washington: Smithsonian Institution).
Landau, L. D., and Lifshitz, E. M. 1959, *Fluid Mechanics* (Oxford: Pergamon).
Marsh, M. C. 1970, *M.N.R.A.S.*, **147**, 95.
Mathis, J. S. 1978, in preparation.
Osterbrock, D. E. 1974, *Astrophysics of Gaseous Nebulae* (San Francisco: Freeman).
Ralston, A., and Wilf, H. S. 1960, *Mathematical Methods for Digital Computers* (New York: Wiley).
Shull, J. M., York, D. G., and Hobbs, L. M. 1977, *Ap. J. (Letters)*, **211**, L139.
Spitzer, L., Jr., and Jenkins, E. B. 1975, *Ann. Rev. Astr. Ap.*, **13**, 133.
Wilson, O. C., Münch, A., Flather, E., and Coffeen, M. 1959, *Ap. J. Suppl.*, **4**, 199.

P. W. J. L. BRAND: University of Edinburgh, Department of Astronomy, Royal Observatory, Edinburgh EH9 3HJ, Scotland

J. S. MATHIS: Washburn Observatory, University of Wisconsin–Madison, 475 North Charter Street, Madison, WI 53706

Astron. Astrophys. 71, 47—50 (1979)

ASTRONOMY AND ASTROPHYSICS

Scattering from a Cylindrical Dust Cloud in an Isotropic Radiation Field

P. W. J. L. Brand

Department of Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, U.K.

Received March 28, 1978

Summary. An approximation to the source function in a uniform cylindrical dust cloud embedded in an isotropic incident radiation field has been used to calculate the scattered and transmitted intensities observed from the cloud. It is shown that for reasonable parameters for the dust (albedo 0.6 and asymmetry factor 0.7), the effect of scattering is to change by a small amount the dust opacity predicted on the assumption that all the observed light is transmitted or foreground emission. This result confirms the magnitudes of dust opacity calculated in a previous paper.

Key words: interstellar clouds — interstellar dust — scattering

1. Introduction

In a recent paper (Brand and Zealey, 1978, referred to as Paper I) the analysis of the edge profile of a dark cloud in front of an H II region was carried out. In that paper the assumption was made that scattering of nebular light into the line of sight was not significant. The following discussion will show that even if the cloud were completely surrounded (isotropically) by the emission region, scattering into the line of sight will not alter significantly the conclusions in Paper I regarding dust opacity in the cloud.

Following Paper I, the assumptions to be made are that (a) the cloud is a cylinder uniformly filled with dust, (b) the incident radiation field is isotropic, and uniform, and (c) the dust has an albedo α and g ($=\langle\cos\theta\rangle$ where θ is the angle through which the radiation is scattered) comparable with typical values in the interstellar medium.

2. Radiative Transfer

This discussion will follow the results of Heaslet and Warming (HW 1966) on radiative transfer in cylindrical geometry.

Send offprint requests to: P. W. J. L. Brand

The equation of radiative transfer in cylindrical coordinates is

$$\sin\psi\left[\mu\frac{\partial I}{\partial\rho}+\frac{1-\mu^2}{\rho}\frac{\partial I}{\partial\mu}\right]=-I+J(\rho). \tag{1}$$

Here $\rho=\kappa r$ (κ being the extinction coefficient), $\cos\mu^{-1}$ is the angle between the radial direction and the projection on the plane normal to the cylinder axis of the ray on which I is specified, and ψ is the angle between the ray and the direction of the axis. The source function $J(\rho)$ is that of the scattered radiation (and here assume isotropic scattering with albedo ω but see below) so

$$J=\omega\bar{I}=\omega\int I d\Omega/4\pi. \tag{2}$$

The flux H in the radial direction is given by

$$H=\int I\mu\sin\psi d\Omega/4\pi \tag{3}$$

and

$$K=\int I\mu^2\sin\psi d\Omega/4\pi, \tag{4}$$

with $d\Omega=\sin\psi d\mu d\phi$ where ϕ is the azimuthal angle round the radial direction. Hence,

$$\frac{1}{\rho}\frac{d}{d\rho}(\rho H)=-(1-\omega)\bar{I} \tag{5}$$

and

$$\frac{dK}{d\rho}+\frac{1}{\rho}\left(2K-\int_0^{2\pi}d\phi\int I\sin^3\psi d\psi/4\pi\right)=-H. \tag{6}$$

Application of the Eddington approximation then gives

$$H=-\bar{I}/3 \tag{7}$$

or

$$\nabla^2\bar{I}=3(1-\omega)\bar{I}. \tag{8}$$

As pointed out by HW, this is not good enough in the non-conservative case ($\omega\neq 1$). These authors, by doing numerical calculations, and noting the results of diffusion

and Eddington approximations, suggested that a best solution is got by solving

$$\nabla^2 \bar{I}(\rho) = \gamma^2 \bar{I}(\rho) \tag{9}$$

with

$$\gamma^2 = 3(1 - \omega) + (1 - \omega)^2. \tag{10}$$

In the cylindrical case the non-singular solution is proportional to the modified Bessel function of order zero,

$$\bar{I} = CI_0(\gamma\rho). \tag{11}$$

The constant C may be determined approximately by boundary conditions at the surface of the cloud, by taking (a) plane parallel and (b) 2 discrete ordinate approximations to get

$$\pm 3^{-1/2} \frac{dI^{\pm}}{d\rho} = -I^{\pm} + J. \tag{12}$$

At the surface I^- is I_{iso}, the isotropic incident specific intensity, and $\bar{I} = (I^+ + I^-)/2$. Subtracting the pair of Equations (12) gives

$$3^{-1/2} \frac{d\bar{I}}{d\rho} = I_{\text{iso}} - \bar{I}. \tag{13}$$

Since $\bar{I}$ at the surface is $CI_0(\gamma\rho_0)$ by Equation (11), letting $\rho = \rho_0 = \kappa r_0$ at the surface,

$$\bar{I} = I_{\text{iso}} I_0(\gamma\rho)/(I_0(\gamma\rho_0) + 3^{-1/2}\gamma I_1(\gamma\rho_0)), \tag{14}$$

where I_0 and I_1 are modified Bessel functions, and the relation $I_0 = I_1$ has been used. Thus we have the source function $J = \omega\bar{I}$ in the cylindrical cloud, and can use it to solve the radiative transfer equation along an arbitrary line of sight. Consider the line of sight a distance r_m from the axis of the cylinder and with it making an angle $(90 - \eta)°$. The source function J is constant on ellipses similar to the intersection of the cylinder by a plane containing both the line of sight and the perpendicular mutual to line of sight and axis. Such a plane is shown in Figure 1. The formal solution of Equation (1) becomes, using the geometry of Figure 1,

$$I_{\text{obs}} = I_{\text{iso}} \exp(-2\kappa L \sec\eta) + \int_{-L\sec\eta}^{+L\sec\eta} J(r_m^2 + l^2\cos^2\eta) \times \exp(\kappa l - \kappa L \sec\eta)\kappa dl \tag{15a}$$

and ratios r_{trans}, r_{scatt} and r_{rot} can be defined by

$$I_{\text{obs}} = I_{\text{iso}} r_{\text{trans}} + I_{\text{iso}} r_{\text{scatt}} = I_{\text{iso}} r_{\text{tot}}. \tag{15b}$$

The second term on the right hand side of (15a), which is the radiation scattered into the beam I_{scatt}, can be written as

$$I_{\text{scatt}} = \exp(-\tau_L \sec\eta) \int_0^{\tau_L \sec\eta} J(\rho_m^2 + \tau^2\cos^2\eta) \times (\exp(\tau) + \exp(-\tau))d\tau, \tag{16}$$

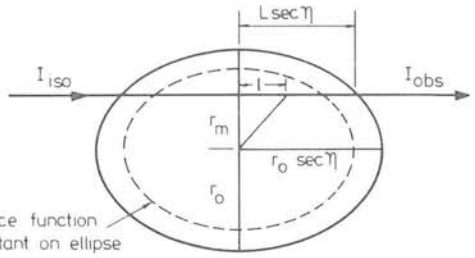


Fig. 1. Section through a cylindrical dust cloud (radius r_0) containing the line of sight, and the perpendicular mutual to line of sight and to axis. The plane is at an angle $(90 - \eta)°$ to the axis. Each cylinder of constant radius (on which the source function is constant) is an ellipse of eccentricity $\sin\eta$

where $\tau = \kappa l$, $\rho = \kappa r$ and $\tau_L = \kappa L$. The length of the line of sight in the cloud is $2L \sec\eta$, and $L^2 = r_0^2 - r_m^2$. The argument of J is $\rho_0^2 - \tau_L^2 + \tau^2\cos^2\eta$. By changing the integration variable from τ to $\tau\cos\eta$ and replacing J by its value given by Equations (2) and (14)

$$r_{\text{scatt}} = \omega \sec\eta \exp(-\tau_L \sec\eta) \times [I_0(\gamma\rho_0) + 3^{-1/2}\gamma I_1(\gamma\rho_0)]^{-1} \times \int_0^{\tau_L} I_0(\gamma(\rho_m^2 + \tau^2)^{1/2}(\exp(\tau \sec\eta) + \exp(-\tau \sec\eta))d\tau. \tag{17}$$

3. Albedo and Phase Function of the Dust Grains

Albedo α and asymmetry factor $g = \langle\cos\theta\rangle$ have been derived recently for the interstellar medium generally (Lillie and Witt, 1976), and for a dark cloud (FitzGerald et al., 1976). Using these results we shall take $\alpha = 0.6$, $g = 0.7$ in the present work. The grains are considerably forward throwing. If a Henyey-Greenstein phase function (Henyey and Greenstein, 1941) is adopted as a reasonable approximation to the data for dielectric grains, then half of all scatterings are within 28° of the incident direction. In view of this it is proposed to represent the phase function by the sum of an isotropic part, and a delta function representing scattering forward. Since the boundary condition is that the specific intensity is isotropic at all points exterior to the dust cloud, small angle scattering into, and out of, the beam should balance to first order. That is, at this level of approximation the scattered radiation entering a given line of sight has either traversed a path very similar to the unscattered radiation (the delta function part) or has been scattered through a large angle (the isotropic part). Thus the intensity along a line of sight can be considered to be due to attenuation of the incident radiation by absorption, and by that fraction of scatterings regarded as isotropic, plus the contribution of radiation scattered isotropically one or more times from any direction into the line of sight. The transfer of this radiation is controlled by an effective opacity due

to absorption and isotropic scattering, as for the direct radiation, and an effective albedo ω can be defined by noting the following. Let the extinction cross section of the dust be $\sigma_e = \sigma_a + \sigma_i + \sigma_f$, the sum of absorption, isotropic scattering, and forward scattering. The ratio of isotropic to forward scattering cross sections is found by noting that if the forward scattering part of the phase function is a delta function, and the rest of the scattering is isotropic, then the asymmetry factor is

$$g = \sigma_f/(\sigma_i + \sigma_f). \tag{18}$$

The albedo is defined

$$\alpha = (\sigma_i + \sigma_f)/(\sigma_i + \sigma_f + \sigma_a) \tag{19}$$

and we define the isotropic albedo as

$$\omega = \sigma_i/(\sigma_i + \sigma_a). \tag{20}$$

Hence

$$\omega = (1 - g)/(\alpha^{-1} - g) \tag{21}$$

measures the effective (i.e. isotropic) scattering relative to the effective (isotropically scattered or absorbed) extinction. If $\alpha = 0.6$, $g = 0.7$ then $\omega = 0.31$. The opacity of Section 2 is given by $n_d(\sigma_i + \sigma_a)$, where n_d is the number density of dust grains.

4. Results and Discussion

The emergent scattered intensity [Eq. (17)] was calculated by numerical quadrature, using a value of $\gamma = 1.60$, corresponding to $\omega = 0.31$. Figure 2 shows the quantity r_{scatt} defined by Equation (15b) as a function of x, the distance from the edge of the cloud in units of cloud radius. By referring to Figure 1, this distance is given in terms of τ_L by

$$x = 1 - [1 - (\tau_L/\rho_0)^2]^{1/2}. \tag{22}$$

The aspect angle $\eta = 0°$ in Figure 2. At low optical radius ρ_0 the amount of scattered light peaks at the center. As ρ_0 increases, the peak scattering occurs near x corresponding to $\tau_L = 1$, that is where the line of sight intersects unit optical depth of cloud. The peak value of r_{scatt} remains nearly constant at 0.13, and as ρ_0 becomes greater than 1, the scattering becomes more like that from a semi-infinite plane parallel atmosphere. Chandrasekhar (1950, Ch. 3 and 5) has dealt with this case, and from his results one may immediately deduce that for isotropic incidence,

$$r_{\text{scatt}}(\text{plane}) = 1 - (1 - \omega)^{1/2} H(\tau_L/\rho_0), \tag{23}$$

where $H(\mu)$ is the H-function tabulated in Table XI of Chandrasekhar (1950). This scattered fraction is also shown in Figure 2.

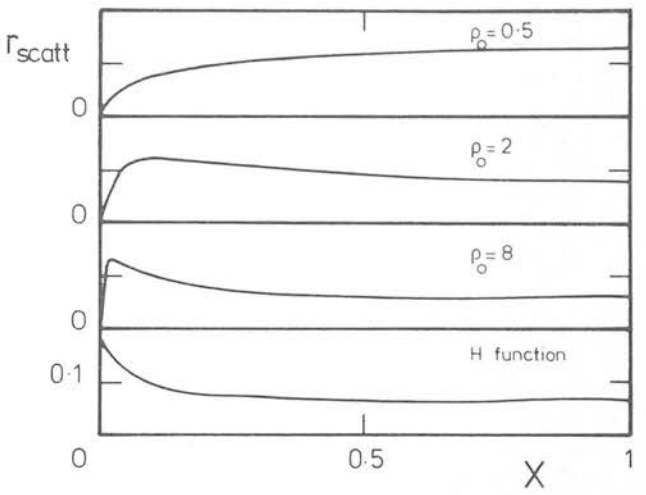


Fig. 2. Ratio r_{scatt} of scattered intensity to incident intensity from a cylindrical dust cloud as a function of fractional radius x, for various optical radii ρ_0. The peak value occurs at line of sight optical depth unity or the largest value if this is less. The lowest curve is r_{scatt} for an infinite plane parallel dust atmosphere calculated using Chandrasekhar's H function

What is observed, however, is not r_{scatt} but r_{tot} [Eq. (15b)] plus any foreground emission there may be, and what must be compared with observation therefore is the difference between off-cloud and on-cloud intensities, normalised by the maximum value of this difference. So in Figure 3 what is plotted is

$$\mathscr{I}(x) = [r_{\text{tot}}(x) - r_{\text{tot}}(1)]/[1 - r_{\text{tot}}(1)], \tag{24}$$

where $r_{\text{tot}}(x)$ and $r_{\text{tot}}(1)$ are given by Equations (15) and (16) with aspect angle $\eta = 0°$. Each curve is labelled with its value of ρ_0. Also indicated in Figure 3 by dashed lines labelled with ρ_0 are the corresponding curves with r_{scatt} put equal to zero, that is

$$\mathscr{I}_{\text{trans}}(x) = [\exp(-2\tau_L) - \exp(-2\rho_0)]/[1 - \exp(-2\rho_0)]. \tag{25}$$

These curves are seen to have similar shapes up to $\rho_0 = 2.5$ for $\mathscr{I}_{\text{trans}}$, and to deviate significantly in shape beyond an $\mathscr{I}_{\text{trans}}$ value of $\rho_0 = 3$. At such values, however, it is unlikely that the method used in Paper I would be successful even if scattering could be neglected.

At the lowest values of ρ_0 scattering has an insignificant effect on the profile. The value of ρ_0 that would be predicted by fitting a no-scattering model to observations of a cloud with the properties described in this paper is $0.8\rho_0$ at values of ρ_0 below 3, and decreases to $0.65\rho_0$ at $\rho_0 = 5$. However, shape differences at these high values of ρ_0 are significant as already noted. In the usable range of ρ_0, therefore, the effect of neglecting scattering in the model fit is to *underestimate* ρ_0 by a factor 0.8.

The effect of varying aspect angle can now be considered. In the case of 'transmission only' profiles, the parameter deduced by fitting the model to observations

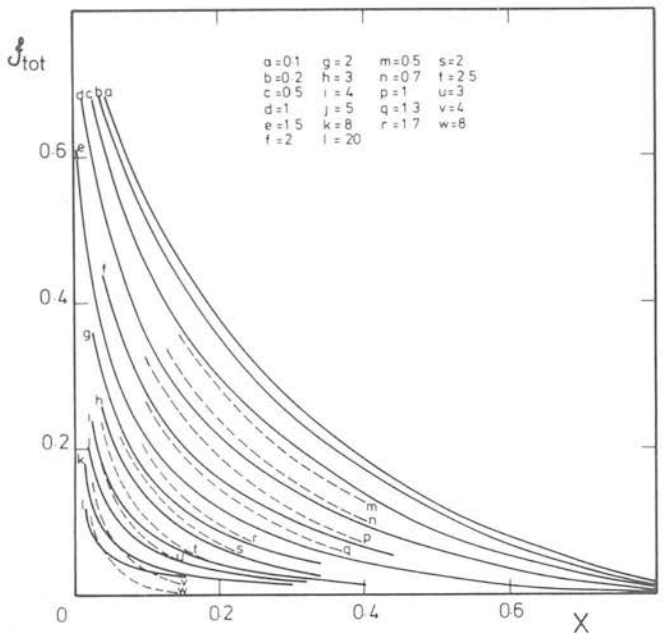


Fig. 3. Normalised intensity profiles $\mathcal{J}(x)$ for various optical radii ρ_0. Also indicated by dashed lines are the profiles from such clouds with no scattering

is $\rho_0 \sec \eta$ because optical depth of any line of sight scales with $\sec \eta$. However, this scaling is lost in the non-linearities of multiple scattering, and Equation (17) has to be re-evaluated for differing values of η. This was done for $\sec \eta = 1.5$ and $\sec \eta = 2.0$, and the results were compared with the $\eta = 0°$ case. The resulting profiles are very similar in shape at all values of ρ_0 to each other and to the $\eta = 0°$ profiles, although of course coincident curves have different ρ_0 values. The entire range of results for $0.2 \leqq \rho_0 \leqq 8$ and $\sec \eta = 1$, 1, 5 and 2.0 could be encompassed in the statement that the "effective ρ_0" estimated by fitting observations of such clouds by a "no-scattering" model is $\rho_0(\sec \eta)^{1.2}$. That is, the effect of scattering is to enhance slightly the effect of obliquity in increasing the apparent optical radius of the cloud.

The conclusion to be drawn is that in the circumstances envisaged here, the error due to the neglect of scattering is of the order of 20% in estimates of dust density, and that the effect of unknown obliquity is an error of at least comparable and probably greater magnitude. In the case that the background is not isotropic (e.g. a cloud in front of a diffuse source) a more general analysis is required but it may be seen intuitively that since the profile of the scattered light is widely different from the transmitted profile and these profiles are additive, a lesser amount of scattered light will have an effect intermediate between no scattering and isotropic scattering. With this last proviso, the results of Paper I, which do not claim accuracy better than 20%, still stand.

Acknowledgements. I am grateful to Drs. D. Morgan, J. Cassinelli and G. Rybicki for useful comments. The friendly hospitality of Washburn Observatory and the Center for Astrophysics, and financial support provided by Science Research Council all helped in the completion of this work.

References

Brand, P. W. J. L., Zealey, W. J.: 1978, *Astron. Astrophys.* **63**, 345
Chandrasekhar, S.: 1950, Radiative Transfer, Clarendon Press, Oxford
FitzGerald, M. P., Stephens, T. C., Witt, A. N.: 1976, *Astrophys. J.* **208**, 709
Heaslet, M. A., Warming, R. F.: 1966, *J. Quant. Spectrosc. Radiat. Transfer* **6**, 751
Henyey, L. G., Greenstein, J. L.: *Astrophys. J.* **93**, 70
Lillie, C. F., Witt, A. N.: 1976, *Astrophys. J.* **208**, 64

Mon. Not. R. astr. Soc. (1981) 197, 217–233

Ionization front interactions and the formation of globules

Peter W. J. L. Brand *Department of Astronomy, University of Edinburgh, Blackford Hill, Edinburgh EH9 3HJ, Scotland*

Received 1981 February 9; in original form 1979 December 10

Summary. It is assumed that an H II region has evolved inside a molecular cloud. The interactions that result from the expanding shell of compressed molecular gas reaching the edge of the cloud are calculated, and the instability of the ionization front to the formation of globules is investigated. The rarefaction wave which is reflected from the contact discontinuity as the leading shock passes through the edge of the cloud, accelerates the ionization front, and since conditions at the front satisfy Capriotti's criterion for instability, the shell breaks up.

The size of the fragment so created is determined by the thickness of the compressed shell. If the shell phase of H II region evolution has proceeded significantly, then globules of up to a fraction of a solar mass may be formed in an H II region caused by a star with an ionizing luminosity of 10^{49} photon s^{-1} in a molecular cloud of density 10^3 cm^{-3}. These globules may survive the ionizing flux from the star, and will be driven from the cloud by the rocket effect.

1 Introduction

The presence of globules in H II regions has been known of since the studies of Bok & Reilly (1947). It is usually assumed that these globules are the result of density inhomogeneities present in the cloud before the creation of the H II region, but recent models of H II regions involving breakout from a molecular cloud have suggested another possibility. Due to the reduction in ram pressure outside a molecular cloud, the compressed neutral shell created during the evolution of the H II region is accelerated, and it appears that a Rayleigh–Taylor instability may disrupt the shell, creating globules.

In consequence I have investigated the operation, during the evolution of an H II region, of this process which will create small globules. However, the process will not create the largest (Bok) globules which must have existed in the region before the generation of the H II region, or which may be created in the instability considered by Giuliani (1979).

The process discussed in this paper occurs at the early 'breakout' stage of the blister (Israel 1978) or champagne (Tenorio-Tagle 1979) model of an H II region at the edge of a molecular cloud.

2 The model

What follows relies on the theory of ionization fronts as developed by Axford (1961) and by Goldsworthy (1961) who with his students and collaborators has also considered interactions between various fronts. The paper by Marsh (1970) also deals with a simple interaction of the type considered in this work.

It is assumed that an early-type star is switched on inside a molecular cloud of proton number density n_0. After quickly reaching its initial Strömgren radius, the resulting H II region begins to sweep up a shell of compressed molecular material. During this stage of evolution, the shell reaches the edge of the molecular cloud, and bursts into the intercloud medium, assumed to be atomic hydrogen at a much higher temperature then the molecular material, and initially in pressure equilibrium with it. During the interaction calculations isothermality is assumed, in the sense that all molecular material is at the temperature $T_{\rm mol}$, all atomic material is at temperature $T_{\rm a}$, and the ionized gas is at temperature $T_{\rm i}$. The corresponding isothermal sound speeds are $c_{\rm mol}$, $c_{\rm a}$ and $c_{\rm i}$. This assumption is discussed later.

In the entire shell-growing phase the ionisation front evolves from critical D-type to very weak D-type. With no interactions occurring, it is possible to write an approximation for the trajectory of the ionization front (Spitzer 1978, p. 254):

$$R_{\rm i} = R_{\rm i0}\, f^{4/7}, \tag{1a}$$

$$U_{\rm i} = c_{\rm i}\, f^{-3/7}, \tag{1b}$$

$$f = 1 + t/t_0. \tag{1c}$$

Here $R_{\rm i}$ is the distance of the front from the exciting star, $R_{\rm i0}$ is the initial Strömgren radius, and

$$t_0 = (4/7) R_{\rm i0}/c_{\rm i} \tag{2a}$$

is the time-scale for the evolution of the shell phase. If $T_{\rm i} = 8000$ K and the ionizing luminosity from the central star is $10^{48}\, S_{48}$ photon $\rm s^{-1}$, then

$$t_0 = 1.65\ \ 10^4 S_{48}^{1/3}\, n_0^{-2/3}\ \text{yr}. \tag{2b}$$

The assumptions built into these equations are (a) that the leading shock is strong, (b) that the compressed shell is narrow so that the shock velocity $U_{\rm s}$ approximately equals the front velocity $U_{\rm i}$, and (c) the interior density is uniform.

Assumption (a) is true until the very late stages, (b) follows from isothermality, and (c) has the following consequence. If $n_{\rm i}$ is the (uniform) number density of protons in the H II region, then ionization equilibrium requires

$$n_{\rm i}^2 R_{\rm i}^3 = \text{constant}, \tag{3a}$$

therefore

$$n_{\rm i} = n_0 f^{-6/7}. \tag{3b}$$

If R is the coordinate of a particular particle, then mass conservation requires

$$n_{\rm i} R^3 = \text{constant}. \tag{4}$$

Equations (3a) and (4) applied to a particle just entering the H II region at the front yield

$$\dot{R} = \tfrac{1}{2}\dot{R}_{\rm i} = \tfrac{1}{2} U_{\rm i}, \tag{5}$$

and so ionized material streams out of the front at half the velocity of the shock, U_s, through the molecular material. This is required to maintain uniform density inside the H II region, and represents the true downstream boundary conditions reasonably well when $t \gtrsim t_0$, that is when pressure equilibrium can be established.

Material is swept into a shell between the leading shock and the ionization front since the mass of the enclosed H II region is less than the mass of the material overrun by the shock. With the thin shell approximation the column density, N_{shell}, is

$$N_{\text{shell}} = (1/3)R_i(n_0 - n_i)$$
$$= (1/3)R_{i0}n_0 f^{4/7}(1 - f^{-6/7}). \tag{6}$$

The shell density is less easy to determine, since it depends on internal conditions in the shell. If the shock is isothermal (that is post-shock cooling is very rapid), the density enhancement is the square of the isothermal Mach number. Isothermality is the best assumption to make in the absence of magnetic fields. However, if a magnetic field is a significant contributor to the energy density ahead of the shock, then it will resist compression and cause the gas to behave adiabatically with $\gamma = 2$ if the field is aligned parallel to the shock front, and if (as is likely so close to an OB star) there is sufficient ionization to freeze the magnetic field in the gas. In this case magnetic pressure dominates in the shell, and the shock is assumed strong so that magnetic pressure and ram pressure can be equated. Suppose β_0 is the ratio of magnetic to gas pressure ahead of the shock, and p_m the magnetic pressure behind the shock

$$p_m \simeq \rho_0 U_s^2 . \tag{7}$$

Now

$$p_m = \frac{B^2}{8\pi} = \frac{B_0^2}{8\pi}\left(\frac{B}{B_0}\right)^2 = \beta_0 \rho_0 c_{\text{mol}}^2 \left(\frac{\rho_{\text{shell}}}{\rho_0}\right)^2 . \tag{8}$$

Thus

$$\rho_{\text{shell}}/\rho_0 \simeq \beta_0^{-1/2} U_s/c_{\text{mol}} \simeq \beta_0^{-1/2} U_i/c_{\text{mol}}. \tag{9}$$

The shell compression $K = n_{\text{shell}}/n_0$ is

$$K = \beta_0^{-1/2}(c_i/c_{\text{mol}})f^{-3/7} \qquad \text{(magnetic)}, \tag{10}$$
$$(c_i/c_{\text{mol}})^2 f^{-6/7} \qquad \text{(isothermal)}.$$

Then the shell thickness s becomes

$$s = N_{\text{shell}}/n_{\text{shell}} = N_{\text{shell}}/n_0 K \tag{11a}$$

or

$$s/R_i = (1/3)(1 - f^{-6/7})/K. \tag{11b}$$

3 The interactions

The ratio of interaction time-scale (the shell crossing time at velocity c_i) to the H II region evolution time is just s/R_i which is given by equation (11b), and is always small. So to simplify the interaction calculation the downstream conditions are held constant, specified by a particular value of f. Furthermore, since the shell is thin, spherical divergence is ignored and the interactions are treated as plane parallel.

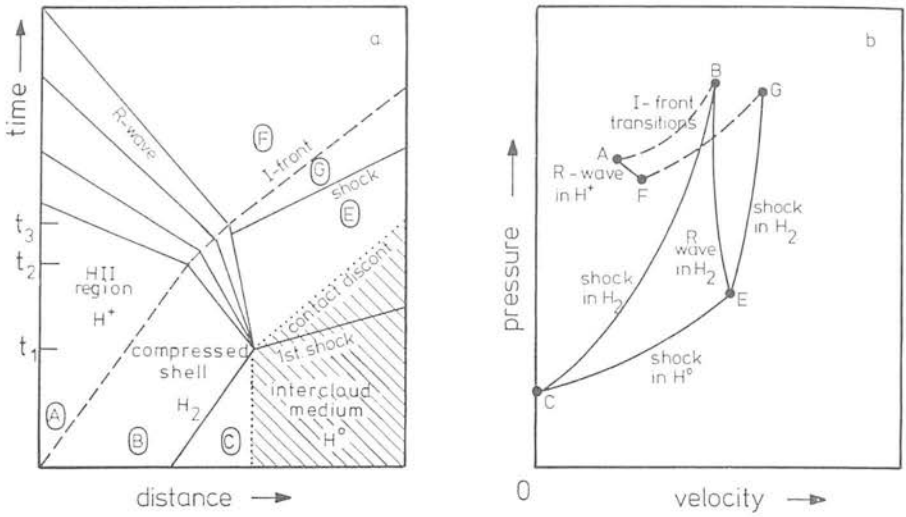


Figure 1. (a) Diagram of the trajectories of the waves that result from the initial shock reaching the contact discontinuity (.....) between molecular cloud and intercloud medium (hatched). At time t_2 the leading edge of the rarefaction wave (fan of lines) reaches the ionization front and at t_3 has passed into the H II region. A shock is reflected. (b) The transition across the waves in (a) map on to the (p, u) plane as discussed in Appendix A.

The method of dealing with interactions is described in Appendix A, and the application of this method to the present problem is described in Appendix B, the solution being shown schematically in Fig. 1. The trajectories in Fig. 1(a) show the first few interactions, which are important in determining the stability of the flow. Fig. 1(b) is the pressure–velocity diagram which relates the velocities for the various waves and fronts in Fig. 1(a).

At time t_1 the shock at the front of the molecular shell B is shown reaching the edge of the cloud represented by the contact discontinuity between molecular hydrogen and the atomic hydrogen of the intercloud medium. The reduction in ram pressure accelerates the shock and a rarefaction wave propagates back through the shell. At time t_2 the leading edge of this wave reaches the ionization front and passes through it into the H II region. At t_3 the passage of the wave is complete, and a second shock is created, and moves forward into the molecular shell.

The various phases of the interaction can be followed on each diagram in Fig. 1 by following the letters labelling the gas between the various waves. In each instance start at the interior of the H II region A and work outwards to undisturbed intercloud medium C.

$t < t_1$: *ABC*	H II, ionization front (IF), compressed shell, shock, undisturbed cloud, contact discontinuity (CD), intercloud medium (ICM).
$t_1 < t < t_2$: *ABEC*	H II, IF, compressed shell, rarefaction wave (RW), less compressed shell, CD, ICM.
$t > t_3$: *AFGEC*	H II, RW, H II, IF, compressed shell, shock, CD, compressed ICM, shock, ICM.

The trajectories calculated from initial values $U_i/c_i = 0.5$, $c_i/c_a = 10$, $c_i/c_{mol} = 10^{1.5}$, are shown in Fig. 2. This figure shows the paths of the waves in a frame comoving with the initial ionization front. The non-centred rarefaction wave causes problems (see Appendix B) since the ionization front is continuously accelerated, but this acceleration is approximated by plotting the front as a straight line with slope corresponding to the mean velocity before and after the interaction.

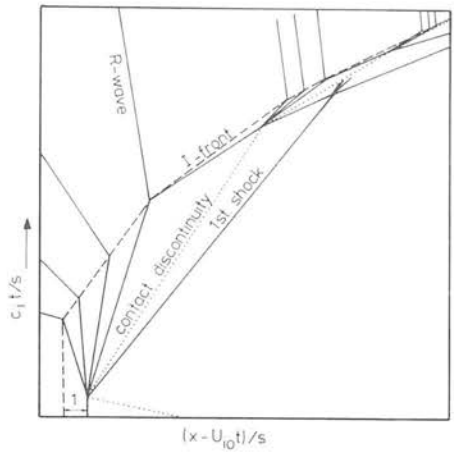


Figure 2. The trajectories for $U_i/c_i = 0.5$, $T_{mol} = 32$ K, $T_a = 160$ K. The axes are scaled with sound speed in ionized hydrogen c_i, and compressed shell thickness s. The x-axis measures distance in a frame comoving with the initial ionization front.

Most of the deflection of the ionization front trajectory occurs during the passage of the front through the first rarefaction wave. During this passage, a dilute field (ionized gas) is accelerating a dense fluid (compressed molecular gas) nearly in pressure equilibrium (within a factor 2 for weak D fronts) and so the interface, in this case the ionization front, is susceptible to Rayleigh–Taylor instability. The instability of the ionization front during the passage of the first rarefaction wave will now be examined.

4 The instability

Capriotti (1973) considered the stability of accelerating ionization fronts, and was able to show in a case where he restricted the front to be of critical D type, that the accelerating shell around a planetary nebula breaks up due to a Rayleigh–Taylor instability, provided that the velocity at which the front eats into a neutral material is less than the velocity relative to the front achieved by an acclerated lump in one recombination time t_{rec}. That is

$$t_{rec}\,\Delta u/\Delta t > 2v_1, \tag{12}$$

where v_1 is the velocity of neutral gas relative to the front, and Δt, Δu are the time of passage of the front through the rarefaction wave, and the consequent change in velocity. The recombination time is approximately $10^5/n_i$ yr. Calculation of Δu and Δt for various velocities U_i/c_i of the ionization front before interaction, and for the different intercloud media specified by T_a, give the values in Table 1. These values are now used to investigate the instability.

Equation (12) can be written

$$t_{rec}\,g > 2v_1 s/c_i^2, \tag{13}$$

where $g = s\Delta u/c_i^2\,\Delta t$ is listed in Table 1. The mass and momentum jump conditions across the D type ionization front (use $v_1 \ll c_{shell}$, where c_{shell} is the relevant speed in the shell, either sound speed or Alfvén velocity) give

$$v_1/c_i = (c_{shell}/c_i)^2\,(U_i/2c_i)/[1 + (U_i/2c_i)^2]. \tag{14}$$

Table 1. The effect of the interaction between ionization front and rarefaction wave.

T_a(K)	U_i/c_i	$\Delta u/c_i$	$c_i\Delta t/s$	$100g$	λ/s
(1)	(2)	(3)	(4)	(5)	(6)
160	0.8	0.148	54.6	0.27	1.3
(c_i/c_a = 10)	0.5	0.132	54.2	0.25	1.2
	0.2	0.108	48.1	0.22	0.8
1600	0.8	0.383	44.0	0.87	2.7
(c_i/c_a = $10^{1/2}$)	0.5	0.327	50.6	0.65	2.6
	0.2	0.210	61.1	0.34	2.0
16 000	0.8	0.734	24.7	2.98	2.9
(c_i/c_a = 1)	0.5	0.525	35.0	1.45	2.9
	0.2	0.273	57.5	0.47	2.5
1 600 000	0.8	1.147	15.8	7.28	2.9
(c_i/c_a = 0.1)	0.5	0.681	24.4	2.48	3.0
	0.2	0.307	54.5	0.56	2.7

Contents of columns: (1) The temperature of the intercloud hydrogen, assuming T_i = 8000 K, and T_{mol} = 32 K. (2) The Mach number of the ionization front at the start of the interaction. (3) and (4) The dimensionless change in velocity and duration of the interaction. (5) One hundred times the dimensionless acceleration. True acceleration = $c_i^2 g/s$. (6) The ratio of maximum unstable wavelength to shell thickness.

Now use equations (1), (3) and (11b), replace $t_{rec} = 1/\beta n_i$ by $t_{rec,0}(n_0/n_i)$, and express R_{i0} and $t_{rec,0}$ in terms of n_0 and S_{48} (the luminosity in units of 10^{48} photon s^{-1}) so that $c_i t_{rec,0}/R_{i0} = 0.036\,(S_{48}\,n_0)^{-1/3}$ assuming T_i = 8000 K. Then Capriotti's condition becomes

$$n_0 S_{48} < [0.108\,gK(c_i/c_{shell})^2 f^{5/7}(1+f^{-6/7}/4)/(1-f^{-6/7})]^3. \tag{15}$$

In equation (15), the expression $K(c_i/c_{shell})^2$ is given in the two cases by (*cf.* equations 7–10)

$$K(c_i/c_{shell})^2 = (c_i/c_{mol})^4 f^{-6/7} \quad \text{(isothermal)}, \tag{16}$$
$$(c_i/c_{mol})^2/\beta_0 \quad \text{(magnetic)}.$$

Since we shall be concerned only with well-established shells, f will be large, and equation (15) can be approximated by (taking $c_i/c_{mol} = 10^{1.5}$)

$$n_0 S_{48} \lesssim 10^9 (108g)^3 f^{-3/7} \quad \text{(isothermal)}, \tag{17}$$
$$(108g/\beta_0)^3 f^{15/7} \quad \text{(magnetic)}.$$

For f = 43 and 215, corresponding to U_i/c_i = 0.2 and 0.1, we have the respective limits on $n_0 S_{48}$ of 3×10^3 and 10^5 in the magnetic case, and 2×10^8 and 10^8 in the isothermal case. Here we have assumed $108g$ or $108g/\beta_0$ is unity (see Table 1).

It can be seen that while in the isothermal case the instability will not be damped by conditions met in real H II regions, in the magnetic pressure dominated case (with β_0 of order unity) only the final stages of evolution will be unstable. In this case also the magnetic field may inhibit the actual development of the Rayleigh–Taylor instability.

The largest wavelength that has grown significantly in time t is $\lambda = (1/2\pi)\,a t^2$, where a is the physical acceleration, and $a = (c_i^2/s)g$ and in this case $t = \Delta t$. Hence

$$\lambda/s = (1/2\pi)\,(\Delta u/c_i)\,(c_i\Delta t/s) \tag{18}$$

Table 2. Masses of globules formed at different evolutionary phases.

U_i/c	t/t_0	R_i/R_{i0}	M_{glob}/M_0 (magnetic)	M_{glob}/M_0 (isothermal)
0.8	0.68	1.35	1.78×10^{-4}	2.78×10^{-7}
0.5	4.04	2.52	0.027	1.08×10^{-4}
0.4	7.48	3.39	0.1447	9.04×10^{-4}
0.3	15.6	4.98	1.03	1.15×10^{-2}
0.2	41.7	8.55	13.8	0.346
0.1	214.0	21.5	970.0	97.0

Velocity, time and radius of shell break-up, with globule mass ratios for magnetic and isothermal shells. To find actual masses, multiply these by 25 $(S_{48}/n_0)\,M_\odot$.

which quantity is given for the particular cases, in Table 1. In nearly every case λ/s is greater than unity. Thus during the acceleration, the shell is likely to break up into lumps of characteristic size equal to the shell thickness.

5 The globules

Assuming the instability has occurred, the fragments will, according to Table 1, column 6, be of characteristic size equal to the shell thickness s. Thus they will have masses of order (taking molecular weight per H nucleus to be unity) M_{glob} given by

$$M_{glob} = m_H N_{shell}{}^3/n_{shell}{}^2$$
$$= M_0 f^{12/7}(1-f^{-6/7})^3/K^2, \tag{19}$$

where $M_0 = (R_{i0}/3)^3 n_0 m_H$ is $(1/36\pi)$ times the initial mass of the Strömgren sphere. Using the Strömgren condition this can be written approximately as

$$M_0 \simeq 25(S_{48}/n_0)M_\odot, \tag{20}$$

where S_{48} is the ionizing luminosity in units of 10^{48} photon s^{-1}, and $M_\odot$ is the mass of the Sun.

Values for the masses corresponding to the isothermal and magnetic ($\beta_0 = 1$) cases are shown in Table 2, assuming shell break-up occurs at the given values of U_i/c_i, and that $c_i/c_{mol} = 10^{1.5}$. The globules so formed are close to pressure equilibrium with the ionized material which now surrounds them, the pressure ratio being $1+(U_i/c_i)^2$, so they will remain approximately the same size as when the shell first fragmented.

6 The fate of the globules

At the time when the shell fragments to form globules, it is moving at velocity U_i, and its radius and the ionized hydrogen density are R_i and n_i.

The fragmentation reduces the impedence of the shell, and the ionized material can flow out between the globules. After a few crossing times, the various interactions will have propagated away, and the flow will be approximately steady. The H II cavity can be approximated by the geometry that applied before shell break-up, and the outer part by a radially divergent flow of ionized material, which terminates in a shock where it is halted by the intercloud medium.

The radially divergent flow will accelerate down a pressure gradient, and may be approximated by the flow from an ionization bounded globule, as described by Kahn (1969).

The effect of the radial flow is to reduce the absorption of ionizing radiation so that the shell fragments (globules) are no longer shielded from a major part of the star's radiation. This will cause enhanced evaporation of the globules, and they will be driven outwards by the rocket effect (Oort & Spitzer 1955; Elmegreen 1976). As discussed by Mendis (1969) the ionized gas leaves the front with density n_e, velocity $2c_i$. The trail of newly ionized gas will lie on the side of the globule nearest the star.

The globule has a neutral column density of N hydrogen nuclei per unit cross-section, which initially equals N_{shell}. Thus

$$dN/dt = -2n_e c_i \tag{21}$$

The expulsion of the ionized material drives the globule outwards with velocity u, starting with velocity U_i.

Conservation of momentum can be written

$$N\,du/dt = -2c_i dN/dt, \tag{22}$$

whence

$$N = N_{shell} \exp\{-(u - U_i)/2c_i\}. \tag{23}$$

The rate of ionized particle production by the globules is $8\pi\gamma R_i^2 n_e c_i$, where γ is the fraction of the shell that remained as globule material at break-up (here taken to be $\sim 1/4$). The outflow rate assuming that the sonic point is at the shell (the fragments acting as a 'nozzle') is approximately $2\pi R_i^2 n_i c_i$, much less than the production rate unless the density of ionized material in the cavity is so great that it chokes off the globule evaporation entirely. The cavity filling time

$$t_{fill} = (4\pi/3) n_i R_i^3 / 8\pi\gamma R_i^2 n_e c_i = (n_i/3\gamma n_e) t_{cross} \tag{24}$$

and is much less than the crossing time $t_{cross} = R_i/c_i$ if $n_i \ll n_e$, so the H II cavity stays full. However, it cannot become too full or the evaporation will be choked off, and so it will approximate to a uniform density region of radius r bounded by the globules in which ionized particle density n_j is given by

$$n_j^2 r^3 = n_i^2 R_i^3. \tag{25}$$

Ionized material produced by the inner shell will not contribute significantly. Requiring mass balance between production, outflow and expansion of the H II region cavity then gives

$$8\pi R_i^2 \gamma n_e c_i = 2\pi r^2 n_j c_i + d/dt\,(4\pi n_j r^3/3)$$

whence with equation (25)

$$n_e/n_i = (1/4\gamma)\,(r/R_i)^{1/2}\,(1 + u/c_i). \tag{26}$$

From equations (26), (23), (22) and (21),

$$\exp\{-(u - U_i)/2c_i\}\; u\,du/dr = (n_i c_i^2/\gamma N_{shell})\,(r/R_i)^{1/2}\,(1 + u/c_i) \tag{27}$$

so

$$X(U_i) - X(u) = A\,[(r/R_i)^{3/2} - 1]. \tag{28}$$

Here

$$A \simeq 2\gamma^{-1}\,(U_i/c_i)^2\,\exp(-U_i/2c_i)$$

Table 3. Destruction time and radius for globules.

f	U_i/c_i	t_d/t_{cross}	R_d/R_i
1.68	0.8	0.04	1.62
5.04	0.5	0.25	2.33
8.48	0.4	0.45	2.88
16.60	0.3	0.88	3.91
42.75	0.2	2.05	6.26
215.44	0.1	7.39	14.88

and

$$X(u) = 2\exp(-u/2c_i) - \exp(\tfrac{1}{2})\,E_1[(c_i+u)/2c_i],$$

where $E_1(z)$ is the exponential integral.
As $u \to \infty$, $N \to 0$ at $r = R_d$ given by

$$R_d/R_i = [1 + X(U_i)/A]^{2/3} \tag{29}$$

t_d is given by

$$t_d = \int_{u_i}^{\infty} u^{-1}\,(dr/du)\,du$$

$$= t_{cross}\,A^{-1} \int_{u_i}^{\infty} \{1 + A^{-1}[X(U_i) - X(u)]\}^{-1/3} \exp(-u/2c_i)\,du/(c_i+u). \tag{30}$$

Values for destruction radius and time are given in Table 3 for the isothermal case. These results for the destruction time t_d can be approximated by $t_d/t_{cross} \cong 0.080\,f^{6/7}$, or

$$t_d/t_0 \simeq (f/4)^{10/7}. \tag{31}$$

Time t_0, for $S_{48} = 40$ and $n_0 = 1000\,\mathrm{cm}^{-3}$, is 5.6×10^4 yr. If $U_i/c_i < 0.2$, $t_d > 1.7 \times 10^6$ yr, while the time for the H II region to evolve to $U_i/c_i = 0.2$ is approximately 2.4×10^6 yr, reaching a radius 8.6 times its initial radius of ~1 pc, that is 9 pc. Thus the lifetime of globules becomes comparable with the main sequence age of an OB star. Under these conditions, the initial mass of an individual globule may be $0.02\,M_\odot$. For higher cloud densities or earlier stages of breakout, globules will be less significant.

The total mass converted into globules, M_{tot}, may be estimated from the shell mass at break-up if a fraction γ of the shell is subjected to the instability. For U_i/c_i not too large, this is (*cf.* equation 2)

$$M_{tot} \approx \gamma M_{i0} f^{12/7} = \gamma M_{i0} (U_i/c_i)^{-4}. \tag{32}$$

With $\gamma = 1/10$, this may be many times the mass of the initial Strömgren sphere, M_{i0}.

7 Conclusions

The multiple interaction as an ionisation front reaches the edge of a molecular cloud (both assumed plane and parallel) have been traced using a simplified analytic technique. In particular the initial acceleration of the front has been demonstrated to be Rayleigh–Taylor unstable, in the case where the shocked gas cools rapidly, and magnetic fields do not dominate. Magnetic fields may completely suppress the instability.

The fragment sizes are determined by two parameters S_{48}/n_0 and U_i/c_i. The later evolution of these fragments is determined by the rocket effect, blanketed by the reservoir of evaporated material in the cavity. Even with this blanketing, the globules are completely evaporated in a few crossing times. However, if the OB stars are buried more than eight initial Strömgren radii inside the molecular cloud when they are born, the globules may exist for a time comparable to the star's main sequence lifetime.

If the density of the cloud is less than $10^3\,\mathrm{cm}^{-3}$, then in the above circumstances the initial masses of the globules may be of the order of $0.02\,M_\odot$, and larger if the H II region were more deeply embedded in the cloud.

Acknowledgment

I thank the referee, A. Whitworth, whose painstaking and constructive criticism considerably improved the presentation of this paper.

References

Axford, W. I., 1961. *Phil. Trans. R. Soc. Lond.*, **A235**, 301.
Bok, B. J. & Reilly, E., 1947. *Astrophys. J.*, **105**, 255.
Bradley, J. N., 1962. *The Physics and Chemistry of Shock Waves*, Methuen, London.
Briscoe, F., 1972. *PhD thesis*, University of Leeds.
Capriotti, E. R., 1973. *Astrophys. J.*, **179**, 495.
Elmegreen, B. G., 1976. *Astrophys. J.*, **205**, 405.
Giuliani, J. L., 1979. *Astrophys. J.*, **233**, 280.
Goldsworthy, F. A., 1961. *Phil. Trans. R. Soc. Lond.*, **A253**, 277.
Israel, F. P., 1978. *Astr. Astrophys.*, **70**, 769.
Kahn, F. D., 1969. *Physica*, **41**, 172.
Marsh, M. C., 1970. *Mon. Not. R. astr. Soc.*, **147**, 95.
Mendis, D. A., 1969. *Mon. Not. R. astr. Soc.*, **142**, 441.
Oort, J. H. & Spitzer, L., 1955. *Astrophys. J.*, **121**, 6.
Spitzer, L., 1978. *The Physics of the Interstellar Medium*, Wiley-Interscience, New York.
Tenorio-Tagle, G., 1979. *Astr. Astrophys.*, **71**, 59.

Appendix A: The (p, u) diagram

The pressure–velocity diagram (*cf.* Bradley 1962) has been so useful in determining the results of wave interactions that its use is described here. The interactions considered are plane parallel isothermal interactions between shocks, rarefaction waves (RW), contact discontinuities (CD), and *weak D-type* ionization fronts (IF). The Rankine–Hugoniot relations across shock fronts, and the corresponding jump conditions across other types of wave or front, can be transformed to a relationship between the pressures and 'lab frame' velocities on either side of the front. In the isothermal case the relations are particularly simple, the pressure ratio being a function of velocity difference.

The range of consequences of the interaction can be limited at once by noting that a shock front moves supersonically relative to gas upstream and subsonically relative to gas downstream, while an RW moves sonically (and a weak D-type IF moves subsonically) with respect to both upstream and downstream gas. Thus an RW cannot overtake or be outstripped by an RW, but will overtake a shock front. A shock front will overtake a shock front or an RW. An IF and a CD behave similarly, and do not overtake anything, so only one shock or RW can emerge in either direction from the interaction, leaving behind a single CD or IF.

Table A1. The relation between pressure and velocity for various disturbances.

Disturbance	P/P_x
Contact discontinuity ($Z = 0$)	1
Shock	$1 + Z^2/2 + Z\,[1+(Z/2)^2]^{1/2}$
Rarefaction wave	$\exp(-Z)$
Ionization front (weak D)	$(1+Z^2)^{-1}$
Ionization front (weak R)	$(c_i/c_x)^2(1+Z^2)^{1/2}$

The ratio (P/P_x) of the pressure downstream to the pressure upstream (P_x) is given as a function of the normalized velocity change $Z = (u - u_x)/c$.

The disturbances are assumed isothermal in each phase (ionized, neutral, atomic, molecular). c is the isothermal sound speed; in the case of the ionization front, the value for ionized gas is used.

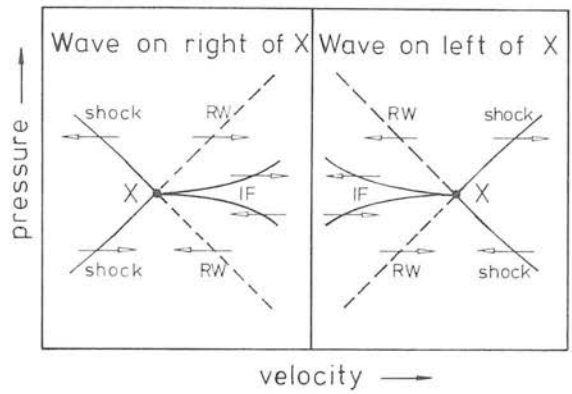


Figure A1. Diagrams showing the pressure and velocity across wave transitions to the left or right of point X, and moving left or right, as indicated by arrows. Rarefaction waves are shown in this and Figs A2, A3 and A4 by dashed lines, shocks and ionization fronts (marked IF) by continuous lines.

The pressure–velocity transitions are given in Table A1. These functions can be used to construct a diagram of transitions from a given point X to conditions on the other side of a wave to the left or right of X. This diagram, which is a graph of P/P_x versus u/c, is shown in Fig. A1. The isothermal sound speed is c. The arrows indicate the direction (left- or rightward) of motion of the wave. The ionized gas sound speed c_i is used to normalize the velocities involved in IF transitions. Note that the slope of the transition locus for weak shocks or RW has magnitude unity.

Consider as an example the interaction between two approaching RW, diagrammed in Fig. A2. The boundary conditions are represented by a and b, the conditions between the incident waves are represented by d, and the conditions between the waves after interaction by e.

The initial RW are represented by the transitions ad and db. After the interaction, the transitions are from a to b, via a left-moving wave to the right of a, and a right-moving wave to the left of b. Fig. A2 shows that these must be RW, and they define conditions at e between the waves. The same notation is used for the other interactions.

Fig. A3 shows the (p, u) diagrams for: (a) colliding shocks; (b) a shock colliding with an RW; (c) a shock overtaking a shock; (d) a shock overtaking an RW; and (e) an RW overtaking a shock.

The first three cases are simple. In each case adb is the initial situation and aeb the final one. In the fourth case, a shock overtaking an RW, Fig. A3(d) shows that a shock is reflected

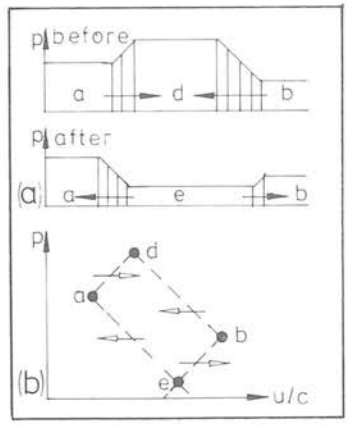


Figure A2. (a) Sketch of two rarefaction waves before and after head-on collision. Pressure is plotted versus distance. (b) The (p, u) transitions corresponding to (a). Comparison with Fig. A1 shows that two rarefaction waves result. The notation is used in Figs A3 and A4.

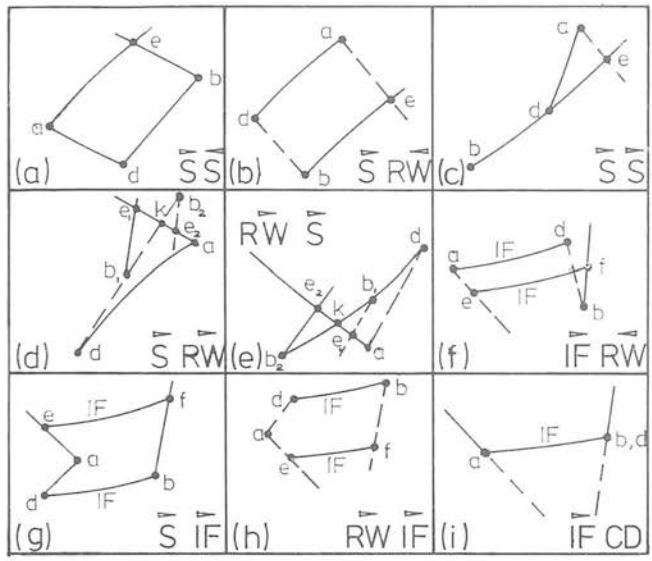


Figure A3. Approximate (p, u) diagrams for the interactions indicated. The direction the wave travels can be deduced by noting that all left-moving waves have negative slope in the diagram.

and either a shock or RW is transmitted. For very weak interactions the condition under which a shock is transmitted is that the pressure behind the incident shock is greater than the pressure in front of the initial RW ($p_a > p_b$). For stronger interactions there is a slight correction to this. By reference to Fig. A3(d), the condition is $p_k > p_b$, where p_k has to be found by solving the relevant equations of Table A1 iteratively. Similar considerations apply to an RW overtaking a shock. Two shocks result if $p_k > p_b$.

Similar analysis will show that a shock hitting a CD produces two shocks if the gas it meets at the CD is denser and cooler, and an RW hitting a CD produces two RW if the gas it meets is hotter and less dense.

In cases involving IFs there is the additional complication of the ionizing radiation. In order that the velocities can be assumed constant before and after interaction, the ionizing

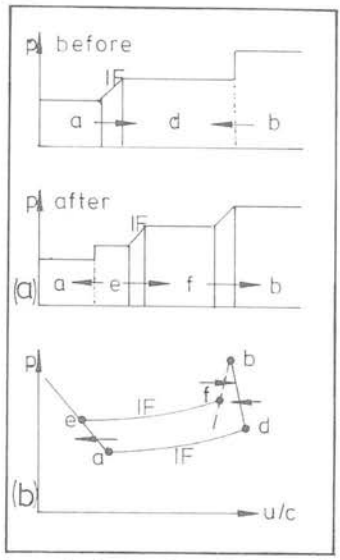


Figure A4. (a) Sketch of a weak D-type ionization front before and after collision with a shock. (b) The (p, u) diagram for (a).

flux emitted by the star must increase linearly with time, since this is how the volume of a plane parallel H II region increases with time if velocities are constant. Over the short period involved in the interaction this is not a stringent requirement, and we shall use it hereafter. The radiation condition (Goldsworthy private communication; Marsh 1970) is the condition that all the ionizing radiation is absorbed within the H II region behind the IF. If $J_* t$ is the number of photons produced per second divided by the recombination coefficient, then we can write a differential condition on the motion of the IF,

$$J_* \Delta t = n_b^2 U_i \Delta t. \tag{A1}$$

This applies to the simplest case of steady motion of the IF with velocity U_i, leaving number density n_b of hydrogen behind. The right hand side of this equation is the integral of n^2 over the volume of H II region newly created during time Δt, while the left hand side is the number per second of new ionizing photons available, divided by the recombination coefficient.

Fig. A4(a) and (b) are the diagrams for the collision of a weak D-type IF with a shock: ad is a rightward moving IF, and db is a leftward moving shock. To decide on the consequences of the interaction a differential approach to the transition equations is used. Since these equations are all monotonic in the relevant variables, the result of the interaction can be deduced by choosing an arbitrarily weak incident shock in this case. This will produce weak resultant waves, and so, whether shocks or RW, they will all obey transition equations of the form

$$p_2/p_1 \simeq 1 \pm (u_2 - u_1)/c, \tag{A2}$$

where $(u_2 - u_1)/c$ is small. All expressions can be expanded as far as first order in such terms.

In the present case the ionisation balance before the interaction is

$$J_* \Delta t = n_a^2 U_{i0} \Delta t$$

and after it,

$$J_* \Delta t = n_e^2 (U + U_i) \Delta t - n_a^2 U \Delta t .$$

Here U is the (backward) velocity of the resultant wave moving into the H II region, while U_{i0} and U_i are the velocities of the IF before and after interaction. In Appendix B it is shown that for weak waves U is the same for shocks and RW and is

$$U = c_i - (u_a + u_e)/2. \tag{A3}$$

From the above equations the radiation condition is

$$n_e^2 (U + U_i) = n_a^2 (U + U_{i0}). \tag{A4}$$

Now define

$$\begin{aligned} \delta_1 &= (u_d - u_b)/c_i, \\ \delta_2 &= (u_f - u_b)/c_i, \\ \delta_3 &= (u_a - u_e)/c_i, \\ \Delta &= (u_b - u_a)/c_i, \end{aligned} \tag{A5}$$

where the δs are much less than unity, while Δ is between zero and unity for a weak D-type IF. Next, note that for D-type IFs, the relative velocity of gas ahead of the front is very small compared with that behind (provided $c_i/c_{mol} \gg 1$), so that (see Fig. A4a)

$$\begin{aligned} U_i &\simeq U_f \\ U_{i0} &\simeq U_d . \end{aligned} \tag{A6}$$

Also, applying equations (A3) and (A4)

$$U/c_i \simeq 1 + \delta_3/2 - u_a/c_i \tag{A7}$$

and

$$n_e^2/n_a^2 = p_e^2/p_a^2 \simeq 1 + 2\delta_3 \tag{A8}$$

so the radiation condition (A4) becomes

$$2\delta_3 + \delta_2 = \delta_1 . \tag{A9}$$

The transition equations can be used to find another relation between the δs. For

$$\begin{aligned} \frac{P_f}{P_b} &= \frac{P_f}{P_e}\frac{P_e}{P_a}\frac{P_a}{P_d}\frac{P_d}{P_b} \\ &\simeq [1 + (\Delta + \delta_2 + \delta_3)^2]\,(1 + \delta_3)\,[1 + (\Delta + \delta_1)^2]^{-1}\,(1 + \beta\delta_1)^{-1}, \end{aligned}$$

where β is the ratio of ionized to unionized gas sound speeds, $\beta \gg 1$, although $\beta\delta_1$ is still small. But

$$\frac{P_f}{P_b} \simeq 1 + \beta\delta_2,$$

so, putting $\alpha = 2\Delta/(1 + \Delta^2)$,

$$(1 + \alpha)\delta_3 - (\beta - \alpha)\delta_2 - (\beta + \alpha)\delta_1 = 0. \tag{A10}$$

Eliminating δ_3 from equations (A9) and (A10) gives

$$\delta_2 = \delta_1 [\beta + (\alpha - 1)/2]/[\beta - (\alpha - 1)/2], \tag{A11}$$

also

$$\delta_3 = \delta_1 \{1 + [\beta + (\alpha - 1)/2]/[\beta - (\alpha - 1)/2]/2\}. \tag{A12}$$

Equation (A11) shows that δ_1 and δ_2 have opposite signs. Now equation (A5) and Fig. A4(b) show that δ_1 is positive, so δ_2 is (provided $\beta \gg 1$) negative. Thus $u_b > u_f$ and a rarefraction wave is reflected back into the neutral gas, while a shock proceeds into the H II region.

Figure A3(f) is the corresponding diagram for an RW colliding with a weak D-type IF, and the above analysis applies here also. However, in this case a shock is reflected, and the RW continues on into the H II region. In the case of a shock overtaking an IF (Fig. A3g), the radiation condition is slightly more complex

$$n_e^2 (U + U_i) = n_a^2 (U + [1 - (n_d/n_a)^2] U_{s0} + (n_d/n_a)^2 U_{i0}), \tag{A13}$$

where U is the (backward) velocity of the reflected wave and U_{s0} is the velocity of the incident shock wave. The analysis goes through as before, and

$$\delta_2 = \delta_1 [2(\beta - \alpha)(1 - \Delta) + (\alpha - 1)]/[2(\beta - \alpha)/(1 + \Delta) + (\alpha + 1)] \tag{A14}$$

and so

$$\delta_3 = 4\delta_1 [1 - \alpha\Delta]/[2(\beta - \alpha)(1 + \Delta) + (\alpha + 1)]. \tag{A15}$$

Equations (A14) and (A15) show that two shocks result from a shock overtaking an IF.

Table A2. Consequences of isothermal interactions.

Before		After		
RW →	← RW	← RW		RW →
S →	← S	← S		S →
S →	← RW	← RW		S →
S →	S →	← RW		S →
S →	RW →	← S		S → (or RW →)[1]
RW →	S →	← S		RW → (or S →)[2]
S →	CD	← S (or ← RW)[3]		S →
RW →	CD	← RW (or ← S)[4]		RW →
IF →	← S	← S	IF →	RW →
IF →	← RW	← RW	IF →	S →
S →	IF →	← S	IF →	S →
RW →	IF →	← RW	IF →	RW →
IF →	CD →		IF	

[1] RW transmitted if $p_k (\simeq p_a) > p_b$.
[2] Shock transmitted if $p_k > p_b$.
[3] RW reflected if b is hotter than a.
[4] Shock reflected if b is hotter than a.

Note the similarities between the IF interactions and the CD interactions above.

Correspondingly, with the same analysis and referring to Fig. A3(h), an RW overtaking an IF produces two RW. These cases should be compared with interactions involving CDs.

Finally (Fig. A3i) we consider a weak D-type IF overrunning a CD between states d and b. The transition lines, initially ad, d = b become ae, ef and fb. Thus the result is either two shocks or two RW. The radiation condition, however, is equation (A4). We replace U_i by U_e and U_{i0} by U_d, and we note that U is the (backward) velocity of the wave reflected back into the H II region so that $(U + u_d)$ is positive. Now suppose that two shocks, ae and fb, are the result of interaction. Then $n_e > n_a$, so $u_f < u_d = u_b$. So fb is an RW! It can only be both shock and RW if it is of zero strength. Thus the resultant transition is ab. Only the relative upstream velocity has changed, and this was so small that it is ignored. The IF swallows the CD.

Similar considerations can be applied to weak R-type IF (which in the extreme case can be thought of as instantly heating everything to the same temperature), and to strong D-type fronts. However, in this latter case an extra constraint related to the allowable sub-to-supersonic transition inside the front must be applied (Axford 1961; Briscoe 1972). These matters will not be considered further here. Although conditions have been assumed to be isothermal, similar analysis could be applied to gas governed by any equation of state.

The interactions discussed here are summarized in Table A2.

Appendix B: Solution of transition equations

The transition equations for the series of interactions delineated in Fig. 1 can be solved in the following way.

First (referring to Fig. 1) approximate $u_B = U_{i0}$ because $U_{i0} - U_B$ is so small (*cf.* equation 15), and $u_A = U_{i0}/2$. Use the notation $f(z)$ for the shock relation in the second column of Table A1. Then the sequence of solution is:

$u_c, p_c, U_{i0} = u_B$ are given

so (shock) $p_B = p_c f(y), \qquad c_{mol} y = u_B - u_C$ (B1)

and (IF) $p_A = p_B \left[1 + (u_B - u_A)^2/c_i^2\right]^{-1}$ (B2)

also (shock) $p_E = p_C f(x), \qquad c_a x = u_E - u_C$ (B3)

and (RW) $p_E = p_B \exp\left[(u_B - u_E)/c_{mol}\right]$ (B4)

so from equations (B3) and (B4).

$$x = (c_{mol}/c_a) \ln\left[(p_B/p_C)(\exp y)/f(x)\right]. \tag{B5}$$

Solving this iteratively for x then gives u_E, p_E.

Next

$p_F = p_A \exp w, \qquad c_i w = u_F - u_A,$ (B6)

$p_G = p_F \left[1 + (u_G - u_F)^2/c_i^2\right],$ (B7)

$p_G = p_E f(z), \qquad c_{mol} z = u_G - u_E,$ (B8)

therefore

$$f(z) = (p_A/p_E)\left[1 + (u_E + z c_{mol} - u_A - w c_i)^2/c_i^2\right] \exp(-w)$$

If the right-hand side is $R(z, w)$ then this is

$$z = (R - 1) R^{-1/2}. \tag{B9}$$

The radiation condition is discussed in Appendix A. There it is assumed that the wave in the H II region is a discontinuity travelling with speed U. This is correct for a shock wave, but an RW is of finite thickness, and absorption in the wave must be taken into account. In the isothermal case, it is possible to show that, as far as absorption goes, one can approximate a *centred* RW by a discontinuity travelling in the direction of the wave with speed equal to the sound speed plus the average of 'before' and 'after' velocities (in the same direction) in the gas. The error in the velocity of this equivalent discontinuity is only $O([u_F - u_A]^2/6c)$ which is small in all cases investigated. This is also true for weak shocks. This approximation is used for the *non-centred* RW involved in the interaction. The radiation condition is $n_F^2 (U_i + U) = n_A^2 (U_{i0} + U)$ and using $n_F/n_A = p_F/p_A = \exp w$, and $U = c_i - \frac{1}{2}(u_A + u_F)$,

$$w = (\tfrac{1}{2}) \ln [(u_E - u_A + zc_{mol} + (1 - w/2)\, c_i)/(u_A + (1 - w/2)\, c_i)]. \tag{B10}$$

Equations (B9) and (B10) are iterated cyclically to solution, which then gives u_F, p_F, u_G, p_G.

If further stages of the interaction are required, then states A, B, C are replaced by states F, G, E and the sequence of solutions is restarted. To find g and λ/s in Table 1 of the paper, we require an estimate of the time $\Delta t = t_3 - t_2$ (see Fig. 2). Although the IF is continuously accelerated from u_B to u_G in this time, only a small error in the estimate will be made by replacing the IF trajectory by a straight line of slope $(u_B + u_G)/2$. Then an elementary geometric construction in Fig. 1(a) shows that

$$c_i \Delta t/s = (c_i/c_{mol})\, (u_E - u_B)/(c_{mol} + [u_G + u_B]/2 - u_E). \tag{B11}$$