Spectroscopic Techniques and Observations in Infrared Astronomy

Andrew Zadrozny

A thesis submitted for the degree of
Doctor of Philosophy of the University of London
and for the
Diploma of Membership of the Imperial College.

January 1990.

Abstract

This thesis deals with several instrumental and observational aspects of astronomical spectroscopy at near infrared wavelengths.

An introductory general review of current instrumentation and techniques for infrared astronomy is followed by an account of aspects of the design of a near-IR cooled grating spectrometer and a multi-channel infrared photometer. A data acquisition system common to both these instruments is described in detail, as are the procedures used for reducing and analysing data produced by these instruments.

The spectrometer has been used to obtain high S/N ratio spectra of bright main sequence and red giant stars. These data are presented, and their analysis is described.

The final two chapters describe spectroscopic observations of the interstellar medium.

The first of these chapters contains an account of an attempt to detect near-IR line emission from the cosmologically significant molecule HD. An upper limit for such emission from the Orion Molecular Cloud is obtained, and its significance discussed.

The thesis concludes with a description of molecular hydrogen observations made towards the OH maser source in W75 N. The measurements were made with a view to gaining a better understanding of the mechanisms responsible for producing the OH and driving the maser emission. The results are discussed, and more recent findings presented.

Contents

Abstract 2

Contents 3

Chapter 1: Infrared Instrumentation 8

1.1 Introduction 8

1.2 Detectors 8

1.2.1 Noise 8

a) Johnson Noise 9
b) Generation-Recombination Noise 9
c) Flicker (1/f) Noise 9
d) Phonon Noise 9
e) Microphonic Noise 9
f) Signal Photon Shot Noise 10
g) Background Photon Shot Noise 10
h) Excess Sky Noise 11

1.2.2 Photodetectors 11

1.2.3 Arrays 12

1.2.4 Bolometers 15

1.3 Photometric Apparatus and Techniques 15

1.3.1 IR Telescopes 15

1.3.2 Signal Detection 16

1.3.3 Photometric Systems 17

1.4 Infrared Instruments 18

1.4.1 CVF Spectrometer 18

1.4.2 Grating Spectrometer 18

1.4.3 Fabry-Perot Spectrometer 19

1.4.4 Fourier-Transform Spectrometer 20

1.4.5 Heterodyne Spectrometer 21

1.4.6 Polarimetry 22

1.4.7 High Resolution Techniques 23

1.5 Surveys 24

1.5.1 Two-Micron Sky Survey 24

1.5.2 Air Force Sky Survey 25

1.5.3 IRAS Survey 25

1.6 Facilities 26

1.6.1 Ground-Based Telescopes 26

1.6.2 Balloons 26

1.6.3 Aircraft 26

1.6.4 Spacecraft 27

a) IRAS 27
b) SL2/IRT 28
c) Hubble Space Telescope NICMOS 28
d) ISO 28
e) COBE 29
f) SIRTF 30
g) Future Sub-Millimetre Observatories 30

Chapter 2: The Four-Channel Photometer and the Cooled Grating Spectrometer 31

2.1 Introduction 31

2.2 The 1.5m Infrared Flux Collector 31

2.3 The Imperial College Four Channel Photometer 32

2.3.1 Optical Design 32

2.3.2 Mechanical Configuration 34

2.3.3 Detectors and Pre-amplifiers 34

2.3.4 Amplifiers and Filters 36

2.3.5 Phase Sensitive Detectors 37

2.3.6 AIM 65 Computer and Interfaces 37

a) Up/Down Counters 38
b) Telescope Control Interface 38
c) Analogue Interface 38
d) Guide Camera Crosshead Interface 39
e) Serial Communications Interface 39
f) Bus Monitor Module 39
g) Battery-Backed RAM Module 39

2.3.7 Control Software 39

2.3.8 Photometry Software 40

2.4 The Imperial College Cooled Grating Spectrometer 41

2.4.1 Optics 41

2.4.2 Mechanical Configuration 43

2.4.3 Detector and Pre-amplifiers 44

2.4.4 Motor Drive 44

2.4.5 Position Encoders 45

2.4.6 Data Acquisition Software 47

2.5 Observing and Data Reduction Procedures 47

2.5.1 Observations 47

2.5.2 Data Reduction 50

a) Wavelength Registration 50
b) Photometric Accuracy 51
c) Determination of Atmospheric Extinction 52

2.5.3 Conclusion 54

Chapter 3: A Search for the HD Molecule in the Orion Molecular Cloud 55

3.1 Introduction 55

3.2 The Cosmological Significance of Deuterium 57

3.2.1 Cosmic Ray Spallation 58

3.2.2 Spallation in Supernova Shocks 58

3.2.3 Cosmological Processes 59

3.3 Interstellar Chemistry of Deuterium 63

3.3.1 Deuterium Fractionation 63

3.3.2 The Spectrum of HD 65

3.4 Measurements of Deuterium Abundance 68

3.4.1 Solar System Deuterium Measurements 68

3.4.2 Observations of the Deuterium 92 cm Line 69

3.4.3 Observations of Deuterium in Diffuse Clouds 69

3.4.4 Observations of Deuterium Bearing Molecules in Molecular Clouds 70

3.5 A Search for HD in the Orion Molecular Cloud 71

3.5.1 Molecular Hydrogen in the Orion Nebula 71

3.5.2 The Instrument 74

3.5.3 The Choice of Line 74

3.5.4 Observing Procedure 75

a) The Peak 1 Source 75
b) Spectral Calibration 75
c) Observation Parameters 76

3.5.5 Data Reduction 77

a) Despiking 77
b) Wavelength Drift 77
c) Weighted Mean 78
d) Gradient Subtraction 78
e) Correlation Detection 79
f) Flux Calibration 80
g) Result 81

3.5.6 Discussion 81

3.6 Conclusion 83

Chapter 4: Molecular Hydrogen Emission From OH Maser Sources 84

4.1 Introduction 84

4.2 Cosmic Masers 84

4.3 Interstellar OH Masers 86

4.3.1 Observations of Type I OH Masers 87

4.3.2 Pumping Mechanisms 88

4.4 The H_2 Observations 90

4.4.1 The OH Maser Source W75N(OH) 90

4.4.2 Observations of the H_2 S(1) Line in W75N(OH) 92

4.4.3 Extinction Correction 95

4.4.4 Comparison With Model 96

References 98

Appendix A: The Four-Channel Photometer Photometry Program 104

Appendix B: The Near Infrared Spectra of Seventeen Bright Stars 114

List of Tables

Table 2.1 The List of Observed Stars. 48

Table 2.2 Spectrometer Limiting Magnitudes. 50

Table 3.1 The v=1-0 Band Dipole Spectrum of HD. 66

Table 3.2 Journal of Observations. 76

Table 4.1 Predicted S(1) Line Flux From W3(OH). 93

List of Figures

2.1 The Four-Channel Photometer Beam Profiles. 33

2.2 The Detector Preamplifier Circuit. 35

2.3 The Data Acquisition System of the Four-Channel Photometer. 36

2.4 Photometry Nod Sequence 40

2.5 The Spectrometer Optics 42

2.6 Spectrometer Line Profile 43

2.7 The Data Acquisition System of the Spectrometer. 45

2.8 Coarse Encoder Interface Circuit. 46

2.9 Spectrum of Vega in the J Band. 49

2.10 A Cross-correlated Spectrum. 51

2.11 Extinction Coefficient at 1.093 μm. 53

2.12 Observed and Predicted Extinction Coefficient at J. 54

3.1 Elemental Abundances From Big-Bang Nucleosynthesis. 61

3.2 A H_2 Map of OMC-1. 73

3.3 A Raw HD Spectrum. 77

3.4 All Coadded Spectra. 78

3.5 The Final Matched Filtered Spectrum. 80

4.1 The OH Λ-Doublet Transitions. 86

4.2 A 6 cm Map of W75N. 92

4.3 The Observed Spectrum Uncorrected for Sky Emission. 94

4.4 The Observed H_2 Spectrum. 95

Chapter One
Infrared Instrumentation

1.1 Introduction

The most commonly used detectors of optical astronomy, the photovoltaic silicon devices, photomultipliers and photographic emulsions, all have cut-off wavelengths at around 1.1 μm. In terms of detector technologies, the infrared portion of the spectrum begins here.

The first infrared astronomical observations were made by Herschel in 1869. During the following ninety years, although infrared astronomy was not pursued with great vigour, steady progress was made. The transmission of the atmosphere was characterised, the temperatures of the planets were established, and some spectroscopy of the planets was done. The history of this period has been summarised recently by Sinton (1986) for example, and in part by Jennings (1986).

Modern infrared astronomy began in earnest in the early nineteen-sixties, when sensitive cooled detectors came into widespread use. Highlights of this era include the development of the germanium bolometer (Low, 1961), the extension of photometric standards to the infrared (Johnson, 1964) and the Two Micron Sky Survey (Neugebauer and Leighton, 1969).

Numerous ground-based and airborne observing facilities followed, which, together with more sensitive detectors, such as the InSb photodiode for short wave work, allowed spectroscopy and extragalactic studies to flourish.

Recent advances include the sky survey conducted by the IRAS satellite, and the emergence of large two-dimensional array detectors.

1.2 Detectors

Progress in infrared astronomy has been underpinned by the continuous advances made in detector technologies. Infrared detectors have increased in sensitivity by eight orders of magnitude over the last thirty years.

1.2.1 Noise

The sensitivity of a detector is determined by its quantum efficiency and its noise characteristics. The definition of quantum efficiency depends somewhat on the detection mechanism being considered, but for a photodetector it is the number of photoelectrons produced per incident photon. Detector noise is often expressed in terms of a Noise Equivalent Power, defined as the incident signal power needed to produce a voltage level equal to the r.m.s. noise voltage. Noise from several sources can be present in a detector.

a) Johnson noise

The Brownian motion of electrons within a resistance produces a randomly fluctuating voltage across it. The noise is "white", that is, constant with respect to frequency, and has the form

$$i = (4kTB/R)^{1/2} \qquad (1.2.1)$$

where i is the equivalent noise current, k is Boltzmann's constant, T is the absolute temperature, B is the measurement bandwidth and R is the resistance. On thermodynamic grounds it is the only source of noise possible in a detector at zero bias, and represents the fundamental sensitivity limit of a photovoltaic detector.

b) Generation-Recombination noise

Charge carriers within a semiconductor form and recombine randomly, causing variations in the resistance of the semiconductor. This results in voltage fluctuations across the semiconductor when a bias current passes through it. The equivalent generation-recombination noise current for an n-type semiconductor is

$$i = 2I/N\,(PB\tau/(1+(2\pi ft)^2)^{1/2} \qquad (1.2.2)$$

where I is the bias current, P and N are the total numbers of free holes and electrons, τ is the electron lifetime, f is the measurement frequency and B the bandwidth.

c) Flicker (1/f) noise

Despite its widespread occurrence in most electronic devices, $1/f$ noise has a poorly understood physical origin. Its name stems from the form of its variation with frequency,

$$i \propto I\,B^{1/2} f^{-1} \qquad (1.2.3)$$

and it is often the dominant noise source at very low frequencies. It should be noted that both generation-recombination noise and $1/f$ noise vanish at zero bias current. Photodiodes are operated at very low reverse bias voltages for this reason.

d) Phonon noise

Phonon noise is present in bolometric detectors. It is caused by the random flow of heat between the detector element and its heat sink, and represents the fundamental limit to the sensitivity of a bolometer. It has the value

$$P = 4T(kG)^{1/2} \qquad (1.2.4)$$

where P is the noise equivalent power, T is the temperature, k is Boltzmann's constant and G is the thermal conductance of the heat path between the detector and heat sink.

e) Microphonic Noise

Microphonic pickup is a common source of noise in practical detector circuits. The term refers to any electrical noise induced by mechanical vibration. It is caused by

flexible electrical connections to the detector and its associated circuitry and can be avoided by mechanically rigid construction of the detector circuit.

f) Signal Photon Shot Noise

As well as the forms of detector noise discussed above, the signal from the source itself is also subject to random fluctuations as a consequence of the discrete nature of the photon. The noise equivalent power of the fluctuations for photovoltaic and bolometer detectors is given by

$$\sigma = \frac{(2P_B h\nu/\eta)^{1/2}}{(1 - \exp(-h\nu/kT))^{1/2}} \qquad (1.2.5)$$

where P_B is the incident radiation power, ν is the measuring frequency, η is the detector's quantum efficiency and T is the temperature of the incident black body radiation. For photoconductors photon shot noise is $\sqrt{2}$ higher than the expression above, because both generation and recombination of photoelectrons contribute to it.

While signal shot noise limited operation is commonplace at visible wavelengths, it is very rarely achieved in ground-based IR observations.

g) Background Photon Shot Noise

Thermal background emission from the sky and telescope structure is generally the dominant source of photon shot noise at infrared wavelengths. In this case the background power incident on the detector is

$$P_B = \epsilon B_\lambda(T)\Delta\lambda t A\Omega \qquad (1.2.6)$$

where ϵ is the effective emissivity, $B_\lambda(T)$ is the Planck function, $\Delta\lambda$ is the optical bandwidth, t the transmission of the cold filter, A the telescope collecting area and Ω the solid angle viewed by the detector.

Substituting (1.2.6) into (1.2.5) yields the expression for thermal background shot noise power,

$$\sigma = \frac{(2\epsilon B_\lambda(T)\Delta\lambda t A\Omega\, h\nu/\eta)^{1/2}}{(1 - \exp(-h\nu/kT))^{1/2}} \qquad (1.2.7)$$

Thermal shot noise is the dominant noise source for broad band ground-based observations beyond about 2.5 μm, and reaches a maximum at 10 μm, that is, at the peak of 300 K black body emission. Sky shot noise is significant at shorter wavelengths as well, where the emission is not of thermal origin, but is produced by fluorescence from hydroxyl in the upper atmosphere. This OH "airglow" is present in the spectral range between about 0.8 μm and 3 μm, and peaks in the H band.

It should be noted that because background shot noise increases with the square root of collecting area, telescope sensitivity only increases as the square root of its collecting area in the background-limited regime.

h) Excess sky noise

It turns out that the fluctuations measured in sky emission exceed those expected from shot noise alone. This has been discussed by e.g. Papoular (1983) and Allen and Barton (1981). The excess noise typically has a $1/f^n$ spectrum, rising extremely sharply above the shot noise level at frequencies below 1 Hz or so. It can be explained as an atmospheric turbulence effect, where volumes of air of varying temperature and emissivity pass through the telescope beam.

1.2.2 Photodetectors

An intrinsic semiconductor acts as a photon detector if incident radiation can excite electrons from the valence band to the conduction band of the semiconductor. That is, such a detector is sensitive to photons whose energies exceed the semiconductor's band-gap, and its wavelength response exhibits a sharp threshold at that energy.

Photodetectors can be operated either in photoconductive or photovoltaic mode. A photoconductive detector consists of a crystal of undoped semiconductor across which a constant bias potential is maintained. Electrons excited into the conduction band cause an increase in current through the device proportional to the incident radiation flux, and this photocurrent is measured as a potential difference across a load resistor.

A photovoltaic detector is a junction diode formed at the interface between p-doped and n-doped semiconductor material. Electron-hole pairs produced in the diode's depletion layer will be separated by the junction's internal electric field before they can recombine, producing a voltage across the diode proportional to the number of incident photons. Photodiodes are normally operated at zero bias to minimise noise, and are connected in series with a load resistor. The photocurrent passing through this resistor produces a voltage across it which is amplified and measured.

Because of their high quantum efficiencies, which can exceed 70%, and the possibility of very low noise operation, photodetectors are among the most commonly used IR detectors.

The two elemental semiconductors, silicon and germanium, have bandgaps of 1.1 eV and 0.68 eV respectively, corresponding to cut-off wavelengths of 1.1 μm and 1.8 μm, (though the latter reduces to 1.6 μm when cooled to 77 K). While silicon is by far the most commonly used photoconductor for visible wavelength detectors, germanium has been little used as a near IR detector material since it offers only a limited additional spectral response. Lead sulphide proved more suitable for this role. It has a cut-off at 3.8 μm and can attain a NEP of 10^{-14} W $Hz^{-½}$. Used in photoconductive mode, it was the prevalent near IR detector before InSb photodiodes became available.

Indium antimonide is an intrinsic semiconductor with an exceptionally narrow bandgap of 0.23 eV at 77 K, corresponding to a cut-off wavelength of 5.5 μm. It first attracted interest as a potential infrared detector material in the early nineteen-sixties,

but it was not until the mid seventies that InSb photodiodes of sufficient quality could be produced to serve as astronomical detectors (e.g. Hall et al., 1975). Because of its low noise characteristics and long wavelength response, the InSb photodiode then quickly displaced lead sulphide as the detector of choice for near-IR applications. These devices have been developed to a high degree of maturity. Commercial InSb detectors are available mounted alongside preamplifiers in hybrid packages, (e.g Smith et al., 1982), and NEPs of 2×10^{-16} W $Hz^{-½}$ are achievable.

While InSb is the narrowest bandgap binary semiconductor compound, some ternary compounds with even smaller bandgaps exist, notably HgCdTe and PbSnTe. Cadmium mercury telluride, in particular, has been heavily developed as a detector material in recent years, mainly as a consequence of military interest. The bandgaps of these ternary compounds are adjustable by varying the Hg/Cd or Pb/Sn ratio and HgCdTe, for example, can be manufactured with a cut-off wavelength of between 4 μm and 14 μm. Its noise properties are competitive with those of InSb, and it seems likely to be one of the most important detector materials of the coming years, particularly for near-IR array applications.

At longer wavelengths, photoconductive detection is still possible by the use of extrinsic semiconductors. These are produced by introducing impurity atoms into the lattice of a semiconductor crystal, so creating additional energy levels within the forbidden energy band. A photon is detected when it ionises an impurity atom, placing an electron in the conduction band of the bulk semiconductor.

Several extrinsic silicon and germanium materials have been developed as astronomical detectors, including, (with cut-off wavelength), Si:In (8 μm), Si:Ga (18 μm), Si:As (23 μm), Si:P (28 μm), Ge:Be (45 μm) and Ge:Ga (120 μm). Several of the short wave materials have sensitivities that current intrinsic detector materials for the 10μm band cannot match, while the long wave detectors like Ge:Ga have unique capabilities in terms of spectral response. Even longer wave photodetectors than these are possible, though. The impurity ionisation energy can be reduced by stressing the crystal, and stressed Ge:Ga detectors under development will extend photoconductive detection techniques to beyond 200 μm.

Current extrinsic detectors have typical NEPs of about 10^{-17} W $Hz^{-½}$, and have been developed in part with orbital astronomy missions in mind, where they allow background limited operation even in a cryogenically cooled space telescope.

1.2.3 Arrays

Most IR astronomy to date has, of necessity, been done with single element detectors. While not a handicap for photometry, this has been very restrictive for most other forms of observation. Mapping of large areas of sky has usually to be done by scanning the telescope, often a prohibitively time-consuming procedure, and spectrometry has been similarly disadvantaged by the low sensitivities available. The use of large two-dimensional arrays for imagery would, in addition to the considerable sensitivity gain, produce data of higher geometric precision, and the observations would be less vulnerable to fluctuations in sky transparency, since all image points would be

observed simultaneously. These advantages would also apply to spectrometry. Large arrays are now coming into widespread use, bringing about radical changes in the design of imaging and spectrometric instrumentation.

For arrays beyond a few elements in size it is impractical to provide separate electrical connections to each detector element. The design of the device must provide for in situ charge integration and some form of multiplexed output. Several approaches to this problem have been investigated.

Charge-coupled devices and charge injection devices fabricated from silicon have proven their effectiveness as detectors at visible wavelengths. One approach to the development of IR arrays has been to make CCD and CID arrays from narrower bandgap semiconductors such as InSb and HgCdTe.

A charge-coupled device is composed of an array of charge storage capacitors. These capacitors are similar in construction to metal-oxide-semiconductor transistors. That is, an insulating layer of semiconductor oxide is grown onto a doped semiconductor substrate, and a metal electrode, called a gate, is deposited upon that. When a reverse bias voltage is applied to the gate, any free minority carriers, generated thermally or by photon absorption, are attracted to the oxide-semiconductor junction, forming a depletion layer in the semiconductor substrate. The minority carriers can build up until the interface potential has dropped to zero and the whole applied voltage is across the oxide layer, when no further charge can be attracted to the junction. The amount of charge that can be stored at each capacitive site is referred to as the well capacity of the device. By applying appropriate voltage waveforms to a pair of gates, the charge stored in one capacitive cell can be transferred to the adjacent one, and by repeating this sequence, the charge can be transferred all the way along a column of storage cells to an output electrode, where it is amplified and read out. It is important that the efficiency of charge transfer from one cell to the next be high, since typical devices can have columns several hundred pixels long. Inefficient transfer reduces sensitivity and results in smearing of the image, as charge is left behind in cells along each column.

The charge injection device, or CID, stores charge in the same way as a CCD, but differs in the way the charge is transferred for readout. Each image cell consists of a pair of metal-oxide-semiconductor capacitors. The gate electrodes are connected together along rows and columns, with one electrode of each pair being connected to a row line, the other to a column line. During integration, a reverse bias voltage is applied to all the column lines, so that one electrode of each pair will accumulate charge, while the row lines, and hence the second electrode of each image cell, are held at the substrate potential. During readout, the charge is first transferred to the second electrode of each pair by reverse biassing the row lines and grounding the column lines. The column lines are then disconnected and allowed to float. A single row line is then grounded, pushing the charge stored at each image cell in that row onto the floating column electrodes. The potential of each column changes by the value of the charge packet divided by the column capacitance, and these voltages are then fed sequentially

to an output amplifier for readout. This readout sequence is then repeated for the next row.

The CCD and CID differ in their suitability as astronomical detectors. The CID is a much simpler structure to fabricate, and was available first. Since CID charge readout does not involve a lengthy sequence of cell-to-cell transfers, charge transfer efficiency is relatively unimportant, and charge smearing cannot take place. Also, "blooming", an effect where charge from a saturated image cell spills into neighbouring cells, is a less severe problem on a CID, where the cells are better isolated from each other. Furthermore, CID charge readout is non-destructive, which allows integrations to be monitored and multiple readouts to be made to reduce noise. Set against these advantages, however, is the much higher readout noise of the CID. This is because the gate electrodes along each column are connected together, producing a very high line capacitance and hence a small voltage change per unit charge. The CCD by contrast places charge onto a single output electrode for readout. Its superior noise properties have made the CCD the preferred architecture for astronomical array detectors, despite the CID's potential advantages in some other areas.

Monolithic charge injection devices up to 32×32 pixels in size have been made from InSb, and smaller 8×8 pixel arrays have been available from several manufacturers for some time. Such devices are inherently noisy, and, because the noise increases with array size (due to higher line capacitance) other array architectures are being pursued with greater interest. The more complex CCD structure has proved difficult to fabricate in materials other than silicon, and monolithic near-IR CCDs are not yet generally available.

The most successful design approach to date avoids the fabrication problems associated with monolithic devices by separating the detection and readout functions onto two separate substrates. An array of photodiodes is formed on a chip of detector material, and this is joined in a sandwich structure to a second chip, made of silicon, which carries the more complex charge integration and multiplexing circuitry. Electrical connections between the two are made by an array of indium solder bumps placed on one of the chips. Such "bump-bonded" hybrid arrays have been made with a variety of detector materials including InSb, HgCdTe and extrinsic silicon. Several multiplexing schemes are possible. For example the detector photodiodes may be connected to the gate electrodes of a conventional CCD structure, or the photocurrent may be integrated on the gates of an array of individually addressable MOSFETs.

Schottky barrier detectors offer another way to exploit the well understood processing techniques of silicon in the production of IR arrays. A Schottky barrier occurs at the interface between a metal and a semiconductor. It is the difference in potentials between the metal's Fermi level and the bottom of the semiconductor's conduction band. A photon can be detected when it excites an electron from the metal into the conduction band of the semiconductor. These structures are of interest as IR detectors because many metal-silicon or metal-metal silicide junctions have Schottky barriers less than 1 eV high. The PtSi-Si junction has a cut-off wavelength at 5.5 μm, for example. Detector arrays can be made by placing a platinum metallisation layer

onto a silicon CCD. Arrays of up to 100,000 pixels have been made this way, but the detection mechanism has an inherently low quantum efficiency, generally well under 10%, which restricts the usefulness of these devices in astronomy.

A common structure used for extrinsic silicon arrays is the accumulation mode CID, or AMCID. These devices consist of an extrinsic silicon substrate onto which a thin silicon oxide layer and then an array of metal gate electrodes have been deposited. Free charge produced in the substrate is stored under the gates during the integration period. A gate is read out by reversing the potential applied to it, pushing its stored charge through the substrate to a FET.

1.2.4 Bolometers

The cooled germanium bolometer is among the most widely used detectors at mid- and far-IR wavelengths. The gallium-doped germanium used as the detector material has a strongly temperature dependent electrical resistance (it varies as the inverse fourth power of temperature). The heating effect of absorbed radiation causes a reduction in the resistance of the detector element, and the change is measured by applying a bias current through a load resistor.

The bolometer consists of a blackened crystal of gallium-doped germanium connected to a heat sink by metal leads. The thermal conductance of these leads is the most important parameter in the design of the bolometer. It determines the responsivity, the thermal time constant, the maximum background loading and the ultimate noise level of the detector (Low and Rieke, 1974). By careful selection of the blackening material bolometers can be made with a very high absorption coefficient and an almost constant spectral response across the whole IR region. In operation, bolometers are generally cooled to 4 K by liquid helium-4, but an order of magnitude increase in sensitivity can be gained by cooling to 0.3 K with liquid helium-3.

1.3 Photometric Apparatus and Techniques

1.3.1 IR Telescopes

A principal design goal for a dedicated infrared telescope is the reduction of thermal background radiation from the telescope structure seen by the detector. This leads to a requirement for design characteristics different from, and in some cases in conflict with those of a general purpose optical telescope.

An important example of such a conflict arises in the use of sky baffles. Whereas optical telescopes make extensive use of baffles to block sky emission from the focal plane, such highly emissive surfaces close to the optical beam would be one of the largest sources of background radiation reaching an infrared detector. Infrared optical systems use cold baffles within the cryostat in place of warm telescope baffles.

Structural supports, especially those for the secondary mirror, also introduce thermal background flux into the detector field of view, and their cross-section as seen by the detector should be minimised.

Another significant source of background radiation from Cassegrain telescopes is emission reflected back onto the detector from the central hole in the primary. The size of the central obstruction can be reduced by adopting a large f ratio (typically $f/35$). A further reduction in background flux can be achieved by placing a small plane mirror, subtending the same angle on the detector as the hole in the primary, in front of the secondary, so aligned as to reflect sky onto the detector via the primary mirror. The detector then sees sky emission in place of the full black body flux from the central hole.

Thermal radiation from outside the telescope aperture must also be controlled; the defining stop of the optical system should have a low emissivity. The aperture of an optical telescope is usually defined by the primary mirror itself, to make full use of its collecting area. At IR wavelengths, however, this allows room temperature radiation from around the primary to fall onto the detector. This can be blocked by introducing a low emissivity stop elsewhere in the optical system, and is often done by using an undersized secondary mirror. The secondary then defines the effective aperture and the detector views only sky around it. Alternatively, a cold field stop may be placed within the cryostat. While this can reduce emission from the stop by more than the use of an undersized secondary, diffraction effects may be severe. Finally, a cooled baffle in front of the field stop can be used to define the effective aperture. If the field of view is to have sharp edges, the field stop must subtend a negligible angle at the baffle. Since the space available within a cryostat is limited, however, this arrangement is only feasable at small f-ratios.

1.3.2 Signal Detection

In the L band ($\lambda = 3.8\,\mu$m), the sky emission through a 15 arcsec. aperture is 9×10^{-11} W m^{-2}. By comparison, the flux from a zero magnitude astronomical source at this wavelength is only 5×10^{-11} W m^{-2}. At longer wavelengths, thermal emission from the sky and telescope structure exceeds that of any astronomical object. This background emission fluctuates both temporally and spatially, yet must be subtracted to very high accuracy if infrared astronomical observations are to be possible at all.

A beam switching technique is used to do this, a procedure where the detector alternately views the source and an adjacent area of sky, with the difference in signals being measured.

To filter out detector $1/f$ noise and the large low frequency components of excess sky noise, the signal must be modulated at a frequency of over 10 Hz or so. Since the telescope as a whole can't be wobbled at so high a rate, the modulation is usually accomplished by wobbling either the secondary mirror or a small mirror placed near the focal plane.

Since only one element of the optical train oscillates in such an arrangement, light from the two beams follows slightly different paths through the optical system. Therefore the telescope emissivity will not be identical in both beams. Also, if a gradient exists in sky emission, sky background levels will differ too. These effects give rise to a "chopping signal", a spurious instrument induced difference in signal levels

between the two beams. The chopping signal can be corrected for by "nodding" the telescope. This involves moving the telescope as a whole every minute or so to switch the source from one beam to the other. Taking the difference in signal levels at these two positions cancels the gradient in sky background to first order, and also cancels the signal caused by the inhomogenous emissivity of the telescope optics.

The chopper-modulated signal can be measured either asynchronously, i.e. as a power detection, or synchronously, so that the signal is rectified in phase with the chopper reference waveform. Both approaches are equivalent in terms of sensitivity at high signal-to-noise ratios, but at low S/N ratios there is a threshold below which a signal cannot be detected asynchronously irrespective of integration time, whereas no such threshold exists for a synchronous detection.

1.3.3 Photometric Systems

Infrared photometric systems were first developed by H.L. Johnson, who took advantage of the emergence of infrared photoconductive and bolometric detectors to extend his UBVRI system to include bands at 1.25 μm, 2.2 μm, 3.4 μm, 5.0 μm and 10.2 μm, labelled J, K, L, M and N respectively, chosen to match windows of high atmospheric transmission (Johnson 1962, 1964, 1965, 1966). Stellar colours in this system are best defined in the J, K and L bands, by the photometric measurements of a large number of bright stars published by Johnson et al. (1966). Subsequent systems added the 1.6 μm H band, and the introduction of InSb detectors meant that the definition of the L passband no longer had to be restricted by the 3.8 μm cut-off of PbS detectors. This led to the adoption of a modified L filter, the L' centred at 3.8 μm, which avoided atmospheric absorption bands shortward of 3.5 μm. Such JHKL'M systems are in widespread use today.

Photometric systems used by different groups differ slightly in filter transmission characteristics, and as a consequence stellar colours in each of these natural systems also differ. As the accuracy of photometric measurements has increased, it has become more important to determine colour transformations between systems, and to define a standard system. Early attempts to homogenize stellar data from several natural systems were made by Koornneef (1983), and more recent efforts are discussed by Glass (1985).

The very transparent 10 μm atmospheric window extends almost a full octave from 8 μm to 14 μm. The standard N filter, centred on 10.2 μm, is over 5 μm wide to take advantage of this. Many applications need better defined effective wavelengths or lower background levels, and most photometers for this region also offer narrower filters that subdivide the 10 μm window into up to three bands centred around 8 μm, 10 μm and 12 μm, with widths between 1 μm and 1.5 μm.

Beyond 10 μm, a magnitude system for the Q (20 μm) band was developed by Simon et al. (1972), and recently reevaluated by Tokunaga (1984). Ground-based techniques have been augumented by the results of the IRAS survey. The IRAS Point Source Catalogue is now a major source of photometric data for the 10 μm to 100 μm region.

1.4 Infrared Instruments

1.4.1 CVF Spectrometer

Circular Variable Filters are tunable interference filters where film thickness, and hence wavelength, varies linearly with angular position on a circular substrate. CVFs typically offer wavelength ratios between 1.33:1 and 2:1 over rotation angles of 90° to 180°, and have peak transmittances between 60% and 80%. Spectral resolutions between 50 and 100 are possible. It is straightforward to incorporate such a filter into the filter wheel of a conventional photometer, giving the photometer a low resolution spectrometric capability. The optical simplicity and high efficiency of the CVF spectrometer have brought it into widespread use for low resolution applications, particularly in the 1 μm to 5 μm region.

1.4.2 Grating Spectrometer

The grating spectrometer is one of the most widely used spectrometric instruments in astronomy. Spectral multiplexing with detector arrays is straightforward, and the limited throughput is rarely a disadvantage, since many astronomical sources are not spatially extended.

The grating equation is

$$a (\sin\theta_m - \sin\theta_i) = m\lambda \qquad (1.4.1)$$

where a is the inter-ruling spacing, θ_i and θ_m are the angles of incidence and reflection respectively, λ is the wavelength and m is the spectral order.

This expression may be differentiated to obtain the angular dispersion, D,

$$D = d\theta/d\lambda = m/a \cos\theta_m \qquad (1.4.2)$$

The resolution , R, is

$$R = Na (\sin\theta_m - \sin\theta_i) / \lambda \qquad (1.4.3)$$

where N is the number of rulings on the grating.

Most of the light incident on a conventional grating, on which the groove faces are ruled parallel to the grating plane, undergoes specular reflection into the zeroth order, where it is wasted for spectroscopic purposes. But if the grooves are ruled in such a way that their faces are at an angle to the plane of the grating, to produce a "blazed" grating, then specular reflection off the face of each groove will not be at the zero order angle, and incident radiation will be diffracted preferentially into one of the higher orders, allowing more efficient use to be made of the incoming energy.

One of the commonest spectrometer configurations is the Littrow mounting, where the propagation vectors for the incident and dispersed light are on the same side of the normal to the grating plane, so that $\theta_i = -\theta_m$, and, when the radiation is incident at the blaze angle, the groove faces are normal to the incident light, a condition known as "autocollimation". This configuration gives the highest resolution,

$$R = 2Na \sin(\theta_i) / \lambda \qquad (1.4.4)$$

Such a spectrometer produces a one-dimensional image format, suitable for linear detector arrays. Several optical configurations can make efficient use of two-dimensional array detectors. The long slit spectrometer uses a long slit to project the image of a thin strip of sky onto the detector. With the dispersion axis perpendicular to the slit, the spectrally dispersed image of the slit lies on one axis of the 2-D array, and the image of the slit itself on the other, thereby exploiting the one-dimensional imaging potential of the grating spectrometer.

A cross-dispersed echelle spectrometer can produce a two-dimensional image format when spectral information alone is required. An echelle is a grating produced with a large blaze angle and widely separated rulings so that it can be operated at large angles of incidence (60° - 70°) and high spectral orders (around 100th order) to achieve very high spectral resolutions ($\lambda/\Delta\lambda \sim 10^4$-$10^5$). Orders can be separated by including a second grating or prism in the optical path with a dispersion axis perpendicular to that of the echelle. The resulting cross-dispersed spectrum has a format well suited to two-dimensional detector arrays.

1.4.3 Fabry-Perot Spectrometer

A Fabry-Perot interferometer consists of two plane parallel partially transparent reflecting surfaces separated by a small distance. Transmission through such an etalon is high when the distance separating the plates is an integral multiple of the wavelength. The transmission T is given by

$$T = (1+4r(1-r)^{-2} \sin^2(2\pi nl\nu \cos\theta))^{-1} \qquad (1.4.5)$$

where r is the reflectance of the etalon surfaces, n is the refractive index of the inter-flat material, l is the etalon plate separation, θ is the angle of incidence and ν is the frequency of the incident light.

The interferometer is scanned in wavelength by varying the plate separation, although scanning by varying n or θ is also possible.

The resolving power R is given by

$$R = 2F\nu nl = Fm \qquad (1.4.6)$$

where m is the order number and F is a parameter known as the reflective finesse, and has the value

$$F = \pi r^{1/2}/(1-r) \qquad (1.4.7)$$

The free spectral range is

$$\Delta\nu_{fsr} = c/2nl \qquad (1.4.8)$$

where c is the velocity of light.

That is, increasing the plate separation has the effect of increasing the resolving power but reducing the free spectral range. The ratio between resolution and free spectral range is determined by the finesse.

A Fabry-Perot interferometer can have a sensitivity advantage over a grating instrument under certain circumstances. Whereas a grating instrument must view the source through a narrow slit, a Fabry-Perot interferometer can admit light from a large solid angle without loss of resolution. It therefore has a higher sensitivity when observing extended sources, an effect known as the throughput, or Jacquinot advantage.

A common configuration for a near-IR Fabry-Perot interferometer consists of an uncooled piezoelectrically scanned F-P etalon placed in front of a CVF spectrometer. The cooled CVF filter serves as the order sorter for the instrument. At high resolutions, where the CVF bandwidth may exceed the free spectral range of the F-P etalon, two etalons in tandem can be used.

1.4.4 Fourier Transform Spectrometer

A Michelson interferometer produces an interferogram by splitting an incident wavefront into two beams, then recombining the two after introducing a variable path difference between them.

The resulting fringe pattern has an intensity I(x) at one output given by

$$I(x) = \int T(\nu)S(\nu)d\nu + \int T(\nu)S(\nu)\cos(2\pi\nu x)\, d\nu$$

$$= I(0) + \int T(\nu)S(\nu)\cos(2\pi\nu x)\, d\nu \quad (1.4.9)$$

where $T(\nu)$ is the transmission efficiency of the optical system, $S(\nu)$ is the spectral intensity, ν is the frequency expressed as a wavenumber and x is the path difference.

The transmitted spectral intensity of the source $T(\nu)S(\nu)$ can be obtained by calculating the Fourier Transform of the interferogram. The Michelson interferometer can therefore serve as a spectrometer, and, under certain circumstances, it has a higher sensitivity than other forms of spectrometric instrument.

In common with the Fabry-Perot interferometer the FTS has a throughput (Jacquinot) advantage over a grating instrument. In the detector noise limited case, however, there is also a multiplex advantage (Fellgett, 1958). This is a sensitivity gain which arises because, in contrast to a scanned grating instrument where spectral elements are measured sequentially, an FTS encodes the whole spectral range onto the detector simultaneously. This results in a signal-to-noise improvement proportional to $N^{1/2}$, where N is the number of spectral elements observed.

The resolution of the instrument depends on the length of the scan and is given by

$$\Delta\nu(\text{min}) = 1/(2N\Delta x) = 1/D \quad (1.4.10)$$

where N is the number of samples from zero path difference, Δx is the path difference between samples and D is the total scan length beyond zero path difference.

The highest frequency at which data can be obtained is determined by the sampling interval,

$$\nu(\text{max}) = 1/2\Delta x \qquad (1.4.11)$$

where ν(max) is the maximum frequency expressed as a wavenumber. This is the Nyquist criterion, and an aliassed spectrum is present above this frequency.

Because of the symmetrical nature of the Michelson interferometer's optical configuration, the input beam can serve as a second output, and vice versa. The output from the second beam is complementary to the first, and has the form

$$I(x) = I(0) - \int T(\nu)S(\nu)\cos(2\pi\nu x)\, d\nu \qquad (1.4.12)$$

The interferogram from a single-input, single output FTS contains a d.c. offset term (I(0) in the equation above) which is proportional to the integrated photometric signal from the source. This term introduces noise since it is subject to fluctuations due to seeing and atmospheric scintillation. The fringe amplitude also includes a contribution from the thermal background. This can greatly exceed the signal from the source and must be corrected for. Both these terms can be removed by adopting a dual input beam, dual output beam configuration for the interferometer. This arrangement exploits the complementary nature of the two beams. The d.c. offset term can be cancelled by taking the difference between the two outputs, and the background can be subtracted by arranging the two input beams to be adjacent to each other on the sky, with the source in one beam and sky alone in the other. The background contribution is then approximately cancelled at each output.

Scintillation effects are avoided by scanning the interferometer at a rate high enough for fringe frequencies to exceed scintillation frequencies.

Several near-IR Fourier Transform Spectrometers, with resolving powers of over 10^5, have been built for use at the coude foci of ground-based telescopes, notably at Kitt Peak Observatory and the Canada-France-Hawaii Telescope. Many mid-IR FTS instruments have also been operated on airborne telescopes. Ridgeway and Brault (1984) summarise some of the astrophysical results obtained. FTS spectrometers based on lamellar grating interferometers, where the wavefront itself, rather than the amplitude, is divided, and those based on polarising Martin-Puplett interferometers have also been used.

1.4.5 Heterodyne Spectrometer

The heterodyne receiver is a coherent detector which uses a square-law receiver, or "mixer", to combine the signal to be detected with a "local oscillator" reference field. The local oscillator frequency is set to be close to that of the signal. The receiver produces an output signal at the difference frequency of the two combined inputs. This "intermediate frequency" is much lower than the frequencies of the two inputs, and it is this signal that is amplified and measured.

The signal and local oscillator waves entering the receiver produce a resultant electric field vector E(t) given by

$$E(t) = E(s)\cos(\omega_s t) + E(LO)\cos(\omega_{LO} t) \quad (1.4.13)$$

where E(s) and E(LO) are the electric field strengths of the signal and local oscillator waves, and ω_s and ω_{LO} are their frequencies.

The response, *i*, of the receiver is proportional to the intensity of the radiation or the square of the electric field

$$\begin{aligned} i \propto E(t)^2 &= E(s)^2 \cos^2(\omega_s t) + E(LO)^2 \cos^2(\omega_{LO} t) \\ &+ E(s)E(LO)\cos((\omega_s - \omega_{LO})t) \\ &+ E(s)E(LO)\cos((\omega_s + \omega_{LO})t) \quad (1.4.14) \end{aligned}$$

The signal and local oscillator frequencies are close to each other, so the intermediate frequency term $\omega_{IF} = \omega_s - \omega_{LO} \ll \omega_s$. The receiver cannot follow the instantaneous intensity variation at frequencies above ω_{IF}, and will respond to the average values of the three high frequency terms in the expression above.

The receiver response is therefore given by

$$\begin{aligned} i &= k\,(E(s)^2/2 + E(LO)^2/2 + E(s)E(LO)\cos((\omega_s - \omega_{LO})t)) \\ &= i(dc) + i(IF) \quad (1.4.15) \end{aligned}$$

where k is a proportionality constant containing the quantum efficiency.

For a reverse-biased diode receiver, optimum sensitivity is obtained when shot noise due to the local oscillator is the the dominant noise source. Under these conditions the minimum detectable signal is given by

$$P(min) = h\nu\,\Delta\nu/\eta \quad (1.4.16)$$

where $\Delta\nu$ is the bandwidth and η is the quantum efficiency. Note that, when $\eta = 1$, this power corresponds to a flux rate of one photon per $1/\Delta\nu$ seconds, that is, one photon per resolution time of the system. While heterodyne detection is in widespread use at radio and microwave frequencies, its application to infrared astronomy has been restricted because the very narrow bandwidths (~GHz) result in low sensitivities, and because of the lack of tunable local oscillators.

Heterodyne detection at mid-infrared wavelengths was pioneered by Teich (1968). A significant modern astronomical application is in long baseline interferometry, (Townes, 1984), where the use of a narrow band local oscillator in place of direct beam combination can ease mechanical stability problems.

In a typical IR heterodyne system, a high bandwidth CdHgTe photodiode serves as the detector and mixer. A 10.6 μm CO_2 laser is generally used as the local oscillator.

1.4.6 Polarimeter

Scattering and dichroic absorption by aligned interstellar dust grains can polarise an incident radiation field. Polarimetry at near infrared wavelengths can be

used to obtain information about the nature of the interstellar environment in regions of high dust concentration.

A near-infrared polarimeter typically consists of a rotating half-wave plate and a wire grid analyser (e.g. Bailey and Hough, 1982) used in conjunction with a conventional IR photometer. Instruments with both warm and cooled analysing grids are in use. The inclusion of a Circular Variable Filter within the photometer allows spectropolarimetric measurements to be made. The signal may be modulated by sky chopping, or by rotation of the half-wave plate itself.

1.4.7 High Resolution Techniques

Labeyrie (1970) showed that speckle interferometry, at visible wavelengths, can be used to recover information from an astronomical image degraded by atmospheric turbulence, at spatial frequencies up to the diffraction limit of the telescope. The technique was later extended to infrared wavelengths (Selby et al., 1979, Sibille et al., 1979) where the lack of large area detectors requires a different observing technique to be adopted; the stellar image is scanned mechanically across a single element detector.

By applying a sawtooth reference waveform to the telescope's chopping mirror, the image is swept repeatedly across a slit whose width is matched to the diffraction limit of the telescope. The period of the sweep is chosen to be less than the atmospheric coherence time, typically 50 ms, so that the speckle pattern remains fixed during a single scan. The speckle power spectrum has a non-zero value up to the telescope cut-off frequency λ/D, where D is the telescope diameter. By ratioing the object power spectrum with that of a reference point source, the power spectrum of the object can be obtained. Resolutions of the order of 0.1", and limiting magnitudes of $m_K = 7$ are achievable with present-day large telescopes.

The introduction of array detectors will make possible two-dimensional speckle observations of much higher sensitivity than current one-dimensional measurements can achieve.

Measurements made in the pupil plane can have some advantages over image plane speckle techniques. Sensitivities are higher in the pupil plane for visibility (but generally not phase) measurements (Chelli and Mariotti, 1986), and the transfer function is independent of seeing, thereby avoiding the calibration difficulties that beset IR speckle observations. A pupil plane rotation shearing interferometer (Mariotti et al., 1987) built to develop the technique consists of a modified Michelson interferometer with orthogonal roof prism reflectors. These rotate the two interferometer beams by 180° relative to one another before recombination. Fringes formed on a detector array give information about the spatial frequency spectrum of the source.

The lunar occultation technique can achieve angular resolutions approaching one milliarcsecond. The angular diameter of an object can be obtained by measuring the diffraction pattern produced when the lunar limb passes in front of it. A spatially extended source produces fringes of lower contrast than a point source. While this

technique offers the highest spatial resolutions currently possible, its application is severely restricted by the infrequent occurrence of suitable occultation events.

By far the largest programme of infrared lunar occultation measurements has been carried at Kitt Peak Observatory by Ridgeway and his co-workers (Ridgeway et al., 1977, Schmidke et al., 1986, and references therein). The angular diameters of over fifty late type giants and supergiants have been determined by observations of lunar occultations at optical and IR wavelengths.

Long baseline interferometry can also potentially reach milliarcsecond resolutions. The visibility of the fringe pattern produced by the interference of beams from two telescopes separated by a distance D yields the power in the object's spatial frequency spectrum at the frequency $f = \lambda/D$.

This technique has been applied successfully at IR wavelengths by di Benedetto and Conti (1983), who measured the angular diameters of five red giant stars at 2.2 μm with the I2T interferometer at CERGA. This instrument uses two 26 cm telescopes mounted on rails, allowing the baseline to be varied between 6 m and 67 m.

An alternative to direct combination of the interferometer beams is the use of a local oscillator as a phase reference. This approach has been taken by Townes (1984), who has built an interferometer for the 10 μm band incorporating heterodyne detection techniques. The instrument consists of two mobile 1.65 m telescopes of very rigid construction, and uses a network of laser interferometers to monitor their mechanical deflection. 10.6 μm CO_2 lasers are used as local oscillators. The laser radiation is combined with energy entering the telescope and the heterodyne detection is made by a fast CdHgTe mixer diode. Since direct interference of the two beams isn't needed, mechanical stability requirements for the interferometer are reduced. Against this must be set the low sensitivity of the instrument, due to the very small bandwidth ($\Delta\lambda/\lambda = 10^{-4}$) inherent in the heterodyne detection technique.

1.5 Surveys

The appearance of the sky varies greatly with wavelength. Few of the brightest sources in any spectral band are present among the brightest sources at wavelengths just two or three octaves away. The large scale sky survey therefore plays a crucial role in the advancement of astronomy in a new spectral region, identifying sources of interest and perhaps discovering new classes of object. Several sky surveys have influenced the evolution of infrared astronomy.

1.5.1 Two Micron Sky Survey

The first large scale infrared survey was the Two Micron Sky Survey (Neugebauer and Leighton, 1969) carried out at CalTech between 1963 and 1969. Three quarters of the sky north of -33° was scanned in the K band, producing a catalogue of over 5600 sources to a limiting magnitude of +3.1.

The survey instrument was a 1.5 m telescope with a linear array of eight cooled PbS photoconductors at its prime focus. The sky was raster scanned in 15 degree strips,

and the signal was modulated by rocking the primary mirror itself. The primary was notable for its method of production. The parabolic optical surface was formed by curing epoxy resin in a rotating dish.

The TMSS remains the most extensive near-IR survey carried out. Several smaller surveys have been done since, though. Recent examples include the Valinhos survey (Epchtein et al. 1987), an extension of the TMSS to the southern hemisphere, and the J, H and K maps of the the Galactic Centre region produced at SAAO (Glass et al., 1987), covering two square degrees at 6×12 arcsec resolution.

1.5.2 Air Force Sky Survey

The sky at mid infrared wavelengths was first surveyed by the Air Force Sky Survey. A 16.5 cm liquid helium cooled telescope, carried on sounding rockets, was launched nine times from sites in the northern and southern hemispheres, accumulating 1800s of observations and covering over 90% of the sky. The instrument used an array of 24 photoconductive detectors measuring at wavelengths of 4.2 μm, 11.0 μm and 19.8 μm. (The 4.2 μm band was replaced by a 27.4 μm one for the southern hemisphere flights.) Each detector viewed a 3.4' × 10.5' field.

Early results were catalogued in The AFCRL Infrared Sky Survey (Walker and Price, 1975). A second version, The AFGL Four Color Infrared Sky Survey (Price and Walker, 1976), contained reanalysed data from the first catalogue and added the results from the last three rocket flights. The third, and final version, the Revised AFGL Infrared Sky Survey Catalog was issued in 1983 (Price and Murdock, 1983). Astrophysical results from the survey have been reviewed by Kleinmann et.al. (1981). The majority of new infrared sources found by the survey are M-type and carbon stars, and objects associated with HII regions and regions of star formation.

1.5.3 IRAS

By far the most extensive mid- and far-IR survey was that carried out by the IRAS satellite in 1983. The spacecraft itself is described later in this chapter. 96% of the sky was surveyed in four spectral bands centred on 12 μm, 25 μm, 60 μm and 100 μm. Several data products from the IRAS mission have been published (Beichman et al., 1984).

The largest of these is the IRAS Point Source Catalog, which contains the 250,000 unresolved objects detected by the satellite. It increases the number of sources observed at infrared wavelengths by over an order of magnitude.

Resolved objects less that 8' across (IRAS resolution ranged from 0.75 × 4.5 arcmin in the short wave channels to 3.0 × 5.0 arcmin at 100 μm) are listed in the Small Extended Sources Catalog. Larger objects, and data for the sky as a whole, are presented as a set of 212 images in the Sky Brightness Images Catalog.

The Extragalactic Sub-Catalog contains sources from the Point Source Catalog and the Small Extended Sources Catalog that are positionally associated with extragalactic objects.

Spectra of those point sources detected by the Low Resolution Spectrometer have also been published (IRAS Science Team, 1986).

1.6 Facilities

1.6.1 Ground-based Telescopes

Several terrestrial telescopes have been built specifically for use at infrared wavelengths, in that they have generally been sited at high altitude, exceptionally dry locations and their designs emphasise low emissivity. They include the 1.5 m Infrared Flux Collector at Izana, Tenerife, the 3.8 m United Kingdom Infrared Telescope and the 3.0 m Infrared telescope Facility on Mauna Kea, Hawaii, and the Wyoming Infrared Telescope.

The design of most modern optical telescopes also makes provision for efficient infrared observation, allowing an undersized chopping secondary mirror to be fitted, and warm baffles to be removed.

1.6.2 Balloons

Even on the highest mountain-top sites, the atmosphere is practically opaque between 30 μm and 300 μm, largely as a consequence of water vapour absorption. The bulk of atmospheric water vapour is confined to altitudes below the tropopause, at 12 km, and the far-IR spectral region can be observed by telescopes above this height.

Helium balloons can operate at altitudes of up to 30 km where, in addition to the lower far-IR opacity, they can take advantage of atmospheric emission 100 times lower than on a mountain-top site.

Many balloon-borne IR telescopes have been operated over the last fifteen years or so (Soifer and Pipher, 1978). Ambient temperature telescopes as large as 1.2 m in diameter, and smaller cryogenically cooled telescopes, have flown. Most have been used for low spatial resolution mapping and sky surveys at far IR wavelengths (up to 300 μm).

1.6.3 Aircraft

Aircraft can also be used as high altitude IR telescope platforms, allowing long duration observations to be carried out at altitudes of up to 15,000 m. The NASA Ames Research Centre has operated a 30 cm telescope on a Lear Jet 24B, and a 32 cm telescope has been flown on a Convair 990 by a group at the Observatoire de Meudon. The largest programme of airborne observations to date has been carried out by the Kuiper Airborne Observatory, an international facility operated by NASA. The KAO is a 0.9 m telescope mounted on a C-141 aircraft. The telescope views the sky through an open port in the top of the fuselage. Its stabilisation system can achieve 1" tracking. A next generation airborne telescope, the Stratospheric Observatory for Infrared Astronomy (SOFIA) is under study by NASA. This 3 metre class telescope would have the sensitivity to detect all the long wave point sources in the IRAS catalog, and would make far-IR spectroscopy of extragalactic objects possible.

1.6.4 Spacecraft

The ideal observing environment is above the atmosphere altogether. Here the background flux is dominated by thermal emission from zodiacal dust (Soifer et al., 1971), and beyond 10 μm this flux is over six orders of magnitude lower than the sky emission at a high altitude ground-based site. Furthermore the vacuum environment allows the telescope as a whole can be cooled to cryogenic temperatures.

These advantages were first exploited by sounding rocket experiments, notably the Air Force Sky Survey, which achieved significant results despite being limited to observations of only a few minutes duration per flight. Orbital spacecraft, on the other hand, can operate continuously for many months, their lifetimes being limited by their supply of cryogenic coolant. Now that mid-IR and far-IR detectors have sensitivities high enough to make good use of the low background conditions above the atmosphere, several orbital IR telescope missions are being undertaken.

a) IRAS

The Infrared Astronomical Satellite was the first cryogenically cooled telescope to be placed in orbit. The 0.57 m aperture telescope, contained in a liquid helium cryostat, conducted an all sky survey in four wavelength bands centred on 12 μm, 25 μm, 60 μm and 100 μm. The spacecraft was launched in early 1983 and operated for ten months before its cryogen supply expired. It surveyed 96% of the sky to completeness levels of 0.5 Jy at 12 μm and 1.5 Jy at 100 μm.

The spacecraft was placed into a 900 km altitude polar orbit. The orbit precessed by about one degree per day to remain close to the plane of the terminator, thereby providing a fixed Sun-Earth geometry throughout the mission. The sky was scanned by rotating the spacecraft about its Sun-facing axis.

Three instruments were carried at the Cassegrain focus of the telescope; the survey array itself, a low resolution spectrometer (LRS) and the chopped photometric channel (CPC).

The survey array was made up of sixty-two detectors, each preceded by a field lens and spectral filter. These were arranged in eight modules, two for each spectral band, in such a way that a source crossing the field of view would be detected by at least two detectors in each band. The short wave detectors, made of Si:As for 12 μm and Si:Sb for 25 μm, each viewed a 0.75×4.5 arcmin. field, while the long wave Ge:Ga detectors had fields of view of 1.5×4.7 arcmin. and 3.0×5.0 arcmin. at 60 μm and 100 μm respectively.

The low resolution spectrometer was a slitless objective prism spectrometer covering the 8 μm to 23 μm wavelength range at resolving powers between 14 and 35. The spectrum was scanned across the detector by the motion of the spacecraft itself. The instrument operated concurrently with the survey array, and acquired a spectrum whenever a sufficiently bright source had passed over the survey detectors.

The chopped photometric channel provided a facility for making measurements of selected sources at higher spatial resolution and higher sensitivity

than was possible with the survey array. Two Ge:Ga detectors viewing 1.2 arcmin. apertures were used for broad band photometric measurements at 50 μm and 100 μm. The provision of an internal chopper allowed the signal to be modulated without scanning, and a flux reference was included for absolute photometric measurements. This instrument was used in a programme of pointed observations carried out in addition to the main survey task.

b) SL2/IRT

The second cryogenically cooled astronomical telescope placed in orbit was the Spacelab 2 Infrared Telescope, flown on the second Spacelab flight. Developed by the Smithsonian Astrophysical Observatory and other institutions, this experiment had a 15 cm diameter primary mirror cooled to 8 K by superfluid helium, and a focal plane detector array derived from the IRAS design.

c) Hubble Space Telescope NICMOS

The HST has warm optics and offers no sensitivity advantage over a ground-based telescope in the thermal infrared. Shortward of 2.5 μm, however, hydroxyl airglow is the dominant noise source for modern ground-based detectors, and there is a sensitivity gain in placing them above the atmosphere. Also, the HST's 2.4 m aperture gives a higher spatial resolution than a ground-based telescope at the shortest IR wavelengths. A near IR instrument has been selected as a HST second generation instrument.

The Near-Infrared Camera and Multi-Object Spectrometer, NICMOS, spans the 0.8 μm to 2.5 μm spectral range using Ge detectors shortward of 1.6 μm and CdHgTe longward. The camera has two 32×32 pixel arrays, one of germanium, the other of cadmium mercury telluride, sharing the focal plane. The multi-object spectrometer uses four linear arrays of 256 Ge detectors, each fed by a fibre-optic probe. Four objects can be observed simultaneously at resolutions between 100 and 10,000 in the 0.8 μm to 1.7 μm region.

The instrument operates at 65 K using solid methane and ammonia cryogens. These will need replenishing after five years on orbit. It is due to be launched in 1994.

d) ISO

The Infrared Space Observatory is ESA's observatory class follow-up to the IRAS survey mission. The 0.6 m liquid helium cooled telescope is due to be launched in 1993 for an eighteen month mission. Although the telescope is similar in diameter to IRAS, ISO will have up to three orders of magnitude greater sensitivity through the use of improved detectors and longer integration times.

A pyramid mirror at the telescope's Cassegrain focus will direct incoming radiation to one of four focal plane instruments. Each instrument views one quadrant of the four arcminute field, and several instruments may observe adjacent areas of sky simultaneously. The instruments' capabilities overlap somewhat for the sake of redundancy.

ISOPHOT, an imaging photopolarimeter, provides the mission with a broad range of polarimetric, spectrophotometric and mapping capabilities. This complex and versatile instrument contains over 200 detectors in four separate subsystems; ISOPHOT-P, photopolarimeter for the 3 μm - 110 μm range, ISOPHOT-S, two spectrophotometers for the 2.5 μm - 12 μm range, ISOPHOT-A, a mapping instrument with three linear arrays sensitive to the 3 μm - 30 μm range, and ISOPHOT-C, a long wave (30 μm - 200 μm) camera using several small photoconductor arrays including a 2×2 array of stressed Ge:Ga detectors for the 200 μm band.

ISOCAM is a camera for the 2.5 μm to 17 μm spectral range. It is the only ISO instrument to use two-dimensional monolithic array detectors. A 32×32 element InSb CID covers the 2.5 μm to 5.5 μm range, while a 32×32 element Si:Ga direct readout array is used for wavelengths between 4 μm and 17 μm. Each array has associated with it a set of wheel-mounted reimaging lenses offering a selection of plate scales, and a filter wheel which includes a CVF. The arrays also share a set of polarising grids.

SWS, the short wave spectrometer, spans the 2.3 μm to 45 μm range at resolutions of over 1000 using two separate grating spectrometers with several small linear detector arrays. A Fabry-Perot interferometer can be inserted into the beam to give resolutions of up to 50,000.

LWS, the long wavelength spectrometer, consists of a grating spectrometer combined with two Fabry-Perot etalons to provide resolving powers between 200 and 10000 in the 45 μm to 180 μm spectral range. It uses a linear array of ten discrete detector elements.

e) COBE

The Cosmic Background Explorer, due for launch in 1989, will perform an all-sky survey of the 3 K cosmic background radiation field. The spacecraft design draws heavily on IRAS experience; the helium dewar is based on the IRAS design, and the spacecraft orbit and surveying strategy are also very similar to those for IRAS. COBE will carry three instruments.

The Differential Microwave Radiometer is mounted outside the cryostat. It will measure anisotropies in the cosmic background in the 31 GHz to 90 GHz range using a differential radiometer. Its dual horn antennae have a 60° separation and 7° field of view.

The other two instruments are mounted within the cryostat and are kept at a 2 K temperature.

The Diffuse Infrared Background Experiment, DIRBE, will search for diffuse emission from the earliest galaxies, and will also detect foreground emission from interplanetary and interstellar dust. This multi-band filter radiometer will have a set of photoconductive and bolometric detectors covering the 1.25 μm to 300 μm spectral

range. It will have a 0.7° square field of view. The instrument has an off-axis Gregorian telescope to meet stringent stray light rejection requirements.

The Far Infrared Absolute Spectrometer, FIRAS, will measure the 3K background spectrum relative to an internal calibration source. The instrument, a polarising Michelson interferometer, will observe the 0.1 mm to 10 mm spectral range at 0.2 cm^{-1} resolution using four bolometric detectors.

f) SIRTF

NASA's as yet unfunded Space Infrared Telescope Facility (Seeley, 1985) is an observatory class IR mission similar in scope to ISO. The 85 cm helium cooled telescope will be among the large orbital facilities designed for Shuttle servicing; on orbit cryogen replenishment will give it a potential lifetime well in excess of five years. This mission follows ISO and will be able to make fuller use of array detectors than ISO could. All three instruments will incorporate large two-dimensional arrays.

The Infrared Array Camera will cover the 2 μm to 30 μm range using arrays of up to 128 × 128 pixels.

A Multiband Imaging Photometer will use small arrays for point source photometry between 2 μm and 200 μm, and will also be capable of wide field imaging in the 200 μm to 700 μm range using a small bolometer array.

The Infrared Spectrograph will offer resolving powers of 100-2000 in the 2 μm to 200 μm spectral range. It is a grating instrument with two-dimensional array detectors.

g) Future Sub-Millimetre Observatories

Two submillimetre heterodyne spectroscopy telescopes, ESA's Far Infrared and Submillimetre Telescope and NASA's Large Deployable Reflector, are under study as future orbital observatories. They will be placed in orbit to avoid the opacity of the earth's atmosphere in this wavelength region. High resolution heterodyne spectroscopy does not require a low background environment, so the collecting optics of both telescopes will operate at ambient temperature. FIRST is to have an 8 m diameter primary mirror, while the proposed LDR would have a 30 m mirror to be assembled in orbit.

Chapter Two
The Four Channel Photometer and the Cooled Grating Spectrometer

Summary

This chapter gives detailed descriptions of two infrared instruments, a multi-channel photometer and a cooled grating spectrometer, discussing their optical and electrical design, data acquisition systems and data reduction procedures.

2.1 Introduction

During 1983 the Imperial College Astronomy Group commissioned two new infrared instruments on the 1.5 m Infrared Flux Collector on Tenerife; a Four-Channel Photometer and a Cooled Grating Spectrometer. These instruments, both designed for use in the $1\,\mu m - 5\,\mu m$ wavelength band, have been in regular operation on this telescope since then, being used most heavily on two long term programmes involving the investigation of carbon star atmospheres and the development of the Infrared Flux Method for determining stellar angular diameters and effective temperatures.

To minimise cost and design effort, the two instruments have many design features in common, (and indeed share the same data acquisition system). They are therefore most conveniently described together in a single chapter.

2.2 The 1.5m Infrared Flux Collector

Both the Four Channel Photometer and the spectrometer were built for use on the 1.5m Infrared Flux Collector. The telescope is designed for operation at infrared wavelengths. It has an undersized secondary mirror, an optical path free of warm baffles, and a location, at Izaña, Tenerife, selected in part for its low atmospheric water vapour content.

The main uses of this telescope involve observing at infrared wavelengths with relatively large sky apertures. Since only moderate image quality is required for these applications, the precision of the telescope's optical figuring need not be as high as that of a general purpose optical telescope. The relaxed optical tolerances allow a lightweight, low cost telescope design to be adopted. A thin primary mirror is mounted in a light supporting structure. The telescope is manually slewed, and stepper motors provide the sidereal drive and the guiding and setting motions.

The telescope motion may be controlled either from a handset or by electrical signals sent from a control computer.

Instrumentation may be mounted at the Cassegrain and Coude foci. The focal ratio at the Cassegrain focus is $f/13.8$, giving a plate scale of $10''\ mm^{-1}$.

2.3 The Imperial College Four-Channel Photometer

The Imperial College Four Channel Photometer measures four photometric channels simultaneously. The photometer has a single sky aperture common to all four detectors. Three of the detectors measure the K, L' and M photometric bands. The fourth measures either the J or H band.

There are several reasons for constructing a photometer of this kind. A multi-channel instrument offers an improvement in observing efficiency, though this must be weighed against the loss of flexibility incurred in terms of available filter combinations. A more important advantage stems from the common entrance aperture for all four detectors. All four signals are affected equally by, for example, telescope guiding errors, and, to a first approximation, by seeing fluctuations. Therefore, photometric colours can be measured to a higher accuracy than is possible with sequential observations by a single-channel instrument.

By scanning the telescope, the photometer can be used to map an area of sky. For this application, the common entrance aperture assures good spatial registration between scans in the four spectral bands. Accurate colour-difference maps can be produced this way, since errors in the telescope motion cannot introduce any misalignment between bands.

The Four Channel Photometer also makes possible multi-wavelength lunar occultation measurements, and introduces new methods of calibrating speckle interferometric data, since the speckle structure viewed by all four detectors during any one scan will have been produced by an identical atmospheric turbulence pattern.

2.3.1 Optical Design

Radiation enters the cryostat through a side-looking ZnSe window. The photometer beam is defined by a cold baffle set in the instrument cover within the cryostat. This beam is brought to a focus at a circular entrance aperture. One of eight apertures between 3" and 30" in diameter can be selected by rotating an aperture wheel. The wheel may also be fitted with a range of slits and grids for speckle interferometric observations.

A sequence of dichroic filters and reimaging mirrors then splits the beam into four spectral bands and brings the separated beams to their respective foci at the detector blocks.

Each detector block, except the J/H unit, consists of an InSb detector preceded by a zinc selenide Fabry lens and an interference filter. Each Fabry lens is focussed for the spectral interval it will operate at, and hence introduces less chromatic aberration than the equivalent lens in a single detector photometer, which must be set to a compromise focus for the whole 1 μm to 5 μm range.

An infrared LED is mounted alongside each detector block in such a way that radiation from it enters the detector block without passing through the spectral filter, so allowing the detector to be "flashed".

The monolithic nature of the detector block assembly reduces the risk of light leakage around the filter.

The J/H detector block differs from this configuration in having, in place of a single fixed filter, four filters mounted on a wheel. Two of these filters are the standard photometric J and H filters, the others are narrow-band J and H filters. The narrow filters are needed to meet a spectral coherence requirement in speckle interferometric observations. (At longer wavelengths the standard photometric filters already satisfy this requirement).

The photometer can be aligned in a straightforward manner by introducing a He-Ne laser beam through the sky aperture along the optical axis. The dichroic filters can then be inserted and aligned one at a time, starting with the last one in the optical chain.

Laboratory beam profile measurements of all four channels are shown in figure 2.1. A 1.5 arcsecond pinhole illuminated by a black body furnace was scanned across a 15 arcsecond photometer aperture. An $f/12$ beam was used to approximate the optical configuration of the IRFC. The scans illustrate the alignment of the four beams.

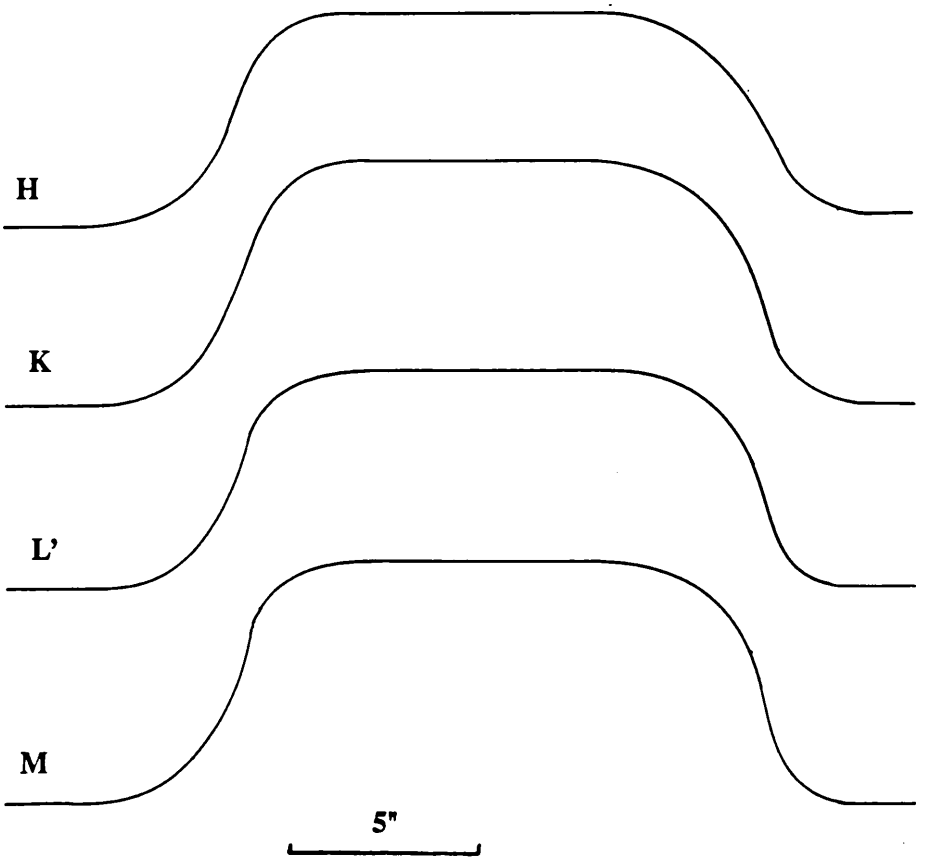


FIGURE 2.1 The Four Channel Photometer Beam Profiles.
A 1.5" pinhole illuminated by a black-body furnace was scanned across the 15" photometer aperture. The scans are on the same horizontal scale, vertical scales are arbitrary.

The photometer is aligned along the optical axis of the telescope by placing a ring of small light bulbs around the circumference of the secondary mirror. The photometer beam is wide enough to accept radiation from these lights, so when the tilt of the photometer has been adjusted to maximise the signal, the detectors are seeing all the lights and the photometer is therefore centred on the secondary mirror.

2.3.2 Mechanical Configuration

The photometer's side-looking cryostat is mounted on a casting at the Cassegrain focus of the IRFC. Also mounted within the casting is a two-mirror focal plane chopper similar to that described by Jorden et al. (1976). One of the the mirrors is semi-transmitting at visible wavelengths, allowing a guide eyepiece to view the unchopped stellar image. A TV guider may also be attached to the casting, and used instead of the eyepiece.

The cryostat outer can is machined from aluminium alloy. Highly reflective Mylar film wrapped around the inner can and around the inner surface of the outer can reduces radiative heat transfer into the cryostat. The photometer itself is mounted on a copper cold surface at the base of the inner can. The photometer's base plate and the mountings for detectors and optical components are made from copper.

The filter wheel and aperture wheel are connected by shafts passing through Wilson seals in the base of the cryostat. Microswitches check when these wheels are in position.

Electrical connections are led into the cryostat via two 19-way Amphenol hermetically sealed connectors, one for the preamp signals and the other for the filter wheel microswitches.

2.3.3 Detectors and Preamps

The detectors used on the Four Channel Photometer are 0.5 mm square InSb photodiodes manufactured by Cincinatti Electronics. Each detector is mounted on a small sapphire disc located precisely at the centre of a mounting disc. The mounting disc is attached to the rear of the detector block, behind the lens and filter.

Each detector is connected to a transimpedance preamplifier circuit placed immediately alongside the detector block (figure 2.2). The preamplifier was designed by Dr. Xin Dian Xiang of Peking University.

When the detector is operated near zero bias potential, noise in the detector circuit is composed almost entirely of the sum of the Johnson noise of the feedback resistor and any excess noise produced by the FETs. The noise properties of the FETs therefore determine the performance of the detector system. The 2N6483 dual JFETs used in the preamp circuit have a nominal noise voltage of $10\,\mathrm{nV\,Hz^{-1/2}}$ at $f = 10$ Hz and T = 300 K, so do not contribute significantly to total system noise. The preamp circuit displays Johnson noise limited performance. The noise voltage has been measured to be $9.1\,\mu\mathrm{V\,Hz^{-1/2}}$ at 78 K with a $10^{10}\,\Omega$ feedback resistor.

Load resistors of $10^{10}\,\Omega$ are used with the J, K and L detectors. A $10^{8}\,\Omega$ load resistor is used with the M filter to avoid saturation by the higher background flux. Since the variation in background level experienced by each detector is small, no provision has been made for switching between load resistors.

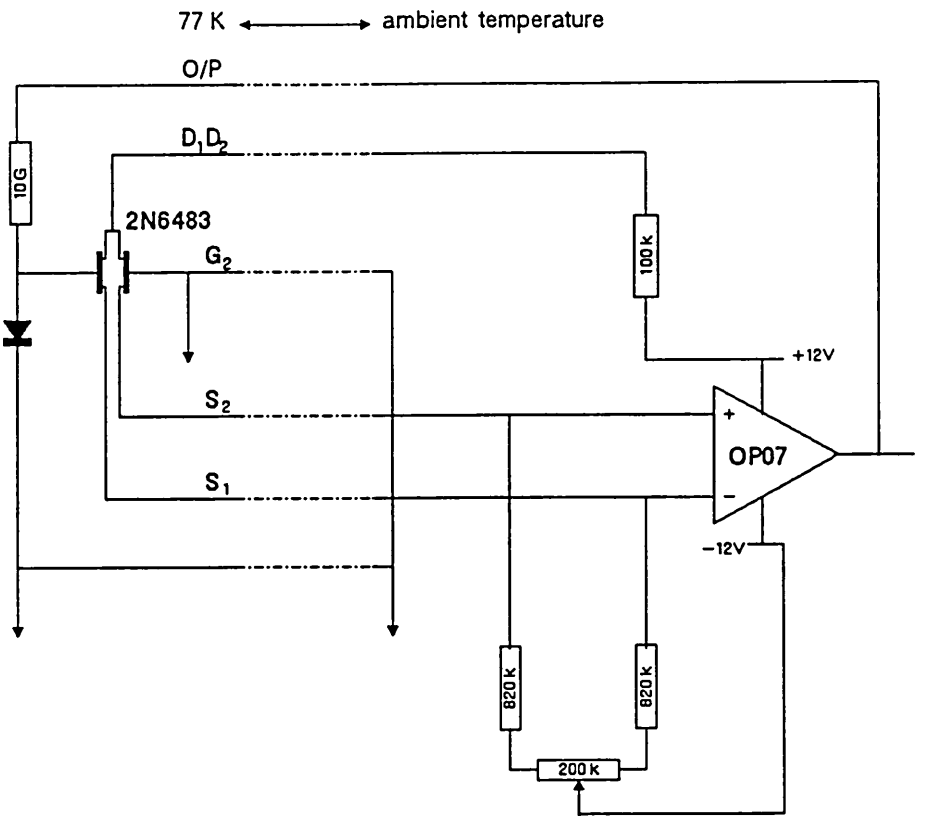


FIGURE 2.2 The Detector Preamplifier Circuit

The two JFETs must have closely matched source voltages, since any offset between them increases the bias potential across the detector. The offset has to be stable to avoid variations in sensitivity, and should be independent of operating temperature.

The circuit design was influenced by the need to eliminate microphonic pick-up. Attention was paid to the stiffness of the preamp structure, flexible leads were avoided and connections between the detector and the circuit board were kept short.

The preamplifiers are powered by a rechargeable lead-acid battery pack to assure a stable supply voltage.

It has long been known that the sensitivity of InSb detectors can be increased by as much as tenfold by illuminating them with intense near-IR radiation. The procedure, known as "flashing", is usually performed by shining a light bulb at the detector through a J filter. This is not possible with the optical configuration of the

Four Channel Photometer, and a more elaborate method has to be used. An infrared LED, operating at 0.9 μm, has been mounted alongside each detector, behind the spectral filter, so that radiation from the LED reaches the detector directly. The detectors are flashed by illuminating these LEDs for about thirty seconds before each night's observing.

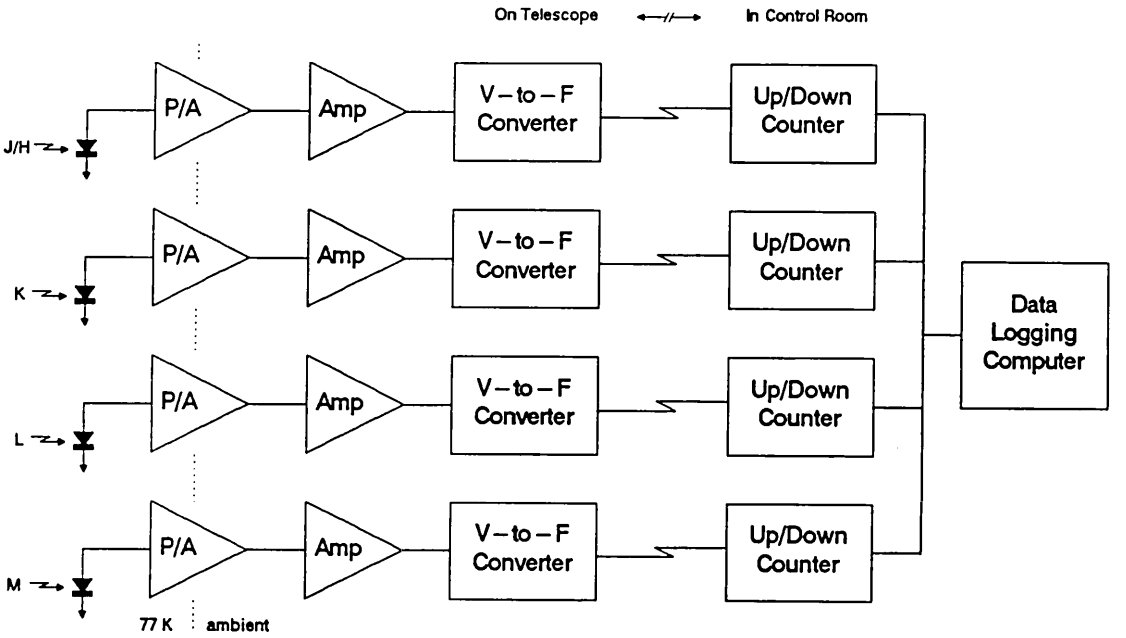


FIGURE 2.3 The Data Acquisition System of the Four Channel Photometer.

2.3.4 Amplifiers and Filters

Four amplifier modules are mounted in the electronics rack, and are connected to the pre-amps. Their gains can be varied in discrete steps between 0.01 and 1000. The amplifiers may be used in d.c. mode, where the frequency response is flat from d.c to 1 kHz, and a.c. mode, where the frequency response cuts off at 10 Hz and 1 kHz. The high frequency cut is chosen to accomodate the highest frequencies encountered in speckle interferometric observations, and to match the frequency response of the detector.

The amplifiers are used in a.c. mode for photometry. For certain other applications, d.c. mapping, speckle interferometry and lunar occultation observations for example, the amplifiers are used d.c. coupled. This can introduce a large d.c. offset if the thermal background is high. The offset must be removed to maintain dynamic range. Since it would be cumbersome to adjust the offsets of the four amplifiers

manually, an automatic facility for removing any unwanted d.c. offset is included in the design. The auto-zero circuit consists of a digital counter linked to the input of a digital-to-analogue converter. The output of the converter is summed with the signal and fed to a comparator. When the "auto-zero" button on an amplifier is pressed, the counter begins to count up from zero, causing the D-to-A output to ramp up and reduce the signal offset. The counter is stopped when the comparator has sensed that the signal voltage has crossed zero, and it is held at that value until the auto-zero function is invoked again. The offset has been found to remain stable to within ±1 mV on a time-scale of hours.

The amplifiers use resistors of 0.1% tolerance throughout. Gain calibrations have found the gains on all the amplifier modules to be within 0.2% of nominal, and stable over a period of eight months, allowing the amplifiers to be used interchangeably without the need for repeated calibrations.

A roll-up filter stage is available after amplification. This introduces an additional gain of 10 at frequencies between about 100 Hz and 1 kHz. It is used mainly for speckle interferometric observations, where the high frequency components of the signal have a much lower amplitude than low frequency components.

2.3.5 Phase Sensitive Detectors

The four phase-sensitive-detection channels consist of voltage-to-frequency converters connected to up/down counters. The output from each amplifier is fed to a V-to-F converter, whose output is a 5 V square wave signal of frequency proportional to the input voltage. The V-to-F converter accepts a 20 V bipolar input range and produces a 100 kHz maximum output frequency. That is, a 0 V input produces a 50 kHz square wave output, a +10 V input produces a 100 kHz output and a -10 V input produces a 0 Hz output. The square wave signal is fed to a digital up/down counter. Counting is synchronised with the chop reference, so that the counters count up for one half of a chop cycle, and down for the other. A phase shift may be introduced between the chop and up/down waveforms to allow for mechanically and electrically induced phase lags in the signal. Although it is possible for the four signal channels to be at slightly different phases, (the M detector has a lower load resistance and hence a different time constant, for example), any such shift would be very small and the same phase shift is applied to all channels to simplify the design and operation of the system. The phase difference between channels proved to be immeasurably small in practice.

The V-to-F converters are mounted alongside the amplifiers in the analogue electronics rack on the telescope. The counters are mounted in the computer rack, which is usually placed in the control room. This configuration allows the signal to be carried the long distance from the telescope to the computer in a form very immune to noise.

2.3.6 AIM 65 Computer and Interfaces

The data acquisition computer for the four-channel photometer was built around a Rockwell AIM 65 microcomputer. This board is based on a 6502 processor

and includes 4 k-bytes of RAM, a current loop serial interface and a parallel port. It also includes EPROMs containing monitor and assembler software and a BASIC interpreter. The control software can therefore be developed on the system that it will run on.

The computer is mounted in a double-height 19-inch case. The upper half of the case contains the AIM 65 board, additional memory cards to expand the RAM to 16 k-bytes and a video controller card. The lower half has a backplane bus into which single-height "eurocard " format cards can be inserted. The bus, called "ASTROBUS", is standard within the Astrophysics Group Electronics Lab. It is based on the 6502 processor bus.

The custom-built cards mounted on this backplane implement a number of interface functions required by the photometer system.

a) Up/Down Counters

The four 24-bit up/down counters for the digital PSD system are carried on two cards in the AIM 65 rack. The counter registers are memory-mapped into the processor's address space, and are read at the end of each integration period. The counters count up and down in alternate chop half cycles, and are briefly disabled in the "dead time" when the chopper is switching. At the end of an integration period they are read out and cleared. The synchronisation signals needed for this are generated not on the counter modules but by counter-timer circuits on the main processor and memory boards. The inputs from the V-to-F converters are fed to the modules' front panels by long cables installed along the telescope structure.

b) Telescope Control Interface

The Telescope Control Interface links the AIM 65 computer to the telescope drive. It allows the telescope to be "nodded" when required by the data acquisition programme. In common with all other parallel input/output ports interfaced to the AIM 65, it is realised using an R6522 Versatile Interface Adaptor. This component provides a straightforward way of implementing memory-mapped I/O for 6500 family processors. The R6522 is connected to the telescope control console through optically isolated line drivers. A nod is commanded by applying a short pulse to the appropriate interface line.

c) Analogue Interface

The Analogue Interface provides four channels of analogue output. It allows the signal levels to be monitored on a chart recorder. It is made from two R6522 VIAs connected to four 8-bit digital-to-analogue converters. The control software selects an 8-bit sequence from the 24-bit up/down counter and places it on the A-to-D input at the end of each integration period.

d) Guide Camera Crosshead Interface

The telescope guide camera is mounted on an X-Y cross-head table driven by stepper motors. The motors can be controlled manually through a handset, but a T.V. Position Interface in the AIM 65 rack allows them to be driven under computer control as well. The interface module provides pulses to the stepper motor control rack to drive the motors, and it can read back the motor position encoders.

e) Serial Communications Interface

Two RS-232 serial communications ports are provided in the AIM 65 rack. One is usually connected to a printer to provide a listing of photometry data. The other is connected to a Cristie CS5 digital cassette recorder. The CS5 receives the same photometry data as the printer, but stores it in a machine readable form, allowing it to be transferred to another computer for analysis at a later date.

f) Bus Monitor Module

A Bus Monitor module is also included. This simply brings several bus signal lines to the front panel for diagnostic purposes.

g) Battery-Backed RAM Module

The final module on the backplane is an 8k-byte battery-backed RAM card. This is used to hold the machine-coded photometer control programme Battery-backed RAM was chosen since it offers permanent memory storage while being easy to modify.

2.3.7 Control Software

The machine-coded control software for the Four Channel Photometer operates the PSD up/down counters and carries out some peripheral interfacing functions.

The interrupt-driven PSD software generates several timing signals needed by the counters. It continues to run in the background even when the high level language photometry programme is present. The control programme receives values for the chop frequency and integration time from the photometry programme and starts the counters running. At the end of an integration period, the programme reads and clears the counters, displays the counter values on the video monitor and on the chart recorder, and leaves these values in memory to be read by the photometry programme. It then begins a new integration.

Several other routines included in the control software are called only by the BASIC photometry programme. These modify the phase of the chop reference signal, move the TV guider crosshead and nod the telescope.

2.3.8 Photometry Programme

The high level language photometry programme, written in BASIC, is stored on audio cassette and is loaded onto the AIM 65 computer at the start of each observing session. The programme integrates the signal from the four detector channels, and provides a user interface allowing the operator to monitor the data as the integration proceeds, and to control the acquisition process. The programme runs interactively in the foreground while the interrupt driven PSD control software continues to run as a background task.

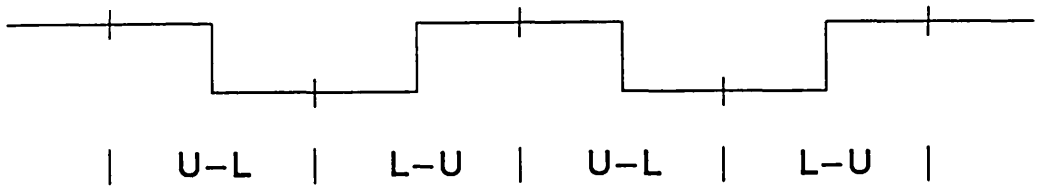


FIGURE 2.4 Photometry Nod Sequence

The photometry algorithm itself provides long term integration of the signals, while compensating for spatial and temporal emission gradients. Whereas nodding the telescope periodically from from one sky beam to the other (here called the upper and lower beams), and subtracting the two signals to obtain an estimate of the stellar flux, does correctly remove the linear gradient in sky emission, it does not treat temporal variations in the sky emission gradient correctly. These might arise when a long integration spans a wide airmass range, or when atmospheric conditions vary rapidly. Such variations introduce a linear drift in the the PSD output, causing the signal difference on a transition from the lower to upper beam to differ from an upper to lower beam transition.

Because of these unequal signal differences, calculating the standard deviation of the combined measurements from both sets of transitions would overestimate the noise. Therefore the mean signal and the standard deviation are calculated separately for each of the two sets of beam differences, the final values being obtained from the mean of the two.

The programme begins by prompting the user for several parameters including the chop frequency, the integration time and the time to be spent in each beam. The PSD phase is also optimised at this point.

An object name, object position and sidereal time are prompted for at the start of each observation. The airmass is calculated from this. The programme then begins integrating the signal. During the observation, the signal and signal-to-noise ratio on each channel are shown on a video display, along with other information relevant to the observation. The intermediate S/N ratios are recalculated and updated on the screen at the end of each nod cycle, and may also be printed out as the integration progresses.

The observation is stopped by a keyboard command. After receiving this command, the programme waits until the end of the next nod sequence, then stops integrating and calculates the final values of the signal, S/N ratio and instrumental magnitude. These values are printed out and stored on a digital cassette. Other keyboard commands allow the observation to be aborted at once, or aborted and automatically restarted.

2.4 The Imperial College IR Spectrometer

2.4.1 Optics

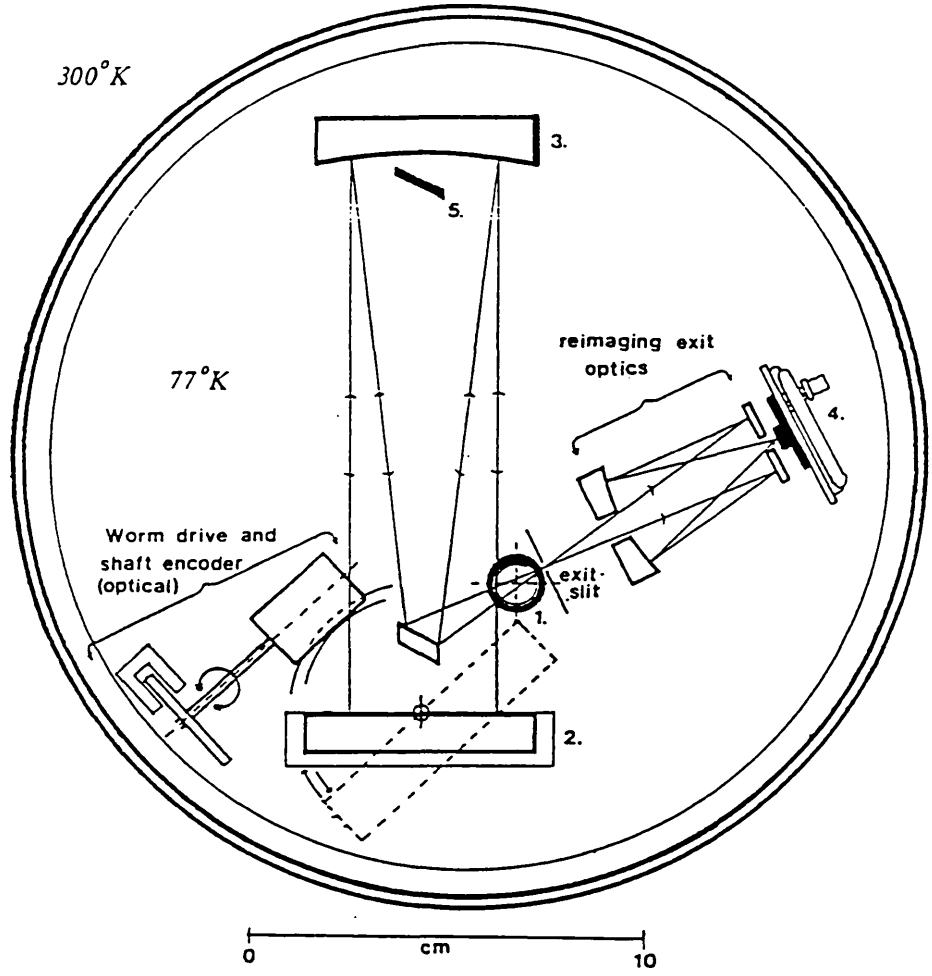
The spectrometer has a Newtonian Littrow mount (Mountain, 1983). Of the possible grating configurations, the Littrow autocollimation mounting has the highest dispersion and throughput. Its reflex optics facilitate a compact design well suited to inclusion within a cryostat. The grating is rotated to scan the spectrum past a single element detector.

The $f/13.8$ telescope beam entering the spectrometer first passes through an order-sorting filter. The seven filters mounted in the filter wheel are the standard photometric J, H, K, L' and M filters, a 1.5 μm cut-on and a 2.5 μm cut-on filter. A blank filter position is also included.

The beam is then converted to $f/4.28$ by a CaF_2 lens within the cryostat. A field lens beyond the focal converter prevents spreading of the beam on the grating. The f/4.28 beam reaches its focus at a 15" circular input aperture, and is then brought onto the spherical collimating mirror via a Newtonian flat. The flat does not reduce throughput since it is positioned in the shadow of the telescope secondary mirror. A cooled stop placed in front of the collimator blocks radiation from the central obstruction.

The collimated beam is dispersed at the grating and is refocussed onto the exit aperture by the collimating mirror. The exit aperture is a 15" wide slit, rather that a circular aperture, to allow some movement of the beam in the direction perpendicular to the dispersion axis. The exit aperture is focussed onto the detector by $f/2$ reimaging optics which are enclosed in a Mylar screen to exclude scattered light.

The grating is a 300 line mm^{-1} replica grating (contracting to 303 line mm^{-1} at 77 K) blazed at 4 μm. Used with a 15" input aperture, it gives a resolution of $\lambda/\Delta\lambda = 600$ at 2.2 μm in second order.



1. Input optics and blocking filters 2. Grating Cell 3. Collimator (spherical)

4. Detector and Pre-amp. 5. Central stop

FIGURE 2.5 The Spectrometer Optics (Mountain, 1983)

The spectrometer is aligned using an $f/13.8$ beam from a He-Ne laser. The position of the beam in the detector plane is observed by exchanging a perspex disc, with the detector position marked on it, for the detector mount. Though the optics remain in fairly good alignment after cooling, provision has been made for the tilt of the collimator mirror to be adjusted from outside the cryostat to correct any small misalignment introduced by cooling.

The optical train of the instrument includes three calcium fluoride elements, seven reflections from gold-coated surfaces and dispersion by the grating, resulting in a transmission of about 43%.

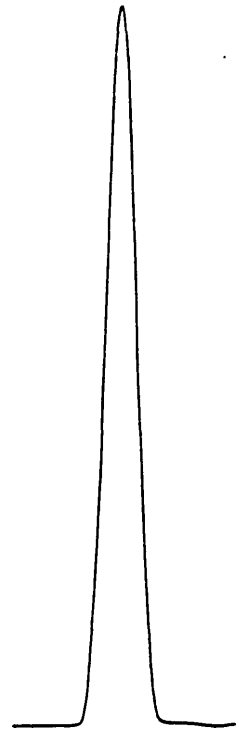


FIGURE 2.6 Spectrometer Line Profile
The triangular profile results from the convolution of the square exit and entrance aperture profiles.

2.4.2 Mechanical Configuration

The spectrometer is contained in a down-looking cryostat mounted in a cage at the Cassegrain focus of the IRFC. Also mounted on the cage are a two-mirror focal plane chopper and a guide eyepiece similar to those used with the Four Channel Photometer.

The cryostat's single inner can is made from stainless steel, and a copper cold surface is brazed to its base. The can is supported by three stainless steel fill tubes. Stainless steel is used in the spectrometer where low thermal conductivity is required. A centering pin in the base of the outer can reduces flexure of the spectrometer assembly within the cryostat. The outer can is made from aluminium alloy. Three Amphenol hermetically sealed connectors bring electrical connections into the cryostat. A 19-way connector carries the encoder signals, a 6-way plug connects the preamp and another 6-way connector is used for signals from the grating limit switches and the filter wheel position switch.

The spectrometer itself is mounted on the copper cold surface and is enclosed by a cold light-tight cover. The spectrometer's base plate and the detector mounting block are made from copper for good thermal contact with the cold surface, the mountings for the other optical components are made from aluminium to save weight.

The small mirrors are glued to their mounts with epoxy adhesive. The collimating mirror and grating are sprung against three point supports in their cells.

The grating cell is rotated with an anti-backlash worm and wheel. The cell is supported by two cross-spring pivots providing accurate rotation centres for the motion of the grating. A flexible bellows coupling links the worm gear to a stainless steel shaft which passes through a Wilson seal in the base of the outer can to a d.c. motor outside the cryostat. Two other shafts are led into the cryostat in a similar way, one for the filter wheel and one for the collimator tilt mechanism.

2.4.3 Detector and Preamps

The detector used with the spectrometer is a Cincinatti Electronics InSb photodiode similar to those used with the Four Channel Photometer. The mounting scheme and the first stage of the preamplifier are also similar to the photometer design described above. Because very low signal levels can be encountered at short wavelengths, the spectrometer's preamplifier includes two additional stages not present in the Four Channel Photometer circuit. The first of these is a d.c. coupled 10× gain stage, and the second is a low pass filter with a 50 Hz cut off. The additional gain is incorporated in the battery powered pre-amp unit, rather than on the mains powered amplifier module, because residual mains ripple on the amplifier power supply can be troublesome at such high gains. The filter is included in case pick-up from the mains or the encoder signal needs to be removed.

Spectrally dispersed radiation from a light bulb outside the spectrometer is not intense enough for "flashing " the detector. Instead, the detector is flashed by a "pea-bulb" placed behind a J filter within the reimaging optics.

2.4.4 Motor Drive

The main difference between the spectrometer's control electronics and those of the Four Channel Photometer is the inclusion of a motor drive and position encoders to control the motion of the grating.

The grating is driven by a 12 V, 3.6 W d.c. servomotor powered by a power amp module in the spectrometer analogue electronics rack. The motor's speed and direction of motion may be controlled from a handset or by the AIM 65 computer. The handset has five buttons (forward, reverse, fast forward, fast reverse and stop) which cause preset voltages to be applied to the input of the power amp. The input can be switched between these handset-selected voltages and the output of an 8-bit D-to-A converter. The converter is fed by a 6522 VIA in the AIM 65 rack and allows the grating to be moved under software control.

Extensive mechanical damage will result if the grating is driven beyond the limits of its motion. Safety features are included in the design of the motor controller to prevent this. Limit switches at the ends of the grating motion generate software interrupts when closed, causing the motor to be stopped. The digital interface is also designed to ensure that the motor will stop if any cable is inadvertently disconnected.

2.4.5 Position Encoders

Two position encoders, a "coarse" encoder and a "fine" encoder, measure the rotational position of the grating.

The resolution of the encoder must greatly exceed the spectral resolution of the spectrometer. The division of one spectrum by another introduces spurious features in highly structured portions of the spectrum if the wavelength registration between the two is imperfect. Any mismatch in wavelength must be about 100 times smaller than the resolution to render this effect negligible (Mountain, 1983). Since there is a flexible coupling between the motor and the worm gear shaft, it would be difficult to meet this requirement with an encoder mounted outside the cryostat. An internally mounted encoder, however, must be capable of operating in vacuum at cryogenic temperatures. An encoder for this purpose was designed within the Astrophysics Group.

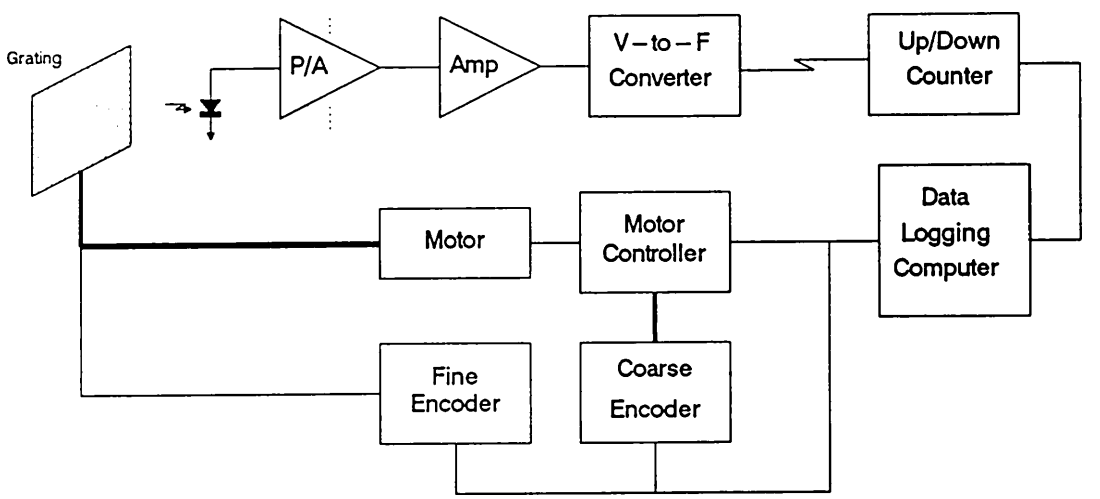


FIGURE 2.7 The Data Acquisition System of the Spectrometer

The encoder consists of two disks of printed circuit board material placed about 1 mm apart. One disk is fixed, the other is mounted on the worm gear shaft and rotates with it. The fixed disk has an array of metallic pads around its perimeter to which a set of sixteen sinusoidal voltages, all having the same frequency but differing in phase, is applied. The second disk also has a pattern of pads and is capacitively coupled

to the first disk. As the worm gear rotates, the phase of the signal picked up by the second disk varies linearly with the angular position of the disk.

The 12 bit encoder measures the position of the worm gear to a precision of 1.3 arcminutes. A spectral line profile at f.w.h.m. is about 500 times wider than the encoder resolution, ensuring that wavelength registration is limited by mechanical imperfections and guiding errors, not by the encoder.

One full cycle of the fine encoder corresponds to a quarter-turn of the worm gear, which in turn rotates about ten times in taking the grating from one end stop to the other. The fine encoder therefore cycles around forty times during a full traverse of the grating. This ambiguity in the encoder reading is resolved by introducing a second, "coarse" encoder to determine which cycle the fine encoder is in.

The "coarse" position encoder is a ten-turn potentiometer mounted outside the cryostat and coupled to the motor. A constant reference voltage in maintained across it with a Zener diode, and the voltage on the wiper is measured by a digital panel meter mounted in the AIM 65 rack. The 3½ digit DPM measures the rotational position with a precision about one hundred times lower than the fine encoder, more than adequate for determining the fine encoder cycle.

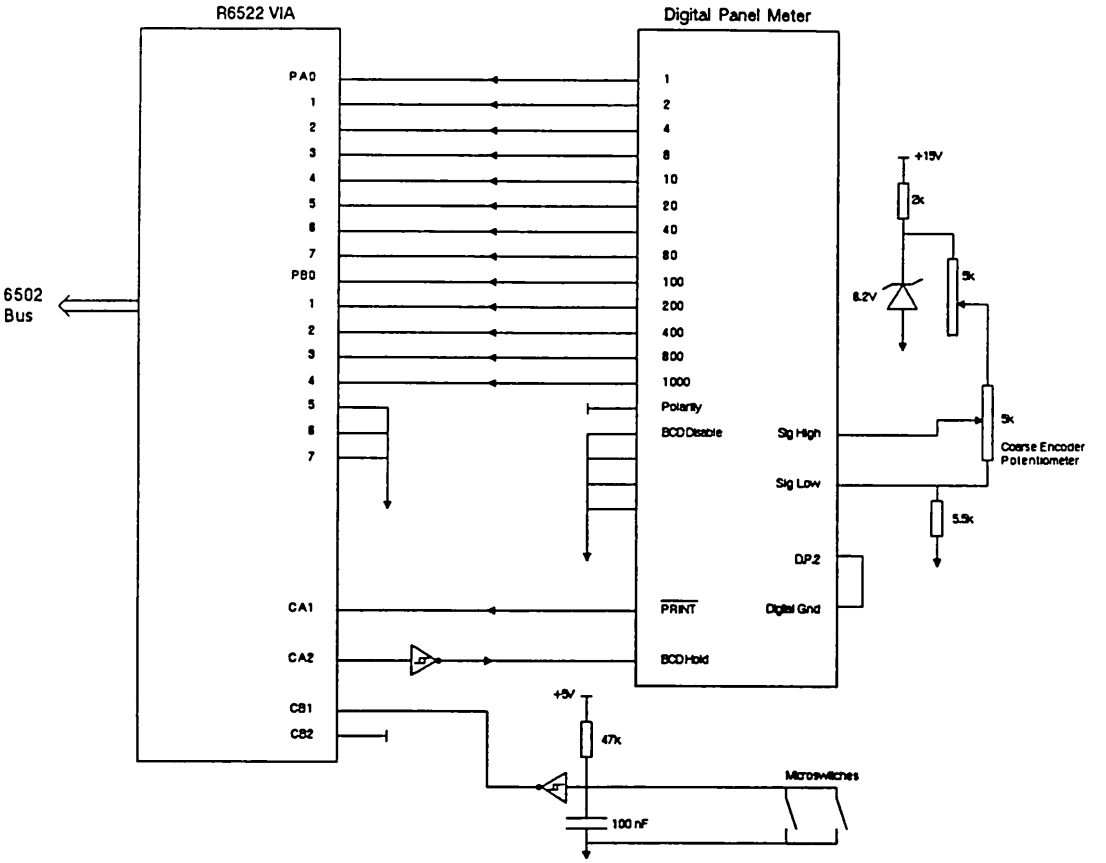


FIGURE 2.8 Coarse Encoder Interface Circuit

2.4.6 Data Acquisition Software

The spectrometer control programme is similar to the Four Channel Photometer's control software in that the main task is the management of the up/down counters and their associated timing signals. The spectrometer uses only one of the four channels, but the software includes additional routines for handling the grating motion. Whereas the photometer control software passes data samples to a high level language photometry programme for on-line processing, the spectrometer's data are stored directly onto a magnetic tape cassette. No on-line processing takes place.

The software waits for a keyboard "start" command, then applies a suitable voltage to the motor to set the grating in motion and begins logging data. At the end of each integration period, the counter and encoder values are read in, the counter value is displayed on the chart recorder and all the data associated with that sample are transmitted along an RS-232 link for storage on the Cristie CS5 data cassette recorder.

Each data sample is stored on tape in the form

XXXX,XXXXXX,XXXX,XXXX <CR LF>

where X represents an ASCII-coded character. The first number is a counter value incremented at the end of every integration period. It serves as a check that no data have been lost, and, since it counts up throughout the night, it is a record of the time of the observation. The second number is the 24 bit PSD counter value, the third number is the fine encoder value and the fourth is the coarse encoder value. Each cassette has a 256 k-byte capacity, enough to store thirty or so spectra.

2.5 Observing and Data Reduction Procedures

2.5.1 Observations

The observations described here were carried out at the 1.5 m IRFC, Tenerife, in June and July of 1985.

Seventeen predominantly main sequence and giant stars, spanning the range of spectral type from B to M, were observed in the J, H, K and L' bands. The ultimate purpose of these observations was to determine the effective temperatures and angular diameters of the stars by application of the Infrared Flux Method (Blackwell et al., 1980). Only the preliminary data reduction, illustrating the general procedure for reducing data from the spectrometer, will be described in this chapter. The observed stars are listed in Table 2.1, and reduced spectra, corrected for atmospheric extinction and instrumental gain, are presented in Appendix B.

The Imperial College spectrometer was used on the IRFC in the configuration described above. The signal was measured on channel number 4 of the Four-Channel Photometer data acquisition electronics. The gains of amplifier unit 4 were calibrated and found to be within 0.2% of nominal, unchanged from previous calibrations.

The chop frequency was 9.3 Hz, chosen to avoid sub-harmonics of the 50 Hz mains frequency. Each sample was integrated for ten chop cycles, giving a 1.08 s

integration time. The grating was driven at a rate chosen to sample a spectral line at between eight and eleven points across its full width, depending on the wavelength. A scan across a spectral band would typically take fifteen minutes to complete.

Star	Sp. Type	m_v	R.A.(2000)	dec.(2000)
BS 15	B8 V	2.1	00^h03^m	+28°32'
BS 5340	K1 III	0.0	14^h16^m	+19°11'
BS 5429	K3 III	3.6	14^h32^m	+30°22'
BS 5531	A2 IV	2.8	14^h50^m	-15°58'
BS 5685	B8 V	2.8	15^h16^m	-9°19'
BS 5854	K2 III	2.7	15^h44^m	+6°26'
BS 6056	M0 III	2.7	16^h14^m	-3°42'
BS 6148	G7 III	2.8	16^h30^m	+21°29'
BS 6337	M3 III	5.0	17^h03^m	+14°05'
BS 6623	G5 IV	3.4	17^h46^m	+27°43'
BS 7001	A0 V	0.0	18^h36^m	+38°47'
BS 7525	K3 III	2.7	19^h46^m	+10°37'
BS 7557	A7 V	0.8	19^h50^m	+8°52'
BS 7754	G8 III	3.6	20^h18^m	-12°33'
BS 7796	F8 I	2.2	20^h22^m	+40°15'
BS 7949	K0 III	2.5	20^h46^m	+33°58'
BS 8781	B9 V	2.5	22^h59^m	+14°40'

TABLE 2.1 The List of Observed Stars

Each star was observed twice, once in each chopper beam, with the grating being driven in the same direction for both scans to avoid any backlash effect on the encoder. Faint stars were scanned several times to improve the signal-to-noise ratio of the measurement.

Wavelength calibration was carried out with a helium spectral lamp. The 1.083 μm helium line was measured in four orders, from second to fifth. Although a calibration was performed at the start of each observing night, it was in fact found that the calibration would remain stable for over a fortnight if the grating was not driven up to an end-stop.

Measurements of the star Vega (BS 7001) provided the flux calibration. This bright, early type star (m_K = 0.0, Sp. = A0V) was observable throughout the night at the time of these observations. A drawback of using a star of such an early spectral type as a flux calibration is the presence of strong hydrogen recombination features in the spectrum. A cooler, G-type dwarf would be more suitable in this respect.

Vega was observed six or so times each night, at airmasses between 1.0 and about 1.8, allowing an atmospheric extinction coefficient to be determined later.

L' band spectra were measured in first order, H and K in second order, and J in third order. The 300 line mm^{-1} grating is not ideally suited to use at J. The filter is

not well matched to the free spectral range in third order, resulting in some overlap between third and fourth orders and hence loss of the extreme long wave portion of J band data. The L' band data were in fact measured through a 2.5 μm cut-on filter, giving a 2.5 μm to 5.0 μm free spectral range. The additional short-wave data, though of little use for this program, were required for a program of carbon star observations being conducted concurrently. Both programs used the same filters to simplify calibration.

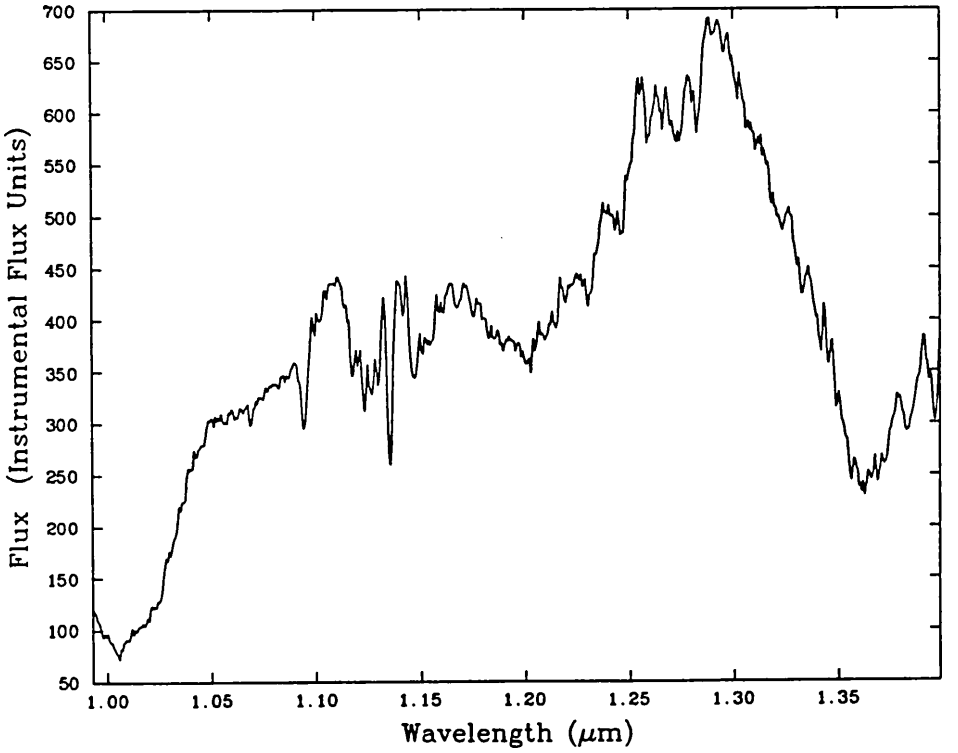


FIGURE 2.9 Spectrum of Vega in the J Band.

Observations were confined to a single spectral band per night. This reduced the time spent on calibrations and avoided the need for frequent refocussing of the telescope. (The spectrometer's beam converter optics contain refractive elements).

The spectrometer was cooled by liquid nitrogen for measurements at L' and M, where the detector is background noise limited. At shorter wavelengths, the nitrogen was pumped solid (to 55 K) to increase the detector impedance and hence improve sensitivity.

The instrument sensitivity was obtained by measuring the flux from Vega at the centre of the filter pass-band and comparing this to the noise level at the same wavelength measured off the source. The one-sigma-one-second limiting magnitudes, that is, the noise equivalent powers expressed as stellar magnitudes, are given in Table 2.2.

Wavelength (μm)	Limiting Magnitude (1σ1s)
1.25	6.1
2.2	5.5
3.8	4.4

TABLE 2.2 Spectrometer Limiting Magnitudes

2.5.2 Data Reduction

The bulk of the data reduction was carried out on Starlink facilities using DIPSO, an interactive spectral arithmetic programme (Howarth and Maslen, 1983).

The data were first transferred, without modification, from the data cassettes they were originally stored on onto ½" magnetic tape, with data from each cassette being stored as a single text file on tape. The data then had to be rearranged into a format suitable for processing by the DIPSO programme.

The encoder values were converted to wavelengths using the helium lamp calibration data, the sidereal time for each observation was obtained from the incremental timer data, and other information about the observation was entered at a terminal. These data were then written to tape in a format readable by DIPSO, each observation being stored as a single file. The first three lines of each file were an information block containing the source name and position, the sidereal time and airmass, the filter used, the amplifier gain and so forth. The rest of the file contained the wavelength and flux values.

a) Wavelength Registration

The first step in the data reduction proper was to ensure good wavelength registration between all spectra. As has already been mentioned, the division of poorly aligned spectra can introduce spurious structure and extra noise into the data. A wavelength shift can be introduced if a source changes its position in the sky aperture. So, if the guiding on one source is at a systematically different position from the guiding on another, the two spectra will be on slightly different wavelength scales.

This effect was investigated and corrected for by applying a cross-correlation technique to the data. A reference spectrum of Vega was chosen for each spectral band, and this was cross-correlated with each object spectrum in turn. (A cross-correlation routine was written and added to DIPSO as a new command).

Before this procedure was performed, however, the spectra were high-pass filtered. This was done because, had the correlation procedure been applied to the unmodified spectra, the cross-correlation function would have been dominated by the contribution from the mean flux levels. The detail of interest would have been difficult to interpret through the underlying triangular background function.

The low frequency signal components were filtered out by convolving each spectrum with a gaussian function ten times broader than the spectral resolution, and

subtracting the resulting smoothed spectrum from the original. Each spectrum was then cross-correlated with its reference spectrum of Vega.

The position of the peak in the cross-correlation function indicated the wavelength displacement of the spectrum, so the wavelength scale was then shifted by this amount to align it with the reference spectrum. Wavelength shifts of up to 0.2 times the resolution were found to occur in the data. The position of the cross-correlation peak could be located precisely (to within 5% of the resolution) by fitting a curve to the upper portion of the central peak. The position determined by this means proved to be insensitive to the shape of the fitted profile, and a gaussian function was fitted to the peak, using DIPSO facilities.

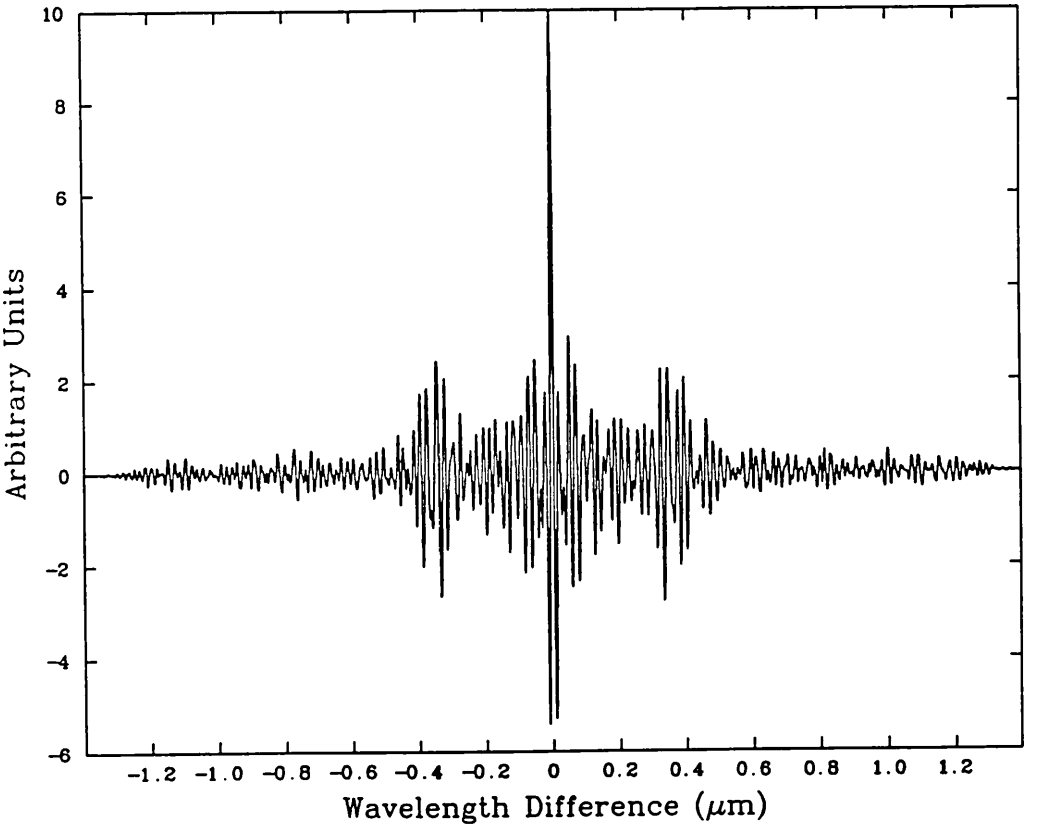


FIGURE 2.10 A Cross-correlated Spectrum.
The cross-correlation of two spectra of Vega in the L' band.

After registration, the spectra were resampled onto a common wavelength grid. The upper and lower beam scans of each pair were then subtracted to remove any d.c. offset, and the spectra of those faint sources for which several scans had been taken consecutively were coadded to improve their S/N ratio.

b) Photometric Accuracy

Photometric accuracy was judged by comparing repeated scans of the same object. On nights when the seeing was relatively good (< 5") agreement between scans was excellent. Except for a few serious guiding errors which were easily

identified, repeated scans agreed to well within the noise. In fact, observations of the brightest stars, such as BS 5340, agreed to within 1% across the whole scan. No evidence was found that, for example, scanning of the grating introduces any drift in optical alignment.

On nights of poorer seeing, however, frequent large discrepancies between scans were found. Among the lowest quality data, fluxes differed by over 10% across more than half of the wavelength interval scanned. These errors, caused by the seeing disk flaring to a diameter greater than the aperture, illustrate the vulnerability of a sequentially scanned spectrometer to poor seeing.

c) Determination of Atmospheric Extinction

The spectra were then corrected for the effects of atmospheric extinction. Each night's set of reference spectra of Vega provided observations across a range of airmass values. Within each set of reference spectra, the logarithm of the flux was plotted against the airmass for each wavelength point. The gradient of the best fitting straight line through these data was taken, after conversion to a stellar magnitude scale, as the extinction coefficient at that wavelength.

This procedure does not treat saturated telluric absorption lines correctly, since extinction does not vary linearly with airmass in these cases.

The observed fluxes of the remaining spectra were then adjusted to their values at one airmass using the extinction coefficients calculated by the procedure described above. They were also scaled to unit instrumental gain.

The determination of the extinction coefficient is susceptible to errors introduced by the form of wavelength misalignment discussed previously. Figure 2.11 illustrates the importance of achieving good registration of the wavelength scales.

The dashed curve in figure 2.11 shows the extinction coefficient as a function of wavelength for a portion of the J band. It was obtained from six scans of Vega taken at airmass values between 1.02 and 1.7. The wavelength registration procedure was not performed. A spurious feature can be seen at 1.093 μm caused by poor registration of the stellar Paschen–γ hydrogen recombination line.

The solid line in figure 2.11 shows the extinction coefficient calculated from the same data after the wavelength scales have been aligned by the cross-correlation technique. The spurious feature is essentially absent.

A method of this sort can succeed only if the source's position in the aperture does not drift significantly during the course of a scan, and if the mechanical precision of the grating drive is high enough to allow full use to be made of the fine encoder's resolution. The results presented here demonstrate that these conditions can be satisfied in practice.

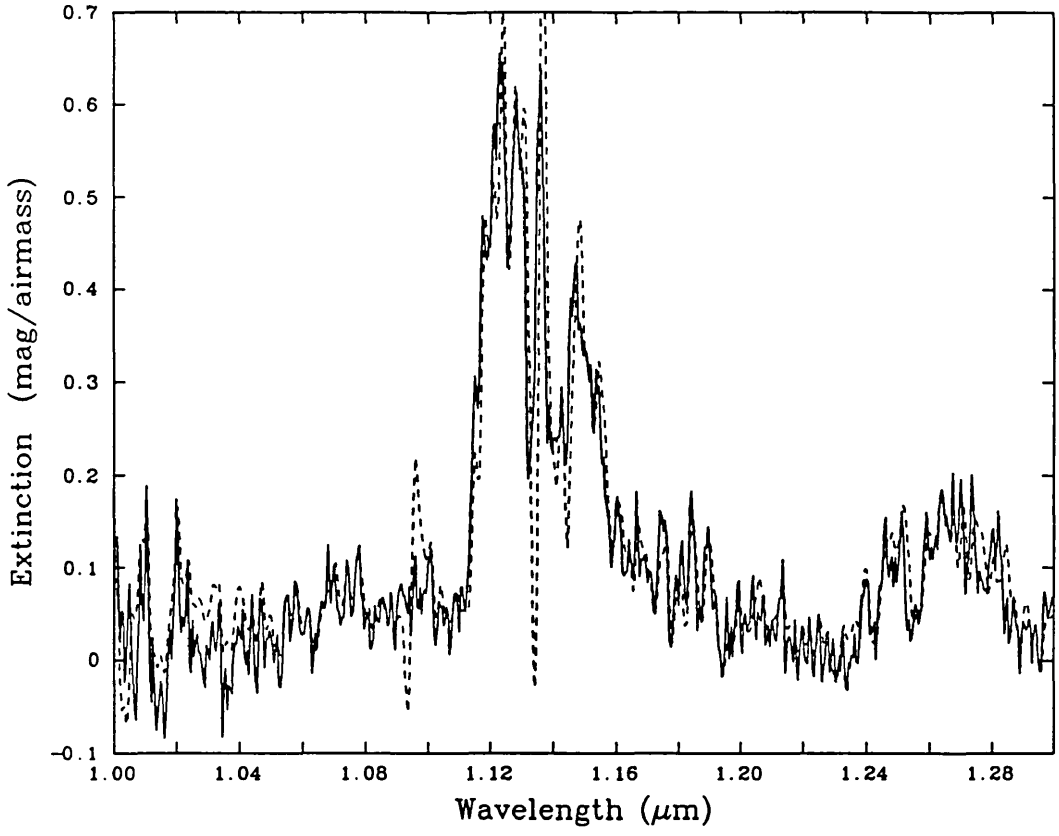


FIGURE 2.11 Extinction Coefficient at 1.093 μm
The solid line represents the extinction coefficient obtained after aligning the spectra, the dashed line represents the extinction coefficient obtained before alignment.

The measured atmospheric transmission was compared with the transmission predicted by a model of the Earth's atmosphere. Atmospheric opacity was modelled using the IRTRANS programme (Traub and Stier, 1976). This programme calculates atmospheric absorption in the infrared from data for over 100,000 molecular transitions of H_2O, CO_2, O_3, O_2, CH_4, CO and N_2O compiled by the Air Force Cambridge Research Laboratory (McClatchey et al., 1973).

Figure 2.12 shows the extinction coefficient, expressed in magnitudes per airmass, measured at J plotted alongside the extinction calculated by the IRTRANS programme. The modelled spectrum includes the contribution from 4 mm of precipitable water vapour, typical for Izaña in June, and has been smoothed to the same resolution as the observed spectrum. There is good qualitative agreement between the two extinction spectra. The measured extinction coefficient is greater than the calculated one in spectral regions where telluric line absorption is expected to be negligible. This is probably because the IRTRANS extinction model does not include the effects of atmospheric dust and aerosols. It seems likely that air-borne dust is the cause of the disparity, particularly as the observations were conducted during a period when exceptionally large amounts of wind-blown Saharan dust were present above the observing site.

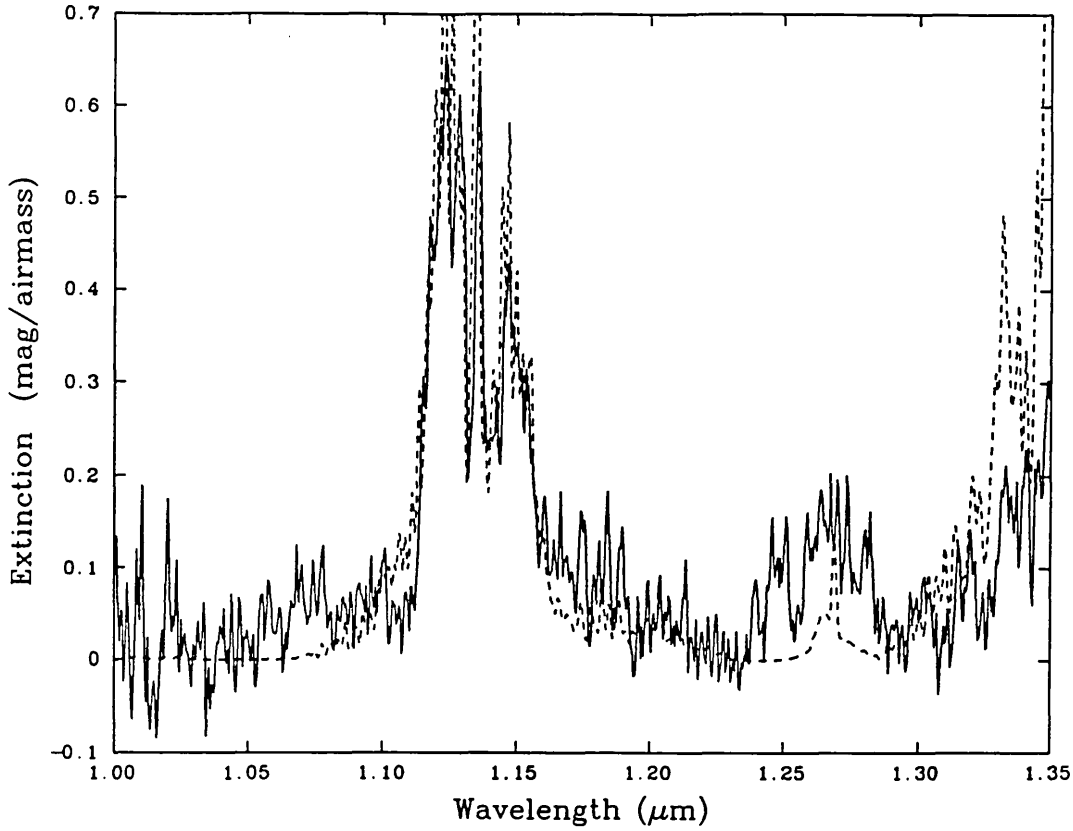


FIGURE 2.12 Observed and Predicted Extinction Coefficient at J.
The solid line is the measured extinction coefficient, the dashed line is extinction coefficient calculated by the IRTRANS model, including the contribution from 4 mm of water vapour.

2.5.3 Conclusion

The Imperial College Cooled Grating Spectrometer was used to obtain spectra of seventeen bright stars. Excellent photometric quality was obtained in periods of good seeing, demonstrating the optical stability of the instrument. Because the instrument scans the spectrum sequentially, however, it is difficult to obtain useful data in periods of poor seeing.

Wavelength registration may vary between scans because of inconsistent guiding. A cross-correlation technique can correct for this very effectively where the source has not drifted across the aperture during the scan.

Chapter Three

A Search for the HD Molecule in the Orion Molecular Cloud

Summary

This chapter describes an attempt to detect a near-IR rotational-vibrational transition of the HD molecule towards the Orion Molecular Cloud.

An upper limit of

$$I_{R(5)} < 2 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$$

for the intensity of the v=1-0 R(5) line of HD at the Peak 1 position in OMC-1 is obtained.

Following a discussion of the background to this program, an account is given of the data reduction procedures used to arrive at the upper limit, and the significance of this result is commented upon.

3.1 Introduction

Studies of interstellar deuterium have provided insights into a number of topics of astrophysical interest. Perhaps most importantly, the isotope is of considerable cosmological significance. No contemporary astrophysical production mechanisms are known for it (deuterium formed in stellar interiors is further processed into helium), so it seems likely that most observable deuterium is primeval. Furthermore, it turns out, from modelling of primordial nucleosynthetic processes, that the amount of deuterium thus formed depends strongly on the baryon density during formation. That is, measurements of cosmic deuterium abundance can provide information not only about the history and physical conditions of the material within which it is observed, but also about the present density of the universe. In view of its importance, then, it is unfortunate that the deuterium abundance has proven difficult to measure reliably, simply because of its low value.

Various observational techniques have been pursued, ranging in wavelength from unsuccessful attempts to detect the 91.6 cm hyperfine line (analogous to the well-known 21 cm hydrogen transition), to UV measurements of Lyman transitions in absorption against nearby hot stars. These latter data, obtained with the Copernicus satellite, have provided the most reliable estimates yet of interstellar deuterium abundances, at least in the local ISM, but even they have recently been found to suffer from serious interpretational difficulties.

By far the largest number of observations, though, have been made at millimetre wavelengths of the rotational transitions of deuterated molecules in interstellar clouds. While these have certainly often yielded abundances for the molecules concerned, it is difficult to derive deuterium abundances themselves from such measurements, since chemical fractionation effects can greatly enhance the relative abundances of deuterated species in these environments, by amounts that depend strongly on the poorly understood physical conditions within the clouds.

In fact, despite the attention given to them, complex molecules hold only a very small fraction of the deuterium present in molecular clouds. Practically all of it is believed to be bound to hydrogen as HD. This molecule serves as the deuterium source in isotope exchange reactions involving larger molecules, and its abundance is not subject to fractionation effects. At first sight though, hopes of detecting it seem slight, due to its low abundance relative to H_2 ($N_{HD} \sim 10^{-5} N_{H2}$). But the abundance ratio itself is not a good indicator of the expected emission ratio because, fortuitously, HD has an electric dipole moment, albeit small, caused by non-adiabatic coupling effects between nuclear and electronic motions, which results in spontaneous emission probabilities over one thousand times greater than the equivalent H_2 quadrupole rates. That is, near IR line emission from HD will be only about one hundred times lower than the H_2 emission from a thermally excited region of interstellar gas.

Until quite recently, though, even this would still have been far too low to be detectable, but infrared spectroscopic instrumentation has improved in sensitivity in recent years to a level where a realistic attempt can now be made to observe such emission from HD.

Measurements of this kind can open up a new observational channel for studies of cosmic deuterium, and produce significant new information unobtainable in any other way. This chapter describes the first attempts at such observations.

3.2 The Cosmological Significance of Deuterium

As already mentioned, observations of cosmic deuterium can throw light not only on local conditions where it is observed, but also on some bulk properties of the universe as a whole, and it is this aspect that has become one of the principal motivations for studying it. This section will examine those topics where considerations of deuterium abundance have a bearing on cosmological issues.

To begin with it is worth recalling the main nuclear reactions that produce and destroy deuterons.

The main production mechanism within stellar interiors is

$$p + p \rightarrow D + e^+ + \nu_e \qquad (3.2.1)$$

This proceeds via the weak interaction and is slow. The deuterons so produced are promptly destroyed by the much faster process

$$p + D \rightarrow {}^3He + \gamma \qquad (3.2.2)$$

This reaction will of course consume any deuterium that was already present prior to the star's formation, with the result that the deuterium abundance in astrated interstellar matter will be reduced from its primeval value. Stellar processing destroys deuterium, it is not a source of deuterium in the interstellar medium.

These two reactions form the main path by which deuterium is processed within stars.

Potential reactions for forming deuterium outside stars include

$$p + n \rightarrow D + \gamma \qquad (3.2.3)$$

(requires neutrons, but no threshold energy)

and, by spallation,

$$p + {}^4He \rightarrow D + {}^3He \qquad (3.2.4)$$

(requires 18.35 MeV threshold energy)

and for destroying it one has

$$D + D \rightarrow n + {}^3He \qquad (3.2.5)$$

or

$$D + D \rightarrow p + T \qquad (3.2.6)$$

and

$$n + D \rightarrow T + \gamma \qquad (3.2.7)$$

Having seen the relevant nuclear processes, possible astrophysical production mechanisms can now be investigated. The reactions above impose quite severe constraints on the physical conditions needed to produce deuterium. Both of the synthesis reactions, for instance, require high energies; to produce the supply of free neutrons in the first case, and to drive the very endothermic spallation reaction in the second. Also, low densities are needed to avoid the three (energetically favourable) destruction processes.

Excluding stellar production, which is not considered a viable way of introducing deuterium into the interstellar medium, two other mechanisms have been studied in detail as possible contemporary deuterium production processes. These two, bombardment of interstellar gas by cosmic rays, (Reeves, Fowler and Hoyle, 1970) and production within supernova shocks (e.g. Colgate, 1974, Hoyle and Fowler, 1973) both satisfy the constraints described above.

Both mechanisms involve spallation reactions. The cosmic ray scheme has protons in the interstellar medium being hit by energetic heavy nuclei, which thereby break up into a variety of lighter species, deuterons amongst them. The second proposal notes that supernova shocks create temperatures high enough to induce nuclear reactions in the ISM, including the proton-helium spallation process mentioned earlier. As with the cosmic ray scheme, numerous other light nuclei are produced as well. Measurements of the abundances of these other nucleides feature prominently in analyses of the viability of both mechanisms.

3.2.1 Cosmic Ray Spallation

Taking the cosmic ray proposal first, it is empirically on very firm ground. Both the cosmic ray flux and the relevant spallation cross-sections have been measured directly, and good estimates for the cosmic abundances of the light elements have also been obtained.

A comparison of observed elemental abundances with the abundances that cosmic ray spallation would produce over the lifetime of the galaxy gives very good agreement between observed and calculated abundances for several species, notably ^{6}Li, ^{9}Be, ^{10}B and ^{11}B, implying that this may indeed be the mechanism by which they are produced (Reeves et al., 1970). On the other hand, the spallation model grossly under-estimates, by one or more orders of magnitude, the abundances of the remaining light nuclei ^{3}H, ^{4}H and ^{7}Li, and, in particular, ^{2}H. In fact, the process can generate only about 1% of the observed deuterium abundance. (Observations of cosmic deuterium are discussed in a later section.)

The reaction rates for these processes are well known, and indeed the four good agreements obtained strengthen confidence in the calculations. The conclusion can be reached with some confidence, then, that the cosmic ray bombardment process is simply incapable of producing the large observed quantities of the latter four isotopes, and other mechanisms must be sought.

3.2.2 Spallation in Supernova Shocks

Is nucleosynthesis in supernova shocks a possibility?

Supernova shocks are associated with temperatures high enough to disrupt atomic nuclei in the interstellar medium. The energy transferred to this matter varies as the square of the shock velocity, and velocities of over 0.13c are needed to split ^{4}He and produce deuterium. Detailed calculations have been carried out to determine the quantities of light nuclei, including deuterium, that will be produced at various shock energies (e.g. Colgate, 1974). These show that, while the production of most species varies in a complicated way with shock energy, that of deuterium increases monotonically, such that the greatest amounts of deuterium will be produced in the most energetic shocks.

Next, the amount of matter exposed to a given shock velocity must be considered. As the shock propagates outwards through the stellar envelope, the gas density decreases while the shock velocity rises. This induces higher deuterium production rates, but in ever smaller quantities of matter. So the total deuterium production predicted for such a shock depends strongly on the finer detail of the supernova model employed. Nevertheless some crude upper limits can be arrived at in a straightforward manner. By making the gross assumptions that all the helium exposed to the shock is converted to deuterium, and that none of this is then destroyed in D+D reactions, or subsequently within stars, this process can account for about 10% of observed deuterium, assuming a galactic supernova rate of one per thirty years. However, both this rate and the fraction of interstellar material exposed to sufficiently energetic shocks are so uncertain that the mechanism cannot be excluded on those grounds alone. More information is available though. The shock-induced reactions, as well as producing deuterium, will produce many other light nuclei , such as ^{7}Li by

$$^4He + ^4He \rightarrow ^7Li + ^1H \qquad (3.2.8)$$

for example, and their production rates will be affected by the above-mentioned parameters too. It should be instructive, therefore, to examine ratios of deuterium relative to those species, as such relative abundances should be less sensitive to variations in supernova rates and energies.

Taking ^7Li as an example again, the shock models predict a $^7Li/^2H$ production ratio of 5×10^{-2}. This is now independent of supernova occurrence rates, though not of shock energies. This value greatly exceeds the observed abundance ratio of 3×10^{-4}. Subsequent destruction of lithium in stars cannot account for this disparity, as this process would burn deuterium faster still, due to its lower Coulomb barrier.

Similarly severe overestimates of production are obtained for the beryllium and boron isotopes. That is, while supernova shocks could produce the high observed cosmic deuterium abundances, to do so they would also have to produce quantities of other light nuclei very much higher than are observed. These light isotope abundances can be used to place a better upper limit on deuterium production in shocks, and they imply that less than 1% of observed deuterium can have been produced this way. This result is no fairly independent of the uncertainties discussed above, allowing such shocks to be excluded as a major source of cosmic deuterium.

In summary, then, it seems that neither spallation by cosmic rays nor within supernova shocks can produce the observed levels of deuterium, and, since no other contemporary production mechanism is known, this conclusion leads naturally to an examination of the possible cosmological origin of deuterium.

3.2.3 Cosmological Processes

There is now a widespread consensus that the universe has expanded from a high temperature, high density state. It is thought that nucleosynthesis in this early epoch is responsible for the production of many of the light isotopes, including deuterium, observable in the present interstellar medium (Wagoner, 1972).

At an age $t = 10^{-23}$s, the size of the communicable universe exceeds the diameter of a nucleon. The interacting particles present in it are the protons, neutrons, leptons, their anti-particles and the photons. These remain in thermal equilibrium until 10^{-6}s. At this point particle energies drop below the rest-mass energy of the proton, and the reverse processes in equilibrium reactions such as

$$p + p^- \simeq 2\gamma \qquad (3.2.9)$$

can no longer take place. The baryons and antibaryons therefore annihilate, leaving only the residual imbalance between the two. The analogous event for electrons occurs at one second, when the reactions

$$e^- + e^+ \simeq 2\gamma \qquad (3.2.10)$$

and

$$e^- + e^+ \rightleftharpoons \nu_e + \overline{\nu}_e \qquad (3.2.11)$$

can no longer proceed in the reverse direction. The electron neutrinos decouple and expand freely from this point.

Also at about this time, the temperature has dropped far enough to allow more complex particles to form, and the era of nucleosynthesis has begun.

It is worth bearing in mind that the deuterium forming reaction

$$n + p \rightleftharpoons {}^2H + \gamma + 2.3\,MeV \qquad (3.2.12)$$

has been taking place since the formation of nucleons from quarks at the $t=10^{-23}s$ point, but the high temperature and photon flux have kept deuterium levels negligibly low until now.

Two other key reactions must also be considered now. They are the weak interactions

$$n + \nu_e \rightleftharpoons p + e^- \qquad (3.2.13)$$

and

$$n + e^+ \rightleftharpoons p + \overline{\nu}_e \qquad (3.2.14)$$

These have kept the proton/neutron ratio in thermal equilibrium at a level given by

$$n/p = \exp-(m_n-m_p)c^2/kT \qquad (3.2.15)$$

and it would seem from this that the neutron abundance will drop indefinitely as the expansion continues. This does not happen for the following reason.

The process is mediated by the weak interaction, and as a consequence the reaction rate falls with the fifth power of temperature (which itself varies as $t^{-\frac{1}{2}}$ in this cosmological model). Therefore as the expansion progresses these reactions slow down, until a time is reached beyond which each nucleon will undergo less than one further interaction over the remaining life of the universe, and the process can be said to have stopped. This happens at a temperature of 10^{10} K, one second into the expansion, and freezes the neutron/proton ratio at a value of 0.22.

Also at this time, the temperature is low enough to stop the reverse reaction in the deuterium production process (3.2.12), which proceeds to completion, absorbing all free neutrons. In fact, some neutrons decay before they have time to react, and the final n/p ratio in nuclei is down to 0.14, from its freeze out level of 0.22.

The nucleosynthesis process does not stop at deuterium, however. Further reactions are energetically favourable, namely

$${}^2H + p \rightarrow {}^3He + 5.5\,MeV \qquad (3.2.16)$$

$${}^3He + {}^2H \rightarrow {}^4He + p + 18\,MeV \qquad (3.2.17)$$

and almost all the deuterium is quickly processed into ^4He.

These reactions stop at the $t=100$ s point, when particle energies drop below the levels needed to overcome their Coulomb barriers.

There are two important points to draw from this account. Firstly, the final helium abundance depends only on the original number of free neutrons available, and this in turn is a function of the temperature at which the n/p ratio "froze out" of equilibrium. There is no density dependence.

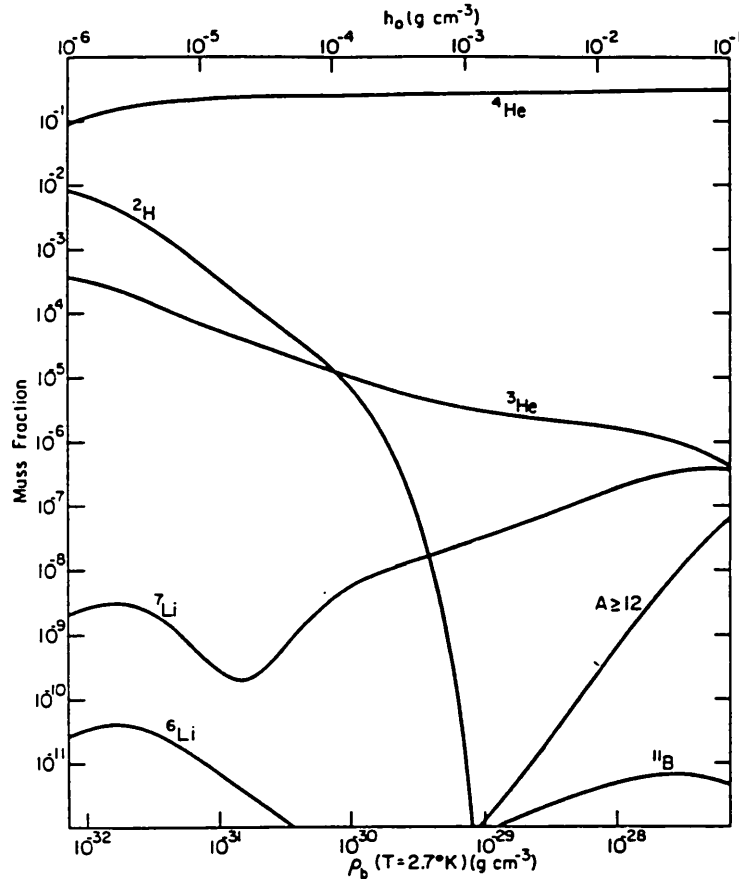


FIGURE 3.1 Elemental Abundances From Big-Bang Nucleosynthesis (Wagoner, 1973). Final abundances produced by standard big-bang models whose only parameter is the present baryon density.

The final ^{2}H and ^{3}He abundances, on the other hand, are critically dependent on baryon density. Given enough time, these isotopes would be completely consumed by the helium-forming reactions. What fraction survives depends on how far these reactions can proceed before they are stopped by the Coulomb barrier effect, and the rates of these processes varies with the density of the reacting baryons.

At high densities, the reactions will go to completion before being frozen, leaving no deuterium or helium-3 at all, and, conversely, lower densities will leave higher residual amounts of these isotopes. It is for this reason, then, that the primeval deuterium abundance is thought to be a sensitive indicator of overall baryon density.

Following the cessation of nucleosynthetic activity at one hundred seconds, the universe continues its expansion with these nuclei, eletrons and photons being the remaining interacting particles in thermal equilibrium. Now, as the universe expands, the matter energy density falls as the cube of the scale factor, whereas the photon

energy density drops as the fourth power, since the photons are being red-shifted. This leads to the photon energy density falling below that of matter at the 2700 year point, giving a matter dominated universe from that time on. The next major event occurs at the 3×10^5 year point, when the temperature has dropped to 3000 K. This allows hydrogen atoms to form, leaving no charged particles. The photons therefore decouple and expand freely henceforth.

All particle species have now decoupled. As the universe expands further, inhomogeneities develop, leading to the formation of galaxies and stars to the present day.

The most important observational evidence for this account is the 3 K cosmic background flux, interpreted as the red-shifted relic of the radiation that decoupled when atomic hydrogen formed.

Of more relevance here is the observed helium abundance, 30% by mass, which agrees well with the predicted value. This abundance depends on the ratio between the cosmic expansion rate and the weak reaction rate, and would be different in most non-standard cosmological models.

As for deuterium, observationally based calculations of its primordial abundance give a mass fraction of 2×10^{-5}, implying an upper limit for the present baryon density of 6×10^{-31} g cm^{-3}, consistent, to within the errors involved, with the observed density of 3×10^{-31} g cm^{-3} based on the virial theorem for galactic masses.

A consideration of the production mechanisms tends to favour a primorial origin for deuterium over a galactic one. This argument, though, has been based solely on theoretical estimates of deuterium production rates via the processes considered. But an empirical test to distinguish between primordial and galactic origins may also be possible. Ostriker and Tinsley (1975) have pointed out that a measurement of the deuterium abundance gradient radially across the galaxy would be such a test. If interstellar deuterium is produced within stars, then its abundance would be higher in astrated interstellar material. On the other hand, deuterium of primordial origin would have a decreased abundance in processed matter.

The degree of stellar processing that interstellar gas has undergone can be independently determined from its metal abundance, and this is known to decrease outwards from the centre of the galaxy. That is, central parts of the galaxy contain more heavily processed material. Therefore, measuring the deuterium abundance gradient across the galaxy would show whether deuterium concentration increases or decreases with the deree of stellar processing, and hence whether or not it has been produced within stars.

Such a measurement has indeed been attempted. Penzias (1979) has measured the ratios of $D^{12}CN$ to $H^{13}CN$, and DCO^+ to $HC^{18}O^+$ in molecular clouds across the galaxy, obtaining results very tentatively supporting a primordial origin for deuterium. However, the interpretation of these results is beset by difficulties due to chemical fractionation processes, and no reliable conclusions about deuterium abundances can be reached from the data.

It is the eventual possibility of carrying out this type of experiment with less chemistry-dependent measurements that is one of the motivations for pursuing the observational strategy described in this chapter.

3.3 Interstellar Chemistry of Deuterium

3.3.1 Deuterium Fractionation

The large relative difference in the masses of the hydrogen and deuterium atoms produces significant differences in their chemical properties. For example, the dissociation energy of the H_2 molecule is 4.478 eV, whereas that of HD is 4.514 eV (Huber and Herzberg, 1979), a difference $\Delta E/k$ of 414 K. Since many dense molecular clouds are at temperatures of this order or lower, this energy difference can lead to a substantial enhancement in the abundances of deuterated molecular species in such environments, a process known as chemical fractionation.

Most deuterium in molecular clouds is thought to be bound as HD, because most hydrogen there is also found in molecular form.

Unlike H_2, which is formed on grain surfaces, HD is formed in the gas phase by the two exothermic ion exchange reactions

$$H^+ + D \rightleftharpoons H + D^+ \qquad (3.3.1)$$

$$D^+ + H_2 \rightleftharpoons HD + H^+ \qquad (3.3.2)$$

where the H^+ has been produced by cosmic ray photodissociation of H_2.

The corresponding reaction between neutrals

$$D + H_2 \rightleftharpoons HD + H \qquad (3.3.3)$$

is thought to be less important as it has an energy barrier, whereas reaction 3.3.2 does not.

Evidence for the dominant role of gas phase reactions in the formation of HD comes from observations made by the "Copernicus" UV satellite of D and HD in diffuse clouds. These observations found high HD/H_2 abundance ratios ($[HD]/[H_2] \sim 10^{-6}$) despite a much higher destruction rate for HD by photodissociation. (Self-shielding of H_2 reduces its photodissociation rate to about 10^{-4} that of HD). Taken with the overall atomic [D]/[H] abundance ratio of 10^{-5} observed in the ISM, this implies a production rate for HD that is 10^3 times higher than that for H_2. Since the process that forms H_2 at grain surfaces is believed to be very efficient (models predict that one third of collisions between H atoms and grains result in the formation of a H_2 molecule) the high HD production rates cannot be accounted for by formation at grain surfaces.

Deuterium enhancement has been observed in several molecules in the ISM, the main ones being DCN, DNC, DCO^+, NH_2D, N_2D^+, CH_3OD and HDO. These can be produced by several routes.

The most significant reaction in the fractionation process, in terms of the number of molecular species subsequently affected, is probably the deuteration of H_3^+ (Watson, 1976),

$$H_3^+ + HD = H_2D^+ + H_2 + \Delta E \qquad (3.3.4)$$

where $\Delta E/k$ has been estimated to be between 90 K and 180 K. The H_2D^+ is removed in part by dissociative electron recombination

$$H_2D^+ + e^- = H + H + D \qquad (3.3.5)$$

and in part by reaction with neutral species

$$\begin{aligned} H_2D^+ + X &= H_2 + DX^+ \\ &= HD + HX^+ \end{aligned} \qquad (3.3.6)$$

where X can represent species such as CO, O, C, N or N_2. The deuterium enhancement can therefore be passed on to a wide range of other molecules via this route. In particular this mechanism can produce the observed species DCO^+ and N_2D^+.

The H_2D^+ produced by reaction 3.3.4 might also account in part for the presence of DCN, HDO and NH_2D via

$$\begin{aligned} H_2D^+ + HCN &= HDCN^+ + H_2 \\ HDCN^+ + e^- &= DCN + H \end{aligned} \qquad (3.3.7)$$

$$\begin{aligned} H_2D^+ + H_2O &= H_2DO^+ + H_2 \\ H_2DO^+ + e^- &= HDO + H \end{aligned} \qquad (3.3.8)$$

$$\begin{aligned} H_2D^+ + NH_3 &= NH_3D^+ + H_2 \\ NH_3D^+ + e^- &= NH_2D + H \end{aligned} \qquad (3.3.9)$$

One difficulty with this route is that the value of ΔE for reaction 3.3.4 is somewhat too low to easily explain the large degree of fractionation observed in the warmer molecular clouds, (Watson and Walmsley, 1982).

Another potential source of deuterium enhancement is

$$CH_3^+ + HD = CH_2D^+ + H_2 + \Delta E \qquad (3.3.10)$$

where $\Delta E/k$ is about 370K. This can produce the observed molecule DCN via

$$\begin{aligned} CH_2D^+ + N &= HDCN^+ + H \\ HDCN^+ + e^- &= DCN + H \end{aligned} \qquad (3.3.11)$$

or by

$$\begin{aligned} CH_2D^+ + e^- &= CHD + H \\ CHD + N &= DCN + H \end{aligned} \qquad (3.3.12)$$

Reaction 3.3.10 also provides another route for producing DCO^+,

$$CH_2D^+ + O \rightleftharpoons DCO^+ + H_2$$

and DCO^+ can also be formed directly by

$$HCO^+ + HD \rightleftharpoons DCO^+ + H_2 \qquad (3.3.13)$$

Of the three routes, reaction 3.3.6 is thought to be the dominant HCO^+ forming mechanism.

Dalgarno and Lepp (1984) have pointed out that atomic D may have a role in deuterium enhancement after all, since the reaction

$$DCO^+ + e^- \rightleftharpoons D + CO \qquad (3.3.14)$$

has the effect of increasing the atomic [D]/[H] ratio in the cloud over the cosmic abundance ratio, because the $[DCO^+]/[HCO^+]$ ratio has already been enhanced by fractionation.

This increases the significance of the reactions

$$D + H_3^+ \rightleftharpoons H_2D^+ + H \qquad (3.3.15)$$

and

$$D + HCO^+ \rightleftharpoons H + DCO^+ \qquad (3.3.16)$$

in the deuterium enhancement process. These reactions had previously been discounted because of the low steady state [H] and [D] concentrations expected from cosmic ray photodissociation alone.

3.3.2 The Spectrum of HD

The asymmetrical nature of the HD molecule causes its spectrum to differ markedly from that of H_2.

In common with all homonuclear diatomic molecules, dipole radiation from the H_2 molecule is very heavily suppressed. Dipole transitions are forbidden by three separate selection rules. The H_2 molecule has no permanent dipole moment, and selection rules forbidding the combination of electronic states of the same parity, and the combination of symmetric and antisymmetric rotational levels also apply. H_2 can only radiate through weak quadrupole transitions and the intensity of its spectrum is about 10^9 times lower than that of a typical heteronuclear molecule.

The HD molecule also has no dipole moment when in its equilibrium position. However, an unusual non-adiabatic effect causes a weak dipole moment to be induced when the molecule is vibrating. This effect is due to the interaction of the electronic motion with the vibration and rotation of the molecule. That is, the electron distribution does not adjust quickly enough to the motion of the nuclei, and the two are always slightly "out of step".

This dipole moment was predicted by Wick (1935) and the dipole rotation-vibration spectrum of HD was first observed experimentally by Herzberg (1950). Durie and Herzberg (1960) subsequently determined the molecular constants of HD from

detailed measurements of the vibration-rotation spectrum, and the dipole moment was later measured by Trefler and Gush (1968).

Though the dipole moment of HD is much smaller than a conventional molecular dipole moment, it nevertheless still exceeds a typical molecular quadrupole moment by several orders of magnitude. Abgrall et al. (1982) computed radiative transition probabilities for the ground electronic state of HD. The lifetimes of rotation-vibrational levels were calculated to be of the order of 10^3s, about three orders of magnitude shorter than the lifetimes of H_2 levels (Black and Dalgarno, 1976).

The infrared rotational-vibrational spectrum of HD differs dramatically from that of H_2. It is dominated by the dipole transitions of the R branch ($\Delta J = -1$) and the P branch ($\Delta J = +1$). Table 3.1 shows the v=1-0 band dipole spectrum of HD (from Abgrall et al., 1982).

R Branch ($\Delta J = -1$)			P Branch ($\Delta J = +1$)		
Line	J'-J"	Wavelength	Line	J'-J"	Wavelength
R(0)	1 - 0	2.690 μm	P(1)	0 - 1	2.823 μm
R(1)	2 - 1	2.633 μm	P(2)	1 - 2	2.898 μm
R(2)	3 - 2	2.581 μm	P(3)	2 - 3	2.980 μm
R(3)	4 - 3	2.535 μm	P(4)	3 - 4	3.069 μm
R(4)	5 - 4	2.494 μm	P(5)	4 - 5	3.165 μm
R(5)	6 - 5	2.459 μm	P(6)	5 - 6	3.267 μm
R(6)	7 - 6	2.428 μm	P(7)	6 - 7	3.377 μm
R(7)	8 - 7	2.402 μm	P(8)	7 - 8	3.494 μm
R(8)	9 - 8	2.380 μm	P(9)	8 - 9	3.619 μm
R(9)	10 - 9	2.363 μm	P(10)	9 -10	3.753 μm
R(10)	11 -10	2.349 μm	P(11)	10 -11	3.895 μm

TABLE 3.1 The v=1-0 band dipole spectrum of HD.

The higher transition rates of HD relative to H_2 can greatly enhance the relative strength of HD line emission in the interstellar medium above that expected from the [HD]/[H_2] ratio alone. In certain environments, the HD spectrum might be expected to be only a factor of 10 to 100 weaker than the H_2 spectrum, despite an abundance ratio [HD]/[H_2] of the order of 10^{-5}.

Several effects can contribute to raising the HD flux relative to H_2. First, the energy levels of HD are lower than those of H_2. In a collisionally excited environment, this results in the collisional excitation rate for each HD level exceeding that for the corresponding H_2 level by a factor of $\exp(\Delta E/kT)$, where ΔE is the energy difference between the HD and H_2 levels. A further consequence of this energy difference is that a given HD level is populated in a larger volume of gas than its H_2 counterpart for a given temperature distribution in the molecular cloud. These two effects are relatively small, and would enhance HD emission by less than an order of magnitude in most circumstances.

A more significant increase results from the differences in radiative lifetimes discussed previously. The molecular densities in dense clouds are generally high

enough to thermalise many H_2 level populations. That is, the high gas density results in the collisional de-excitation rate of H_2 exceeding its radiative de-excitation rate. Many excited H_2 levels therefore decay collisionally, rather than radiatively, thus reducing line emission from H_2.

The collision rate per molecule is given by

$$R = n(H_2)\sigma_c v_{th} = 5\times10^{-14}\, n(H_2)\ T^{1/2}\ s^{-1} \qquad (3.3.17)$$

where $n(H_2)$ is the volume density of H_2 molecules, σ_c is the H_2-H_2 collision cross-section and v_{th} is the molecular thermal velocity.

At a temperature of 2000 K, typical of shock-heated gas in, e.g. the Orion Nebula, the collision rate exceeds the radiative decay rate at H_2 densities $n(H_2) \geq 10^6$ cm^{-3}. Since this density is comparable to those observed in molecular clouds, it is likely that collisional de-excitation dominates in the shock-heated gas in these regions.

Collisional de-excitation, however, is not significant for HD. The radiative lifetimes of HD levels are about 10^3 times shorter than those of H_2, and consequently the radiative decay rates of HD levels exceed collisional decay rates at the densities encountered in molecular clouds. Therefore collisional de-excitation remains negligible for HD, and each collisional excitation still results in radiative decay.

Collisional de-excitation reduces line emission from H_2 by a factor

$$\frac{t_{H_2}(r)}{t_{H_2}(c) + t_{H_2}(r)} \qquad (3.3.18)$$

where $t_{H_2}(r)$ and $t_{H_2}(c)$ are the radiative and collisional decay rates. Since no collisional de-excitation of HD takes place, the photoemissivity ratio is increased by the inverse of this term.

The greatest $I(HD)/I(H_2)$ enhancements can occur in environments where higher vibrational-rotational levels are excited not collisionally, but radiatively by a strong flux of ultraviolet photons.

A H_2 molecule absorbs UV photons at wavelengths of lines in the Lyman and Werner bands between 912 Å and 1108 Å. Absorption of an ultraviolet photon leaves a H_2 molecule in an excited electronic state. This leads to a radiative transition either into the vibrational continuum of the ground electronic state, i.e. to the photodissociation of the molecule (Stecher and Williams, 1967), or into a bound vibrational-rotational level of the ground electronic state.

A transition to a bound state is followed by a cascade of radiative decays at infrared wavelengths through the vibrational-rotational levels (Black and Dalgarno, 1976). In the limit of a low UV radiation field 90% of UV photons at the wavelengths of the H_2 Lyman and Werner transitions result in cascade, 10% in photodissociation. In stronger UV fields however, multiple pumping becomes significant (Shull, 1978). A H_2 molecule already in an excited vibrational-rotational state can absorb a further UV photon before it has returned to the ground state. This cuts short the cascade and hence reduces the number of infrared photons emitted per absorption of a UV photon. In still stronger UV fields, repumping can occur while the molecule is in a high enough

vibrational state for direct photodissociation to take place, reducing the H_2 abundance itself.

Multiple pumping becomes significant at UV fluxes $F/F_0 > 10^4$, where F_0, the mean interstellar UV flux, equals 3×10^{-8} photons $cm^{-2}\, s^{-1}\, Hz^{-1}$. Direct photodissociation from high vibrational levels sets in at UV fluxes $F/F_0 > 10^5$.

For HD levels, on the other hand, radiative lifetimes are three orders of magnitude shorter than for H_2 levels, so the flux of UV photons has to be three orders of magnitude higher before multiple pumping of HD becomes significant. That is, fluorescent emission from HD molecules in a molecular cloud continues to increase in proportion to the ultraviolet flux up to UV fluxes 1000 times higher than those at which repumping of H_2 sets in. Under these conditions, $I(HD)/I(H_2)$ enhancements of 10^3 to 10^4 are possible. This, assuming an abundance ratio $[HD]/[H_2] = 10^{-5}$, implies that HD fluorescent emission of 10^{-2} to 10^{-1} that of H_2 may occur in such environments (Hartquist et al., 1982).

To summarise, line emission from HD can exceed the value expected from the $[HD]/[H_2]$ abundance ratio by several orders of magnitude, both in collisionally excited and in radiatively excited environments. In the collisionally excited case, collisional de-excitation suppresses emission from H_2, but not HD, at high gas densities, and in the radiatively excited case, repumping and photodissociation suppresses emission from H_2, but again not from HD, in strong UV radiation fields. In both cases the much shorter radiative lifetimes of HD levels, a consequence of the dipole moment of HD, are responsible.

3.4 Measurements of Deuterium Abundance

This section will review previous deuterium abundance determinations. They will be presented in a rough order from the earth outwards, with most attention being paid to studies of interstellar matter.

3.4.1 Solar System Deuterium Measurements

Starting with a brief look at terrestrial and solar measurements, sea water has a D/H ratio of 15×10^{-5}, which is also the ratio found in water from meteorites. Chemical fractionation enhances deuterium concentrations here though, and these values don't represent the true D/H ratio in matter as a whole.

Spectroscopy of the sun's photosphere has yielded a very low value for deuterium abundance, with an upper limit of 0.4×10^{-5} being reported by Grevesse (1970), for example. This is to be expected, since, as discussed previously, deuterium is rapidly converted to helium-3 within stars. Hence it is measurements of this helium-3 that can give information about the original, protosolar deuterium abundance.

The $^3He/^4He$ ratio has been measured both in meteorites (Black, 1972) and in the solar wind. Geiss et al., (1972) used aluminium foil collectors placed on the lunar surface to trap solar wind particles. Values of $^3He/^4He = 5 \times 10^{-4}$ were obtained in

both cases, implying an upper limit of 5×10^{-5} for the D/H ratio. This limit is the value the ratio would have if all ^{3}He were made by stellar processing. It is free of any cosmological input. Introducing assumptions about the primeval helium-3 abundance allows a better estimate, and not simply an upper limit, to be obtained, and a protosolar D/H ratio of $(2 \pm 1) \times 10^{-5}$ has been derived from these measurements, though this is now cosmological model dependent.

Turning next to measurements made of Jupiter's atmosphere, observations of deuterated methane (Beer and Taylor, 1973) have produced a D/H ratio of $(5.1 \pm 2.2) \times 10^{-5}$. The large error quoted reflects the difficulties involved in trying to correct for chemical fractionation and atmospheric effects.

A different approach to observing Jovian deuterium was taken by Trauger et al. (1973), who obtained the abundance of the HD molecule by measuring a dipole absorption line at visible wavelengths. The detection of this very faint feature was achieved by using a spectrometer design which accepted light from the entire Jovian disk while compensating for the planet's large rotational Doppler shifts, thereby allowing the full spectral resolving power of over 10^5 to be preserved. They obtained a value of $D/H = (2.1 \pm 0.4) \times 10^5$. This result is largely model-independent, free from chemical effects, and is based on a direct measurement of presumably unprocessed material from the proto-solar nebula. It is therefore particularly well-founded, and agrees well with the values obtained from solar system helium-3 measurements.

3.4.2 Observations of the Deuterium 92 cm Line

Beyond the solar system, deuterium has been searched for in the interstellar medium, both in atomic form and bound in molecules.

Most of the early attempts at detecting it in interstellar matter were based on searching for the 91.6 cm ground state hyperfine transition of atomic deuterium. Numerous such observations were carried out throughout the nineteen-fifties and early sixties, usually searching for the line in absorption against the strong radio continuum emission from either the Galactic Centre or Cassiopeia A sources. No detections were made, despite ever increasing sensitivities, and these efforts culminated in Weinreb's upper limit of $D/H < 7 \times 10^{-5}$ towards Cas A (Weinreb, 1962).

Further observations were carried out by Cesarsky et al. (1973) some ten years later, towards the Sagittarius A Galactic Centre source. They obtained a possible detection, tentatively putting forward a D/H ratio lying between 5×10^{-4} and 3×10^{-5} in this direction (the uncertainty being due to the poorly known optical depth of the intervening hydrogen). However, two more recent measurements towards the Galactic Centre, made with the Ooty facility by Sarma and Mohanty (1978), and Anantharamiah and Radhakrishnan (1979), do not confirm this feature, and report upper limits of D/H $\sim 5 \times 10^{-5}$, similar to Weinreb's.

3.4.3 Observations of Deuterium in Diffuse clouds

The advent of orbital ultraviolet telescopes in the early seventies, in particular the OAO-3 "Copernicus" satellite, made possible new techniques for observing

interstellar deuterium. Lyman transitions could now be observed in absorption against nearby hot stars, as could UV molecular transitions of HD.

Such measurements had to be confined to the brightest, hottest and least reddened stars in the sky, a limitation imposed by the low sensitivity of the Copernicus spectrometer at the wavelengths of the higher hydrogen Lyman lines.

In the first of these observations, Rogerson and York (1973) measured several atomic deuterium and hydrogen lines against the nearby (81 parsec distant) B-type giant beta Centauri, obtaining a D/H ratio of 1.4×10^{-5} along this line of sight. The same authors subsequently extended this work to include four further lines of sight, two of them towards other B stars about a hundred parsecs away, but also two towards hotter O and Wolf-Rayet stars, giving lines of sight up to one thousand parsecs long. These results yielded an average D/H ratio of $(1.8 \pm 0.4) \times 10^{-5}$ for the local interstellar medium. Further, higher accuracy observations allowed more sophisticated analyses to be made, and apparent directional variations in the D/H ratio could be discerned. However, doubt was cast on all these results when Vidal-Madjar and coworkers' (Vidal-Madjar et al., 1983) observations towards the B-giant ϵ Persei showed the deuterium features along this line of sight to be varying by a factor of about three on a time scale of a few hours. They interpreted this as being caused by hydrogen Lyman lines in the high velocity stellar wind being blue-shifted far enough to contaminate the corresponding interstellar deuterium Lyman transitions. They reanalysed previous data in light of this effect, and revised downwards their estimate of the local interstellar D/H ratio from 2×10^{-5} to 5×10^{-6}. However, the possibility of residual systematic error in this estimate cannot be excluded.

The Copernicus satellite was also used to observe molecular transitions of interstellar HD. As in the previous case, the lines were measured in absorption against nearby early type stars. These observations, carried out by Spitzer et al. (1973), were made in the context of a broader investigation of molecular hydrogen. HD/H_2 ratios of about 10^{-6} were measured, although there was a spread of over an order of magnitude in the values obtained along different lines of sight. Self shielding of the H_2, resulting in lower photo-dissociation rates for H_2 relative to HD, makes obtaining the D/H ratio itself from these results difficult.

3.4.4 Observations of Deuterium Bearing Molecules in Molecular Clouds

Millimetre-wave studies of the rotational transitions of deuterium-substituted molecules in Giant Molecular Clouds constitute the most numerous group of observations of deuterium in the interstellar medium. Thirteen such molecules have been observed to date. The brightest and most extensively studied are DCO^+, DCN, (e.g. Penzias, 1979) and DNC (e.g. Turner and Zuckerman, 1979). Others include CH_3OD, NH_2D (Rodriguez-Kuiper et al., 1978), HDO (e.g. Beckman et al., 1982), DC_3N (Langer et al., 1980), DC_5N (Schloerb et al., 1981), HDCO (e.g. Loren and Wootten, 1985), N_2D^+, CH_2D^+, and most recently C_2D (Combes et al., 1985) and H_2D^+ (Phillips et al., 1985). Abundances of deuterated species have been found to be enhanced by as much as three orders of magnitude relative to the ISM [D]/[H] ratio.

Deuterated molecules have been found in several dense clouds, but two objects in particular have received the greatest attention, Orion Molecular Cloud-1, a nearby example of a warm cloud where the temperature, T = 75 K, approaches the energy difference of some deuterium exchange reactions, and the Taurus Molecular Cloud, a cold cloud with a temperature, T = 10 K, far below that of the exchange reaction energies.

Comparison of the degree of deuterium enhancement found at these two temperatures provides useful information about the deuterium exchange reactions responsible for the deuteration of the observed species.

For example, the relative abundance of deuterated formaldehyde is similar in both OMC-1 and TMC-1, with $[H_2CO]/[HDCO] = 10^{-2}$ (Loren and Wootten, 1985) suggesting that a reaction with a high exothermicity ($\Delta E/k \gg 75$ K) is important in the formation of HDCO. This indicates that the fractionation route for formaldehyde starts with deuterium exchange in CH_3^+, ($\Delta E/k$ = 370 K), rather than deuterium exchange in H_3^+, ($\Delta E/k$ = 90 K to 180 K).

On the other hand the $[DCO^+]/[HCO^+]$ ratio differs significantly in these two environments. $[DCO^+]/[HCO^+] < 2\times10^{-3}$ in OMC-1 (Penzias, 1979) but $[DCO^+]/[HCO^+] = 1.5\times10^{-2}$ in TMC-1 (Guelin et al., 1982). Deuterium enhancement is much higher at the lower temperature, suggesting that DCO^+ is formed mainly from H_2D^+.

Studies of deuterated molecules have also been used to attempt the observational test of the origin of deuterium proposed by Ostriker and Tinsley (1975). This test sets out to determine whether deuterium is of cosmological or stellar origin by measuring the deuterium abundance gradient radially across the galaxy.

From observations of metal abundances, the amount of processing within stars undergone by interstellar matter is known to increase towards the centre of the galaxy. If deuterium is produced in stars, then its abundance should follow the same trend. If it is of cosmological origin, then its abundance should decrease in regions of high metal abundance, that is, towards the centre of the galaxy.

Attempts have been made to determine the galactic deuterium gradient from observations of the [DCN]/[HCN] and $[DCO^+]/[HCO^+]$ ratios in several molecular clouds (Penzias et al., 1977, Penzias, 1979). These ratios show an increase with galactic radius, which the authors interpret as evidence for a cosmological origin. Others, however (Beckman et al., 1982) caution that the interpretational uncertainties in such measurements render that conclusion speculative.

3.5 A Search for HD In OMC-1

3.5.1 Molecular Hydrogen in the Orion Nebula

At visible wavelengths, The Orion Nebula is seen as a region of reflection nebulosity illuminated by four young O and B type stars known as the Trapezium cluster. It is part of a large molecular cloud lying about 500 parsecs from the sun.

A bright 2 μm pointlike source, the Becklin-Neugebauer object, lying about one arcminute south of the Trapezium stars, was discovered by E. Becklin in 1967. It attracted interest at a candidate protostar because of its low colour temperature, T_{col}=650 K. Measurements of the BN object at 20 μm found it to be embedded in a region of extended mid-IR emission. This region, the Kleinmann-Low nebula, has a colour temperature of 70 K, (at the time of its discovery the coolest object known) and extends about 30" around BN. Maps of the KL region in the 4.8 μm to 20 μm bands, made at higher spatial resolution, identified several further pointlike sources with low colour temperatures, IRc2 to IRc9, embedded in the nebula.

Mid-infrared spectrophotometric observations (Downes et al., 1981) found the source IRc2 to have an exceptionally deep silicate absorption feature, indicating that it is heavily obscured. The large extinction implied by these measurements suggests that IRc2 has a luminosity of L = 10^5 L_{sun}. It is the most luminous object in the Orion-KL complex. IRc2 is now thought to be a massive (M = 50 M_{sun}) protostar and the source of most of the luminousity in the region.

It was in the Orion Nebula that the first infrared detection of molecular hydrogen in the interstellar medium was made. Gautier et al. (1976) used a Fourier Transform Spectrometer to observe the BN object and part of the KL nebula in seven lines of the H_2 v=1-0 band. H_2 emission was detected from both sources. The data suggested that the vibrational-rotational levels were collisionally excited, and that the gas was at a temperature of about 2000 K, markedly higher than the observed CO temperatures (6 K - 75 K) or the dust temperature in the KL nebula (70 K).

Subsequent observations (Grasdalen and Joyce, 1976, Beckwith et al., 1978) have found the H_2 emission to have a spatial extent of almost two arcminutes. The emission does not peak near the BN object, but rather is strongest in two lobes of unequal brightness lying on opposite sides of the KL nebula, and roughly centred on the heavily obscured source IRc2. Figure n shows the map by Beckwith et al. (1978) of the Orion Nebula in the H_2 v=1-0 S(1) transition, measured with a 5" spatial resolution. The larger and brighter of the two emission lobes, named Peak 1, lies about 20" NW of KL, the fainter, Peak 2, lies about 15" SE of KL. The v=1-0 S(1) line intensities at the emission peaks in these two lobes were found to be 7.5×10^{-3} ergs s^{-1} cm^{-2} sr^{-1} and 5.0×10^{-3} ergs s^{-1} cm^{-2} sr^{-1} respectively. For a H_2 temperature of 2000 K, this implies a H_2 column density of the order of 10^{19} cm^{-2}. Various molecular observations suggest that the volume density in the region is of the order of 10^6 cm^{-3}, so the thickness of the emitting region is about 10^{13} cm, five orders of magnitude smaller than the projected size of the emission on the sky (~0.3 pc). That is, the H_2 emission emanates from a thin sheet of material.

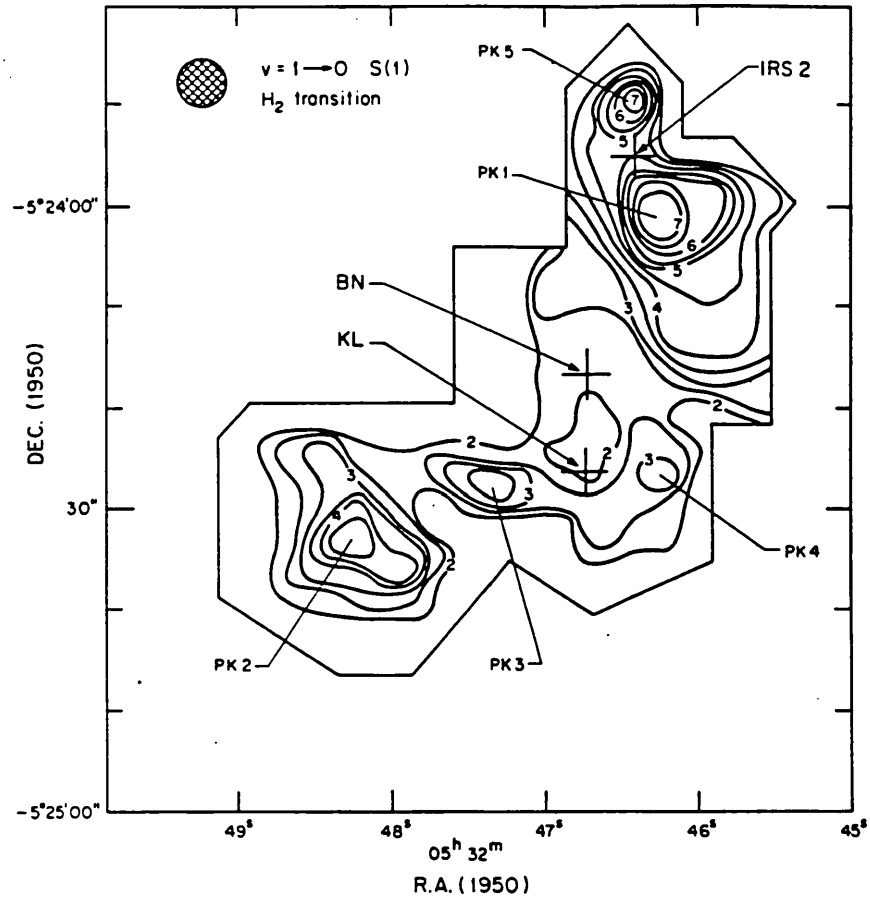


FIGURE 3.2 A H_2 Map of OMC-1 (Beckwith et al., 1978).
The molecular hydrogen emission in Orion as seen in the v=1-0 S(1) transition with 5" spatial resolution. Contour intervals are 6.1×10^{-4} ergs s^{-1} cm^{-2} sr^{-1}. The extent of the area mapped is enclosed by the straight line boundary.

Velocity-resolved observations of the H_2 v=1-0 S(1) line (Nadeau and Geballe, 1979) show the line wings to extend over 100 km s^{-1} from the line centre. The molecular hydrogen is expanding at velocities of typically 40 km s^{-1}, and up to 100 km s^{-1}, about a centre located near IRc2. The two lobes of H_2 emission correspond to the two streams of a bipolar flow. Millimetre-wave observations of the rotational transitions of other molecules (e.g. Wright et al., 1983) confirm the picture of a high velocity bipolar outflow centred on IRc2. A dynamical age for the system of 2×10^3 years can be obtained from the size and expansion velocity of the flow.

Millimetre-wave observations (Plambeck et al., 1982) have also identified a 19" × 11" ring of expanding high density molecular gas centred on IRc2 and perpendicular to the axis of the bipolar outflow. Beyond this ring there appears to be a differentially rotating disk of more tenuous molecular material.

The picture that has emerged of the Orion-KL region (e.g. Hasegawa, 1986) is of a massive $M = 50\,M_{sun}$ protostar, IRc2, driving a supersonic bipolar outflow. Strong shocks occur at the interface between the outflow and the surrounding molecular gas, giving rise to intense spectral line emission from vibrationally excited

H_2. The protostar is ringed by the disk of high density gas from which it probably formed.

3.5.2 The Instrument

The instrument used for the observations was a Fabry-Perot spectrometer in tandem with a CVF photometer. The high resolution of a Fabry-Perot spectrometer was needed to isolate the HD lines from nearby H_2 and telluric lines. This instrument was chosen in preference a grating spectrometer because of the extended nature of the source. The higher throughput of the Fabry-Perot spectrometer allowed an 15" beam to be used, whereas the UKIRT grating spectrometer, CGS 2, is restricted to a 5" aperture.

The CVF photometer used for the observations is a common user UKIRT photometer called UKT1. It's filter wheel includes a Circular Variable Filter for the 2 μm band.

The Fabry-Perot etalon has a nominal resolution of 100 km s^{-1}. The etalon gap d is 106.5 μm, corresponding to a free spectral range $\Delta\nu$ of

$\Delta\nu = 1 / 2d = 46.95\ cm^{-1}$.

The etalon is scanned in wavelength by piezoelectric actuators. The tilt and separation of the etalon are measured capacitatively. The etalon gap is kept constant despite the varying orientation of the instrument by a servo control mechanism. The servo stabilisation is provided by a commercial instrument produced by Queensgate Instruments. This instrument senses variations in the etalon gap and adjusts the voltages on the piezoelectric actuators to compensate.

3.5.3 The Choice of Line

Measurements of several HD lines would be needed to determine the physical parameters of the source. It became apparent early in the observations, however, that it would not be possible to detect more than one line in the time available. Therefore the whole observing session was devoted to a single line to obtain the best possible upper limit.

The HD line selected for the observation was the R(5) line at 2.459 μm. The choice of line to observe was determined by two requirements. The line had to be in a spectral region of low telluric absorption, and there could be no other bright emission lines from the source within a few resolution elements of it.

The HD v=1-0 R branch, (i.e. with $\Delta J=-1$) lies between about 2.3 μm and 2.69 μm. Although the low order lines below R(7) lie beyond the long wave cut-off of the photometric K band filter (2.4 μm), the atmospheric transmission at longer wavelengths remains high enough to allow lines down to R(4), at λ = 2.494 μm, to be observed.

Observations of the HD R-branch are also hampered by the proximity of the v=1-0 Q branch (ΔJ = 0) of H_2. Many HD R branch lines have wavelengths very

close to H_2 Q branch transitions. The HD R(4) line, for example, lies within 0.006 μm of the H_2 Q(7) line, and the HD R(5), R(6) and R(7) transitions differ in wavelength by less than 0.005 μm from the H_2 Q(5), Q(3) and Q(1) transitions respectively. The nominal resolution of the Fabry-Perot etalon at 2.5 μm is 0.0014 μm, so the lowest order HD R-branch lines observable through the atmosphere all lie within four or so instrumental line-widths of molecular hydrogen lines, which would be expected to be at least two orders of magnitude brighter than the HD lines in the source being observed.

The P branch, ($\Delta J = +1$) of HD lies between 2.82 μm and about 4.5 μm. The low order P branch lines lie in the region of very structured telluric absorption beyond the short wave wing of the L band. The P branch is also in a region of higher background emission, reducing the sensitivity of observations with a warm etalon.

Although the HD R(5) spectrum is contaminated by the wing of the H_2 v=1-0 Q(5) line, it was nevertheless chosen as a good compromise between atmospheric transparency, background emission, and interference from nearby emission lines.

3.5.4 Observing Procedure

a) The Peak 1 Source

Since the same physical processes are assumed to populate the excited vibrational levels of both the H_2 and HD molecules in the region, the peak of HD emission is expected to coincide with the peak of H_2 emission. The observations were therefore made towards "Peak 1" at R.A.= $5^h\,32^m\,46.2^s$, dec.= -5° 24' 01" (Beckwith et al., 1978), the point of most intense molecular hydrogen line emission in the Orion nebula. The "Peak 1" source was located and centred in the beam by scanning the 2.1218 μm H_2 v=1-0 S(1) line, the brightest of the molecular hydrogen lines in the source, at several points near the expected position of the peak.

b) Spectral Calibration

The initial spectral calibration of the instrument was carried out in two parts. With the Fabry-Perot etalon withdrawn from the beam, the CVF filter was calibrated using several lines from a krypton spectral lamp. The variation of wavelength λ with angular position θ of the CVF may be described by a relation of the form

$$\lambda = a\theta + b \qquad (3.5.1)$$

where a and b are instrumental constants.

The etalon was then introduced into the beam, and calibrated using a pair of krypton lines at 2.430 μm and 2.427 μm. These lines were chosen for their proximity to the HD R(5) line. They lie in the adjacent spectral order of the etalon. The parallelism of the etalon gap was determined by observing the symmetry and width of the instrumental profile.

The spectral calibration was monitored throughout the observations by repeating the krypton line measurements at the beginning and end of each observing session, and at regular intervals in between.

c) Observation Parameters

The spectral line was sampled at fourteen points 55 km s^{-1} (4.6×10^{-4} μm) apart. (The Fabry-Perot etalon can be scanned in increments of 3.66 km s^{-1}. The chosen sampling interval corresponds to fifteen of these steps). The line profile was sampled at about three points across its Full-Width-Half-Maximum, and the scan length was chosen to ensure that it extended far enough into the line wings to obtain a good estimate of the continuum flux level.

The signal was modulated by chopping the secondary mirror. An East-West chop of 50 arcsec. amplitude was used. The large chop amplitude was needed to prevent the very extended source from contaminating the reference beam.

Table 3.2 shows a journal of the observations. The "left beam" has the reference beam to the east of the source, the "right beam" has the reference to the west.

Date	Beam	No. of Scans
20th Jan.	Right	133
22nd Jan.	Right	51
22nd Jan.	Left	77
23rd Jan.	Right	69
Total		330

TABLE 3.2 Journal of Observations

The source was observed each night until it was about three hours below the meridian. That is, it was observed in the airmass range AM = 1.1 to AM = 1.5 .

The integration time per point was 5s, and a total of 330 scans of the R(5) line were obtained over the three nights of observation. So in all, the observations amounted to 6 hr 25 min of integration time on the line.

The scans were not stored singly. Rather, the on-line reduction software carried out some further integration of the signal before storage to disk. Sequences of either six or ten consecutive scans were coadded and stored, together with their associated error values, as a single disk file. The instrument software also maintained the long-term coadded signal to allow the course of the integration to be monitored.

Two bright stars were used to calibrate the flux, BS 1239 (m_v=3.8, Sp.= B3V) and BS 3314 (m_v=3.90, Sp. = A0V).

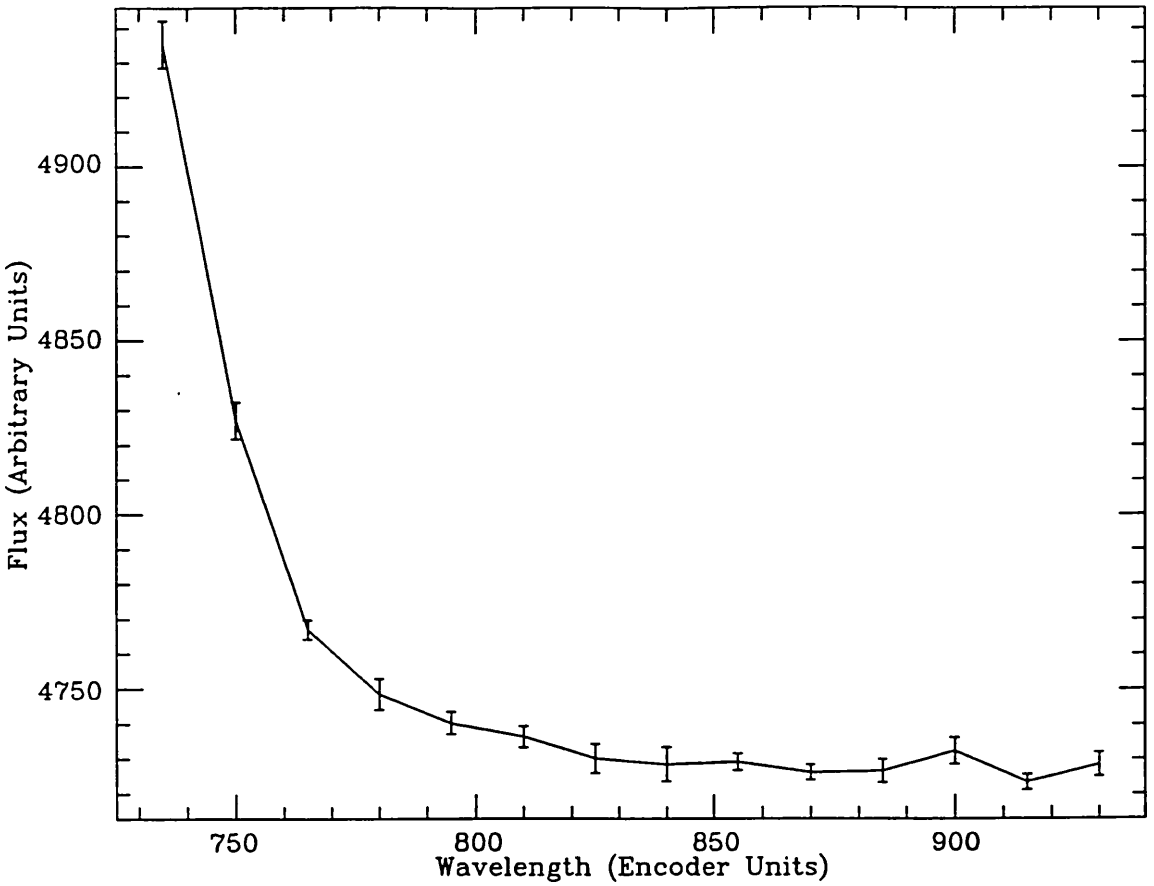


FIGURE 3.3 A Raw HD Spectrum.
The spectrum is centred on the position of the HD R(5) line. It is composed of ten coadded scans, each point was integrated for 50 s, and 1σ error bars are displayed. The strong gradient is the long-wave wing of the v=1-0 Q(5) line of H_2.

3.5.5 Data Reduction

a) Despiking

Due to a defect in the data acquisition system, some very large noise spikes were present in the data. Each spike was confined to a single sample, but generally exceeded the r.m.s. noise by over an order of magnitude. The sensitivity of the observation would have been significantly reduced had the spikes been left in the data. Rather than discard the contaminated scans altogether, the spikes were removed by performing a linear interpolation across adjacent samples. The spectra were sufficiently oversampled to allow this.

b) Wavelength Drift

After removal of the noise spikes, the scans were corrected for drifts in the wavelength calibration of the Fabry-Perot etalon. The calibration had been monitored during the observations by making regular measurements of the krypton spectral lamp. The worst drift observed within a single night's observations was 12 km s^{-1}, about 7%

of the resolution. The smallest observed drift was under 1.5 km s^{-1}. These drifts did not degrade resolution significantly and were not corrected for. The night-to-night variations, however, were larger. These calibrations differed by up to 40 km s^{-1}, and the coadded spectra from the separate nights were shifted onto a common wavelength scale before being coadded.

c) Weighted Mean

The spectra were not of uniform quality. Some sets of coadded scans suffered from degraded sensitivity. (The noisiest scans had a 55% lower sensitivity than the best). Periods of poor seeing or variable atmospheric transparency, for example, might increase noise temporarily. To take account of this, a weighted mean of the scans was computed. Each sample in the stored spectra (each spectrum consisted of six or ten scans coadded at the time of the data acquisition) was weighted by its inverse variance.

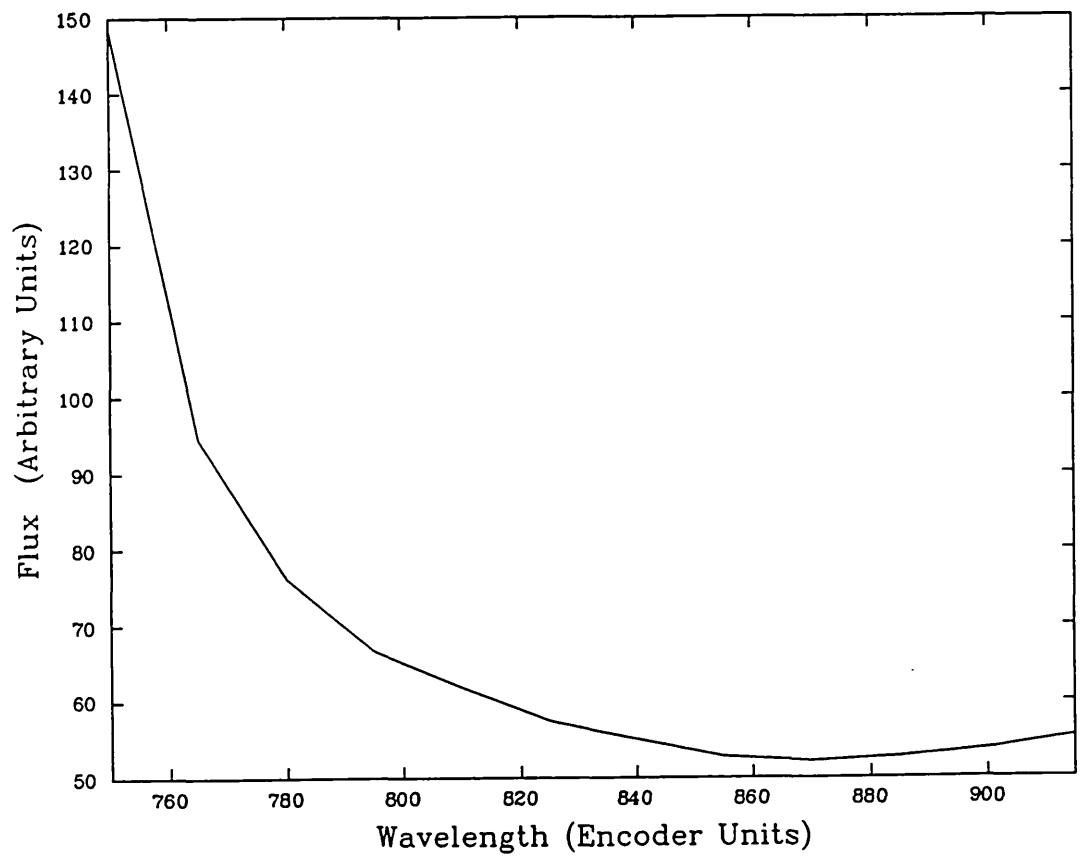


FIGURE 3.4 All Coadded Spectra

d) Gradient Subtraction

The HD v=1-0 R(5) line frequency differs from the frequency of the H_2 v=1-0 Q(5) line by only 7 cm^{-1}. The resolution of the Fabry-Perot etalon is 2.2 cm^{-1}. Because of the proximity of the H_2 Q(5) line, a strong gradient is present on the continuum of the HD line scans, caused by the long wave wing of the H_2 line.

The instrumental profile has a broad wing. The transmission of a Fabry-Perot etalon (neglecting absorption) has the form

$$I_t/I_i = 1 / (1 + F \sin^2(\delta/2)) \qquad (3.5.2)$$

where I_i and I_t are the incident and transmitted intensities, F is the finesse and δ is the phase shift undergone in traversing the optical path length of the etalon.

The line wing was approximated by a power law function. A fourth power least-squares fit was subtracted from the spectrum to remove the gradient without introducing features resembling the instrumental profile.

e) Correlation Detection

The signal to noise ratio of the resulting spectrum can be enhanced by passing the signal through a matched filter.

A matched filter has an impulse response equal to a time-reversed version of the signal waveform to which it is matched. That is, convolving the matched filter's impulse response with the input signal is equivalent to auto-correlating the input. The matched filter has an important application in the detection of signals of known shape in the presence of noise. It has the property that, when the noise power spectrum is white over the frequency range occupied by the signal, the signal to noise enhancement produced by a matched filter is the best that can be attained with any linear filter (e.g. Lynn, 1982).

For the case of a discretely sampled signal, the signal to noise enhancement is given by

$$(\Sigma(a_i^2))^{1/2} / a_{pk} \qquad (3.5.3)$$

where the impulse response of the matched filter has the sample values a_n, ... , a_3, a_2, a_1 and where the largest of these samples has the value a_{pk}.

The signal being searched for in this case is the instrumental profile of the Fabry-Perot etalon. Its matched filter can be approximated by a Gaussian function of FWHM equal to that of the instrumental profile. A matched filter of this form, with FWHM = 166 km s^{-1} and sampling interval = 55 km s^{-1} produces an expected signal to noise enhancement of 1.5 .

The filtering operation reduced the noise by 2.2 times. The discrepancy with the expected value is within the error expected for a spectrum of twelve samples.

The standard deviation of the resulting, final spectrum, expressed relative to the H_2 v=1-0 S(1) line flux at the same position, was

$$\sigma = 2.3 \times 10^{-4} F_{S(1)}$$

No feature resembling the filtered instrumental profile could be seen at the expected position of the R(5) line. The line had not been detected.

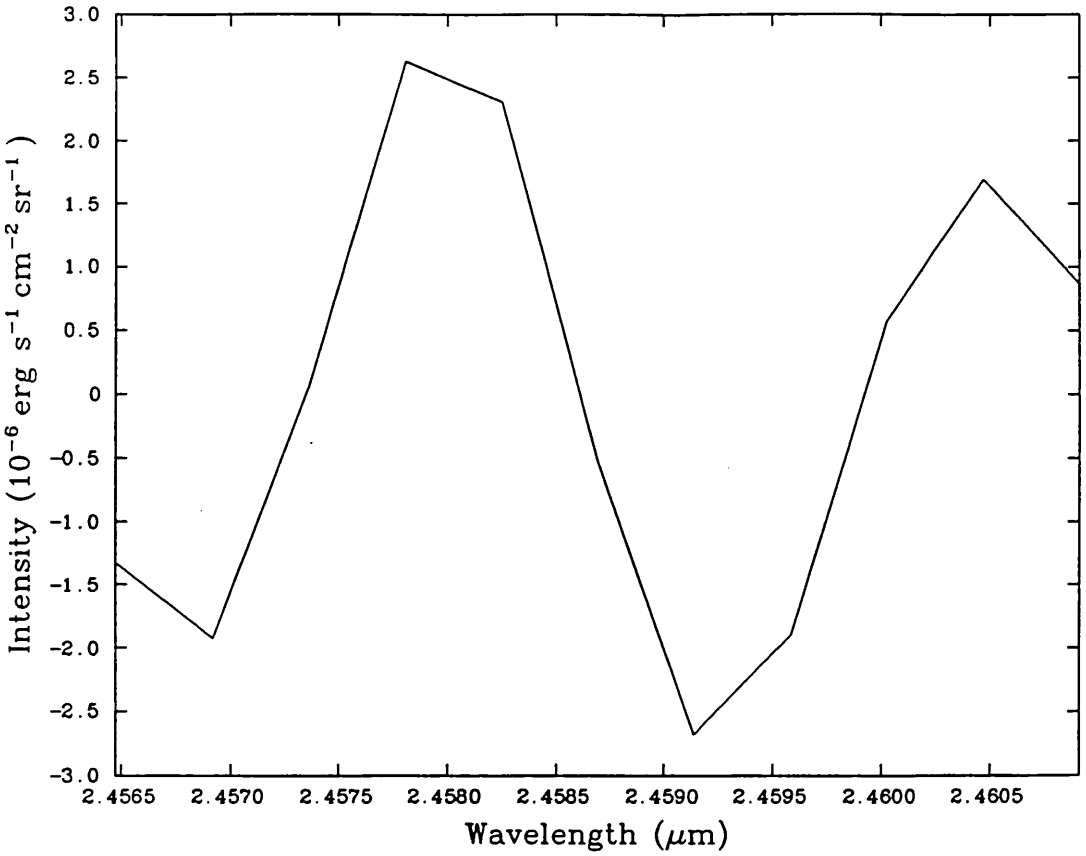


FIGURE 3.5 The Final Matched Filtered Spectrum

The three-sigma upper limit for the HD R(5) line intensity at Peak 1 is then

$$F_{R(5)}/F_{S(1)} < 7\times10^{-4}$$

f) Flux Calibration

Two sources of flux calibration were used, the H_2 v=1-0 S(1) line at Peak 1, and the bright star BS 3314, a photometric standard at UKIRT.

The flux in the S(1) line was obtained from the 5" resolution map of Beckwith et al., (1978). The intensity within an 15" diameter circle centred on Peak 1 was integrated, giving a total S(1) flux of $F = 2.6\times10^{-11}$ erg s^{-1} cm^{-2} within the aperture.

BS 3314 has a K magnitude $m_K = 3.94$. The spectral flux of a zero magnitude object at 2.2 μm is $F = 3.9\times10^{-7}$ erg s^{-1} cm^{-2} μm^{-1} (Johnson, 1966). The spectral flux from BS 3314 at 2.43 μm, assuming a Rayleigh-Jeans spectrum, is therefore $F = 7.0\times10^{-9}$ erg s^{-1} cm^{-2} μm^{-1}, and the flux in the instrumental bandwidth is $F = 9.4\times10^{-12}$ erg s^{-1} cm^{-2}.

The value of the instrumental flux unit obtained from the S(1) calibration is

$$1 \text{ unit} = 3.3\times10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$$

the value from the stellar calibration is

$$1 \text{ unit} = 2.2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2}$$

The 33% discrepancy may be attributed to several sources of error. The S(1) map of Beckwith et al. (1978) has a 10% error assigned to the observations, and obtaining the 15" aperture-integrated flux from the map introduces a further 5% error. The measurements of the S(1) flux were repeatable to 8%. The BS 3314 spectra have a 3% formal error, and the instrumental bandwidth was measured to about 5% accuracy.

The stellar calibration is substantially more accurate, and is adopted in the following calculations.

g) Result

Using the stellar flux calibration given above, the three-sigma upper limit for the flux from the HD R(5) line in a 15" aperture centred on Peak 1 is

$$F_{R(5)} < 2 \times 10^{-14} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ corresponding to an intensity upper limit}$$

$$I_{R(5)} < 4 \times 10^{-6} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$$

The external extinction at 2.12 μm towards Peak 1 is 1.4 magnitudes (Beck and Beckwith, 1983). The intensity upper limit corrected for extinction is then

$$I_{R(5)} < 2 \times 10^{-5} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ sr}^{-1}$$

From the error bars on the original six scan and ten scan spectra, the instrument's one-sigma-one-second sensitivity was 4.8×10^{-13} erg s^{-1} cm^{-2}. The total integration time per point was 1650 s. Assuming that the S/N ratio increases as the square root of time, and taking the matched filter into account, the final error, extrapolated from the error on short integrations, is expected to be

$$1\sigma = 5.3 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$$

whereas the spread of points in the final spectrum, obtained from an observation more than six hours long, and spread over three nights, is

$$1\sigma = 6.0 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$$

just 13% lower. That is, if spectral and flux calibrations are performed regularly, integrations of several hours duration can be made without loss of sensitivity.

3.5.6 Discussion

In the limit of very low density, when neither H_2 nor HD levels are in thermal equilibrium and every excitation leads to radiative decay, the line ratio, assuming equal formation rates for the H_2 and HD levels, is given by

$$\frac{f(HD)\, A(HD)\, \tau(H_2)}{A(H_2)\, \tau(HD)} \qquad (3.5.4)$$

where f(HD) is the [HD]/[H_2] abundance ratio, A(HD) is the spontaneous emission probability for the HD v=1, J=6 - v=0, J=5 transition, A(H_2) is the spontaneous emission probability for the H_2 v=1, J=3 - v=0, J=1 transition, τ(HD) is the inverse

radiative lifetime of the HD v=1, J=6 level and $\tau(H_2)$ is the inverse radiative lifetime of the H_2 v=1, J=3 level. The values adopted for these quantities are

f(HD) = 2×10^{-5},
A(HD) = 1.9×10^{-6} s^{-1}, (Abgrall et al., 1982),
$A(H_2)$ = 3.5×10^{-7} s^{-1}, (Turner, Kirby-Docken and Dalgarno, 1977),
τ(HD) = 6.7×10^{-5} s^{-1}, (Abgrall et al, 1982),
$\tau(H_2)$ = 8.3×10^{-7} s^{-1}, (Black and Dalgarno, 1976).

The line ratio at low densities is then 1.3×10^{-6}.

At very high densities, where collisional de-excitation dominates over radiative de-excitation for both H_2 and HD transitions, the line ratio becomes

$$f(HD) \frac{A(HD) \int 9 \exp(-E(HD)/T(z)) \, n_{H2}(z)/Z_{HD}(z) \, dz}{A(H_2) \int 21 \exp(-E(H_2)/T(z)) \, n_{H2}(z)/Z_{H2}(z) \, dz} \qquad (3.5.5)$$

where E(HD), the energy of the upper (v=1, J=6) HD level, is E(HD)/k = 7750 K, (Abgrall et al., 1982), $E(H_2)$, the energy of the upper (v=1, J=3) H_2 level, is $E(H_2)/k$ = 6950 K, $n_{H2}(z)$ is the H_2 number density as a function of distance z along the line of sight, T(z) is the temperature, and Z_{HD} and Z_{H2} are the partition functions. At temperatures $T \gg (E(HD) - E(H_2))$, i.e. $T \gg 800$ K, this expression has the value 5×10^{-5}.

That is, the HD R(5) / H_2 S(1) photoemissivity ratio has the value 1.3×10^{-6} at low densities, where radiative de-excitation is the dominant decay route for both H_2 and HD levels, and rises to 5×10^{-5} at very high densities, where collisional de-excitation dominates for excited levels of both molecules. At intermediate densities, where radiative de-excitation still dominates for HD, but both radiative and collisional de-excitation contribute significantly to the decay of H_2 levels, the photoemissivity ratio rises linearly with density.

The density at which collisional de-excitation sets in for HD may be estimated from the H_2 collisional excitation rates given by Draine, Roberge and Dalgarno (1983).

The collisional excitation rates of HD are not well known. It will be assumed, however, that the rate of formation by collisional excitation of the H_2 (v=1, J=3) level is equal to the collisional formation rate of the HD (v=1, J=6) level. This approximation may be adequate because, while the H_2-H_2 potential and the H_2-HD potential differ at long range, they are similar at short range. Since only very hard collisions can lead to vibrational excitation or de-excitation, the difference in the potentials of the two systems at large separations should not cause their respective vibrational excitation rates to differ significantly. Also, the difference in the level structures of the H_2 and HD molecules may cause only a small difference in their excitation rates.

The rate coefficient $<\sigma v>$ for the v = 1-0 transition in an H_2-H_2 collision is given as

$$<\sigma v> = 6.6\times10^{-14}\, T \exp(-79.99/T^{1/3}) / (1-e^{-5987/T})\ cm^3\ s^{-1} \quad (3.5.6)$$

where $<\sigma v>$ is the product of cross-section and velocity averaged over a Maxwellian distribution.

Taking T = 2000 K, appropriate for the Orion Molecular Cloud, gives the value

$$<\sigma v> = 2.4\times10^{-13}\ cm^3\ s^{-1}.$$

The collisional deexcitation time constant is then

$$\tau = 1 / <\sigma v> n_{H_2}\ s \quad (3.5.7)$$

Equating this to the radiative lifetime of the HD v=1, J=6 level yields an H_2 density $n_{H_2} = 3\times10^8\ cm^{-3}$. This is the density at which the HD v=1, J=6 level attains thermal equilibrium.

The upper limit obtained for the HD R(5) / H_2 S(1) line ratio, $R < 7\times10^{-4}$, exceeds the line ratio in the high density limit, $R = 5\times10^{-5}$, so the observations cannot be used to place new constraints on the density of the shocked gas, or on deuterium abundance in the ISM.

3.6 Conclusion

To attempt a determination of the deuterium abundance ratio in the ISM, and to learn more about the physical conditions in the Orion Molecular Cloud, near-IR vibrational-rotational emission from the HD molecule has been searched for at Peak 1 in OMC-1. An upper limit for the intensity of the HD v=1-0 R(5) line at Peak 1 in the Orion Nebula has been obtained. The three-sigma upper limit, corrected for extinction, is

$$I_{R(5)} < 2\times10^{-5}\ erg\ s^{-1}\ cm^{-2}\ sr^{-1}$$

This limit is too high to provide additional constraints on the physical conditions in the gas. Future observations of fluorescent H_2, where HD emission is expected to be much stronger, and the use of more sensitive, array detector based spectrometers, might be expected to fare better.

Chapter Four

Molecular Hydrogen Emission From OH Maser Sources

Summary

This chapter describes observations of molecular hydrogen towards the hydroxyl maser source in W75N. They were made with a view to gaining a better understanding of the mechanisms responsible for forming OH and driving the maser emission in compact HII regions.

The model proposed by Hartquist and Dalgarno for interstellar OH masers predicts that they exist in hot gas heated by the magnetic precursors of shocks preceding expanding HII regions. The gas is cooled partially by H_2 line emission, and measurement of this can test the model.

Measurements of the v=1-0 S(1) line have been obtained towards W75N(OH), and seem to be consistent with this model.

4.1 Introduction

Cosmic masers have now been under study for just over twenty years, and the subject continues to attract widespread research interest.

The H_2 observations presented here were concerned with an observational test of some recent models which attempt to account for the OH maser emission found near many compact HII regions. While these models and observations will be discussed in some detail, the preceding general discussions of astronomical masers will be as brief as possible, consistent with presenting an adequate contextual framework for the later material.

Several general reviews of the subject have been produced. A.H. Cook's monograph "Celestial Masers" (1977) provides a broad introductory text, and more recent reviews of observational work (Reid and Moran, 1981) and physical processes (Elitzur, 1982) are also available.

4.2 Cosmic Masers

The existence of naturally occurring masers in the interstellar medium was first established when Moran et al. (1968) used VLBI techniques to show that many of the then recently discovered cosmic OH sources had angular diameters in the milliarcsecond range, corresponding to brightness temperatures of over 10^{12} K, despite dynamical temperatures of only 100 K being implied by the widths of the OH emission lines. Since then over seven hundred sources of maser emission have been discovered in the sky, from five different molecular species and three distinct classes of object.

That masers are so common in the interstellar medium is a consequence of the low density, which prevents the energy levels of many molecules from thermalising, and of the large path lengths available, which allow appreciable gains despite the low densities.

In addition to hydroxyl, maser emission has been observed from water, silicon monoxide, ethanol and CH, in astronomical sources near the two extreme ends of the stellar evolutionary cycle. On the one hand, maser activity has been found to be associated with compact HII regions in OB associations, regions of active star formation. On the other, masers have been observed in the circumstellar envelopes of evolved M giant and supergiant stars. Extremely powerful OH maser emission has also recently been found in several extragalactic sources. That only these types of environment are known to exhibit maser activity should be treated with caution, since no comprehensive all-sky survey of maser emission has ever been carried out, and masers have been searched for mainly where the chances of success seemed high from previous observations.

The main observational test that a radio source is a maser remains the one used by Moran. The brightness temperature is obtained from the angular diameter and flux, the molecular dynamical temperature is obtained from the line width, and the two are compared. A brightness temperature significantly greater than its dynamical counterpart implies maser activity.

Known astrophysical masers span two orders of magnitude in wavelength. OH masers have the lowest frequencies, with emission seen principally from four transitions at around 1.7 GHz. Water masers emit at 22.2 GHz, ethanol from a series of lines near 25 GHz, and SiO at 43 GHz, 86 GHz and 129 GHz. The silicon monoxide transitions are the only ones observed from excited vibrational states of a molecule.

Masers are referred to as "stellar" or "interstellar" to distinguish those around evolved stars from those near HII regions. Stellar maser activity has only been observed from OH, H_2O and SiO. Variability timescales and the proximity of masers to their exciting star both vary with maser frequency. The more energetic water and silicon monoxide masers are found a few stellar radii from their central stars and display variability on a timescale of days. Hydroxyl masers, on the other hand, lie several hundred stellar radii away, and vary over many months.

The spectra of stellar hydroxyl masers have a characteristic double peaked structure due to blue and red shifted components at velocities between 10 km s^{-1} and 50 km s^{-1} relative to their central star. This is interpreted as implying that the masers are in outflowing material. Furthermore those OH masers around long period variable stars have been seen to vary in phase with the stellar light curve, suggesting that the pumping mechanism is radiative in this case. This is one of the few instances where direct observational evidence to distinguish pumping schemes exists. In general, the processes producing population inversion in astrophysical masers are very poorly understood. Models abound, but few observational tests have been suggested. The maser variation in these sources is roughly linear, not exponential, with the pumping flux, suggesting that the masers are saturated. Although most strong masers are believed to be saturated, masers associated with long period variables provide the clearest evidence of this.

Interstellar masers are associated with very compact, expanding HII regions around newly formed O or B stars in molecular clouds. Hydroxyl, water, silicom

monoxide, ethanol and CH masers have been observed in these environments. Maser activity is present only during the earliest stages of the star's life, with H_2O masers, for example, starting to emit when the HII region is still undetectably small, and lasting for about 10^5 years beyond that. Interstellar OH masers differ markedly from stellar ones, with maser emission coming from different transitions in the two cases, suggesting that they do not share a common pumping mechanism. (Interstellar OH masers are discussed in more detail later.)

Ethanol maser emission had until recently been seen only from the Orion Molecular Cloud, one of the closest and most studied interstellar maser sources. Menten et. al. (1986) however now believe they have detected it in ten further sources.

This concludes the brief general look at astronomical masers. Interstellar OH masers will now be described in more detail.

4.3 Interstellar OH Masers

Following the general overview given above, this section will give a more detailed description of OH masers associated with recently formed HII regions.

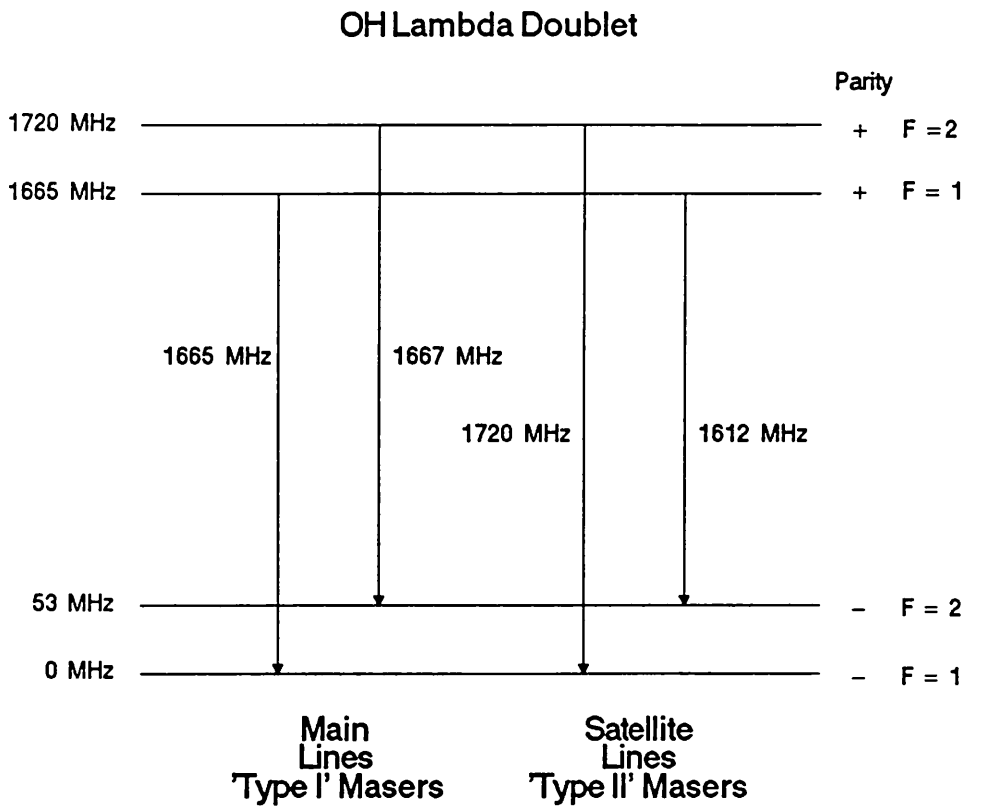


FIGURE 4.1 The OH Λ-Doublet Transitions

The maser emission arises from Λ-doublet hyperfine transitions in the ground rotational state of the hydroxyl molecule. (Weak maser emission from higher rotational states has also been observed, but won't be discussed here.) The ground

rotational state is split into four components as a result of the hyperfine interaction between the electrons and H nucleus, and lambda doubling. Lambda doubling arises in diatomic molecules with dissimilar nuclei, where electronic momenta of opposite sense about the internuclear axis correspond to unequal energies when the molecule is rotating. (The ground state of OH has J=3/2, not zero, due to an unpaired electron being present.)

These four levels give rise to four transitions (fig. 4.1) at frequencies of 1665 MHz, 1667 MHz, 1612 MHz and 1720 MHz, corresponding to wavelengths of around 18 cm. The former two, called "main line" transitions, have spontaneous transition probabilities five to ten times greater than the latter two "satellite line" transitions. The main line transitions conserve the total angular momentum quantum number, F, while the satellite transitions don't. Main and satellite line maser emissions are associated with different astrophysical sites. Main line emission is observed principally near compact HII regions, and these sources are classified as "Type I", while satellite lines dominate the emission from evolved stars, and these are classified as "Type II" sources.

4.3.1 Observations of Type I OH Masers

VLBI observations have shown that the maser radiation from Type I sources is emitted from a large number of separate components, each about 10^{14} cm across, spread over an area up to 10^{17} cm in diameter around the central HII region. Typical line widths of around 2 kHz have been found for these sources. It is hard to tell, in most cases, whether the lines are narrower than would be expected from the temperature of the region. Unsaturated maser lines will be narrower than the original incoming radiation field because the maser gain in the line centre is much higher than in the wings. This will cause the line centre to grow in intensity faster than the wings, narrowing the line. A saturated line, on the other hand, will usually broaden back to its original thermal profile since the centre will saturate first, while the wings continue to grow. Goldreich and Kwan (1974) pointed out that this need not happen, however, as mechanisms exist to maintain the narrowed profile even in saturation. The only indisputable example of line-narrowed interstellar OH maser emission is the W3(OH) source, where a line-width corresponding to a 5 K temperature has been observed. Otherwise widths roughly consistent with the local temperatures have been found, bearing in mind that maser lines will be narrower than non-maser lines from the same region anyway, since they are not broadened by turbulence. (They are subject to a requirement for velocity coherence.)

OH maser velocities relative to thir central source are invariably quite low, (< 10 km s^{-1}), compared with, say, the 100 km s^{-1} velocities found for H_2O masers in the same regions. The OH masers generally seem to be in material not accelerated by the HII region's expanding shock front. In fact in some sources maser spots seem to be slightly red-shifted relative to their HII region. Since the HII region is opaque at the maser frequency, these masers are in front of, and therefore falling in towards, the central star, perhaps implying they are within a remnant accreting envelope. In other cases, though, evidence for outflow and rotation has been found. Accurate positional

measurements by Ho et. al. (1983) found most OH maser spots to be at the edges of their HII regions, often associated with only one dense component in the region.

The power radiated by individual maser spots cannot be determined, since the angular spread of the beams is not known. However, if the spots are assumed to beam in random directions, and hence that the total maser emission from an HII region as a whole is roughly isotropic, the total integrated power from the maser complex can be estimated. Values of 2.1×10^{21} W for W3(OH), and 1.1×10^{20} W for W75N(OH) have been obtained for the 1665 MHz transition, corresponding to photon fluxes of 2×10^{45} s^{-1} and 1.1×10^{44} s^{-1} respectively. (These very large photon fluxes will later be seen to impose severe constraints on radiative pumping schemes).

As already mentioned, interstellar OH masers seem to be saturated, although the evidence for this is not particularly firm. The best observational support for this comes from line width measurements, which often seem consistent with a thermal (i.e. not narrowed) profile, implying saturation. A plausibility argument in favour of saturated masers is also often put forward. It is pointed out that most maser models are stretched to their limits to account for the very high powers emitted by these sources, and the most efficient possible route must be taken at every stage of the process. A saturated maser converts inverted levels to photons more efficiently than an unsaturated one, and it is claimed that the high observed powers can only be produced if the masers are saturated.

Magnetic fields of a few milligauss in the masing regions have been deduced from observations of Zeeman splitting of the maser lines. This effect is also thought to be responsible for the very high (near 100%) polarisations of many Type I OH masers. (Type II masers, by contrast, are practically unpolarised).

4.3.2 Pumping Mechanisms

A suitable mechanism for OH maser emission must be able to form the OH and supply it with energy, in particular inverting the lambda-doublet populations. The pumping mechanism is not known, but this is not considered a serious obstacle to understanding other aspects of the phenomenon, because OH is relatively "easy" to mase. OH maser wavelengths are by far the longest of the celestially observed masers. The transition energies are much lower than the ambient temperature. As most interstellar molecules are found to be very far from thermal equilibrium anyway, since the low densities involved make collisional thermalisation a very slow (10^5-10^7 year) process, inverting the OH level populations does not require particularly unusual physical conditions.

Many possible pumping schemes have been proposed, but these can't yet be excluded observationally.

Radiative pumping is thought to be unlikely in interstellar OH masers. The maser photons are of very low energy, and , since at least one pump photon is required for each maser photon, the process is very inefficient. Typical maser regions emit about 10^{44} 18 cm photons per second, and it seems unlikely that enough suitable higher energy photons would be available. Andresen (1986) though, has recently pointed out

that photodissocation of H_2O by the vacuum ultraviolet flux from the central O/B star can produce inverted OH populations in the maser transition, and can also produce the required maser powers, because the exceptionally broad (550 Å wide) H_2O absorption band involved allows a large fraction (about 30%) of the star's energy output to contribute to the pumping process.

Many collisional pumping schemes have put forward, based on excitation of OH by species such as H, e or H_2, but the chemical and physical conditions of the regions are too poorly known to choose between them.

The problem of OH formation has been addressed by several authors. Notably, Elitzur and de Jong (1978) consider its formation behind the HII region's advancing shock front. They show that atomic oxygen reacts to form water immediately behind the shock front, and this water then photodissociates into OH. Although this model has gained a degree of acceptance, there are several difficulties associated with it. Van Dishoeck and Dalgarno (1983) have shown that the OH produced by this process would itself quickly photodissociate in the same region where it was formed. Furthermore, much observational evidence suggests that OH maser emission comes from slowly moving material unaccelerated by the advancing shock, whereas this model has the OH being produced behind it.

In response to this, Hartquist and Dalgarno (1982) proposed an alternative model that attempts to resolve these problems. They pointed out that heated gas exists ahead of the shock front as a result of ambipolar diffusion, and then demonstrated how high OH concentrations might build up in such material. By forming the OH in unaccelerated material, this scheme avoids the dynamical problem faced by the Elitur-de Jong model. Their argument runs as follows.

The magnetic field outside the HII region will be compressed by the magnetic precursor of the MHD shock. As field lines straighten, ionised atoms will stream through neutral species, a process called ambipolar diffusion, resulting in heating of the gas. Now, the main neutral component of the ISM beyond the HII region is H_2, the main ion is S^+, so the heating will be mainly due to S^+ streaming through molecular hydrogen. Cooling will be largely by H_2 line emission, as pointed out by Black and Hartquist (1979), originally in the context of the Elitzur-de Jong model.

Also, energetic S^+ will be removed by the reaction

$$S^+ + H_2 \rightarrow SH^+ + H \qquad (4.3.1)$$

$$SH^+ \rightarrow \text{rapid recombination}$$

So, since S^+ is depleted as the magnetic precursor advances through it, its density will be highest at the front of the precursor, and lowest at its rear. Therefore the highest heating rate due to ion-neutral streaming, and hence the highest temperature, will be at the front of the precursor. This point should be noted, as it will now be shown that such a high temperature region is where the OH would form. That is, it will form well ahead of the accelerated material.

The model considers a volume of hot, say 1000 K, gas surrounded by large column densities of cool gas and bathed in UV radiation. H_2 in the cool gas will be in its lowest vibrational state, and ultraviolet radiation at the wavelengths of the Lyman and Werner transitions from this level will be absorbed. Those wavelengths corresponding to transitions from higher vibrational states are, however, free to pass unattenuated through the cool gas. When this flux reaches the hot, vibrationally excited gas, rapid photodissociation of H_2 takes place across a length of about 10^{14} cm, beyond which the gas becomed self-shielding. But those H_2 molecules in their ground vibrational state will be unaffected by this process, since the UV flux that would excite them has already been blocked out by the cool gas.

This model therefore creates a large amount of atomic hydrogen, while preserving a substantial population of molecular hydrogen in the same region thanks to self-shielding of its ground vibrational state by the cool surrounding gas. H and H_2 densities will be of the same order of magnitude in the 10^{14} cm wide region over which the transitions from excited vibrational states become self-shielding.

Three oxygen-hydrogen reactions are important in hot gas;

$$O + H_2 \rightleftharpoons OH + H \qquad (4.3.2)$$

$$O_2 + H_2 \rightleftharpoons 2OH \qquad (4.3.3)$$

$$OH + H_2 \rightleftharpoons H_2O + H \qquad (4.3.4)$$

It turns out from these reactions that the amount of OH produced will be greatest when the atomic and molecular hydrogen densities are equal. That is, in the model considered here the greatest OH densities will occur in the 10^{14} cm wide region discussed above. This length is typical of OH maser spot diameters.

As already mentioned, ambipolar diffusion provides a mechanism for producing the volumes of hot gas required here. Cooling will take place by line emission from rotational and vibrational-rotational transitions of H_2, and by OH line emission (Black and Hartquist, 1979).

The predicted presence of detectable molecular hydrogen emission from OH maser sources offers the opportunity of observationally testing the model by measuring any such emission. This was the motivation for the observing program that will now be discussed.

4.4 The H_2 Observations

4.4.1 The OH Maser Source W75N(OH)

W3(OH) and W75N(OH) are among the brightest and best studied Type I OH maser sources. Both show marked similarities to the star-forming region in Orion, in that they contain compact HII regions associated with infrared sources and H_2O and OH maser emission. They were the first two Type I OH sources to be mapped by

VLBI techniques. W75N(OH) was chosen for this study because W3(OH) is too far north (dec. = +61°) to be observable from UKIRT.

W75N is a part of the Cygnus X HII complex. This association of large molecular clouds, situated at a distance of 2 kpc (Dickel et al., 1969) contains the cluster of HII regions in W75N, the compact HII region DR21, H_2O maser sources, the OH maser sources W75N(OH) and W75S(OH), and a number of infrared, radio and molecular emission sources.

A map of CO emission (Dickel et al., 1978) shows the complex to be composed of two distinct molecular clouds. The northerly one, with a radial velocity of +9 km s^{-1}, is associated with W75N. The other, twenty arcminutes to the south, has a radial velocity of -3 km s^{-1} and contains W75S and DR21. Dickel et al. have suggested that the interaction between the two clouds is responsible for the star-forming activity in the region.

The OH maser source itself, W75N(OH), was discovered independently by Rydbeck et al. (1969) and Zuckerman et al. (1969). With a total flux density at 1665 MHz of 30 Jy (Haschick et al., 1981) it is one of the strongest type I OH maser sources. The maser emission has been studied interferometrically by several observers (e.g. Harvey et al., 1974, Haschick et al., 1981, Baart et al., 1986). A consistent view of the source's kinematics has not emerged. Some observers find evidence for a rotating disk, while others believe the maser spots are embedded in a radial outflow. The magnetic field in the region is of the order of 3 mG, high enough to "soften" the shock structure.

Continuum observations at 6 cm (Habing et al., 1974) have found the position of the OH maser source to coincide with that of an ultra-compact HII region, HII(B). Another compact HII region, HII(A), brighter and more extended that the first, lies about thirty arcseconds to the north-east of HII(B) (Harris, 1974). A detailed re-observation of W75N in the 1665 MHz maser line and in the radio continuum (Haschick et al., 1981) has confirmed the positional coincidence, to within 0.5", of the maser source and the HII region HII(B).

W75N is a source of extended far-infrared emission. The 20 μm emission peaks near HII(A) (Wynn-Williams et al., 1974), while at 53 μm the peak lies closer to HII(B) and the maser source (Harvey et al., 1977).

Molecular hydrogen emission from the region has not been mapped in detail, but a single CVF measurement towards HII(A) (Fischer et al., 1980) did detect thermally excited H_2 at that position. Taken with measurements of the IR continuum (Harvey et al., 1977) a value for the S(1) to continuum emission ratio of S(1)/IR = 3×10^{-5} is obtained. This is similar to the value in Orion, S(1)/IR = 1.2×10^{-5}, and is characteristic of star-forming regions in general (Joseph et al., 1984).

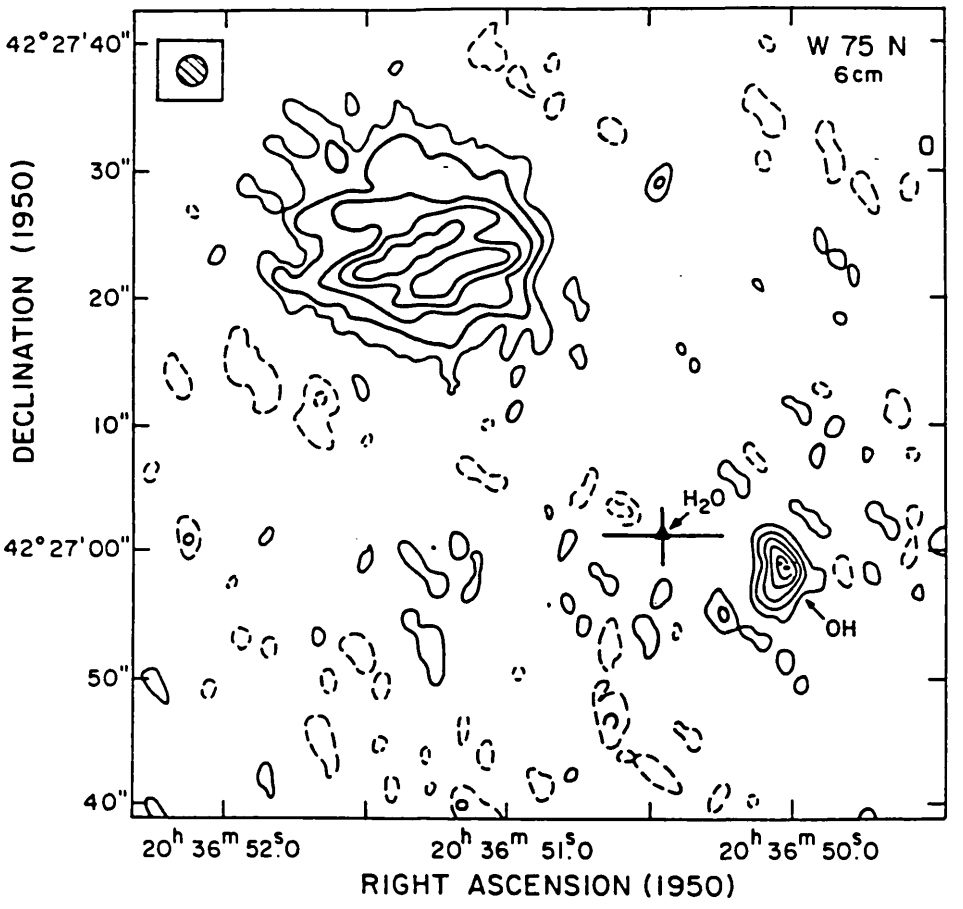


FIGURE 4.2 A 6 cm Map of W75N (Habing et al., 1974)
The H_2O maser position is plotted from Genzel and Downes, (1977). The contour levels are -1, -0.5, 0.5, 1, 2, 3 and 4 mJy per beam area. The northerly source is HII(A). The more compact southerly source, coincident with the OH maser position, is HII(B).

4.4.2 Observations of the H_2 S(1) Line in W75N(OH)

Black and Hartquist (1979), in their paper discussing possible H_2 line emission in OH maser regions, carried out a detailed calculation of the fluxes expected from several rotational and vibrational-rotational lines of H_2 at various temperatures, using the W3(OH) source as an example. They concluded that these lines should be detectable in the nearest OH maser sources, and that the line widths would be comparable to the radial velocity dispersions of the maser features, that is, 10 - 30 km s^{-1}, distinguishing H_2 emission in material heated by ambipolar diffusion from that heated by shocks, which typically has line widths of about 100 km s^{-1}.

The predicted v= 1-0 S(1) line fluxes from W3(OH) are shown in Table 4.1.

The calculations assumed a magnetic field of 5×10^{-3} gauss for the region, (as determined by Moran et. al., 1978) and that the molecular hydrogen emission came from a shell of radius 7.0×10^{16} cm and thickness 7×10^{14} cm, as inferred from interferometric observations.

T(K)	n(cm^{-3})	Flux(W cm^{-2})
1100	9×10^{5}	5.6×10^{-20}
700	2×10^{6}	1.5×10^{-21}
400	4×10^{6}	4.9×10^{-25}
250	9×10^{6}	7.4×10^{-30}

TABLE 4.1 Predicted S(1) Line Flux in W3(OH) (Black and Hartquist, 1979).

A full test of the model would require that several lines be measured to establish a temperature for the gas, and that at least one line be velocity resolved so that an investigation of the kinematics of the region could be carried out.

The results presented here were obtained as preliminary observations for a program of this kind. The data presented here were obtained as a service observation at the 3.8 m United Kingdom Infrared Telescope on the 17th July, 1984.

The instrument used was a Fabry-Perot interferometer operated in tandem with a CVF photometer. The Fabry-Perot etalon has a nominal resolution of 100 km s^{-1} FWHM, though this was broadened to 120 km s^{-1} for this observation because of the wide 18" aperture used. The CVF photometer was the UKIRT instrument called UKT9. This near-IR photometer uses an InSb detector and has a Circular Variable Filter for the 2 μm band and a set of standard photometric filters.

The signal was modulated by frequency switching the Fabry-Perot etalon. The etalon gap was switched between a reference position and a position on the H_2 line. The reference wavelength was chosen to be free of emission lines from the sky and the source.

Frequency chopping has advantages over sky chopping in instances where a source is (or is suspected to be) so extended that an uncontaminated sky reference cannot be obtained within the throw of the chopper. The technique is particularly well suited to Fabry-Perot spectrometers, where it can be straightforward to modulate the etalon gap. Frequency chopping does not cancel telluric emission features, however, so its use is confined to spectral regions relatively free from atmospheric emission lines.

The observations were made at the position of the OH maser group in W75N(OH) as given by Haschick et al. (1981),

i.e. R.A.(1950) = $20^h\ 36^m\ 50^s.0$, dec.(1950) = 42° 26' 58".

Four points spaced 10 arcseconds apart were first measured along a north-south line at the centre wavelength of the H_2 line to examine the extension of the source and to locate the emission peak. The source appeared point-like at the resolution of the 18"

beam, and the emission maximum coincided with the OH source to within the 2" pointing error of the telescope.

A spectral scan of the line was then made at the position of the maximum. A spectral interval centred on 2.1218 μm was sampled at sixteen points 36.6 km s^{-1} apart. The 1-0 S(1) line of H_2 was therefore sampled at about six points and a nearly equal length of continuum was measured on either side of it. The integration time was 4 s per point.

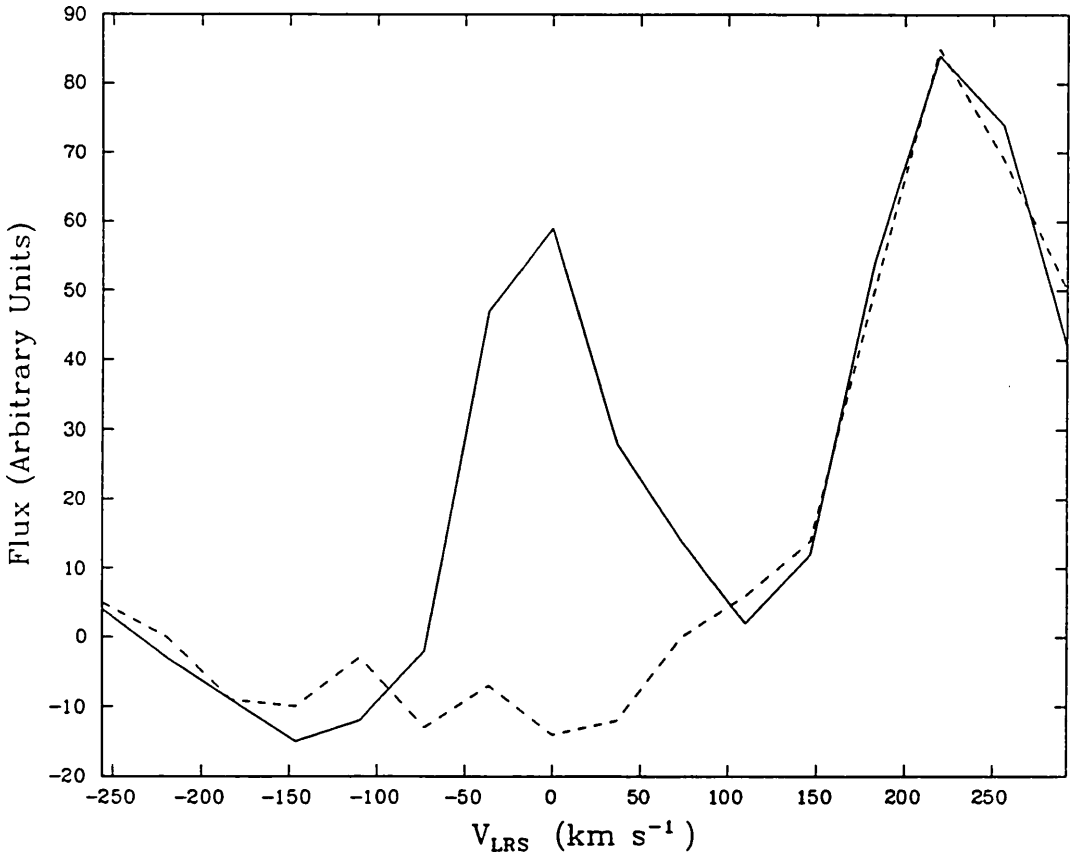


FIGURE 4.3 The Observed Spectrum Uncorrected for Atmospheric Emission.
The solid line is the source+sky spectrum, the dashed line is the spectrum of the sky alone.

The telluric emission spectrum was removed by scanning an empty field 300" north of the source and subtracting it from the source spectrum. Figure 4.3 shows the source spectrum and the sky spectrum, figure 4.4 shows their difference, that is, the H_2 line with the telluric emission feature removed. The local rest standard was defined relative to the centre of the H_2 line.

Spectra of the S(1) line in the planetary nebula NGC 7027 were taken before and after the observation of W75N(OH). These served to check the instrument set-up and were also used to calibrate the data.

The measured flux in the 1-0 S(1) line was $(4.3 \pm 0.8) \times 10^{-20}$ W cm^{-2}.

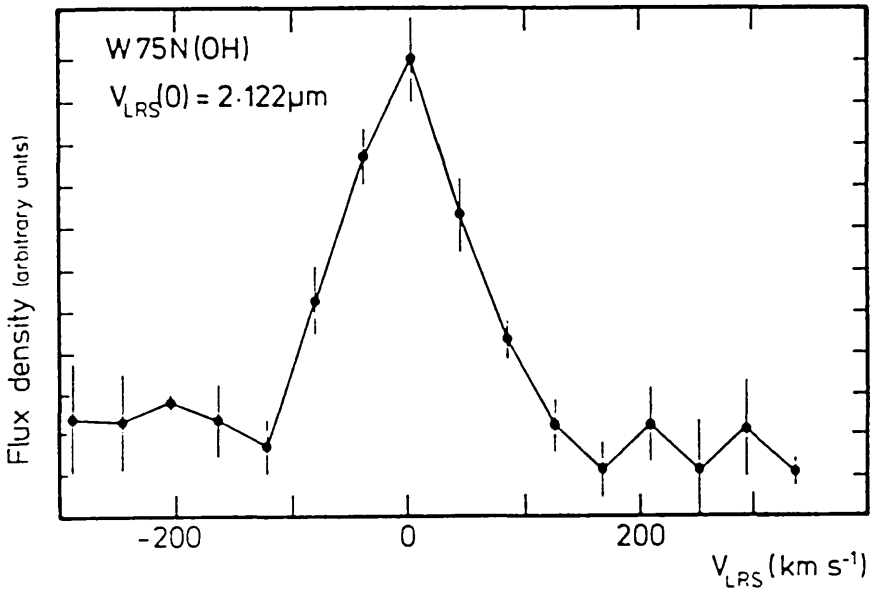


FIGURE 4.4 The Observed H_2 Spectrum (Mountain et al., 1985).

4.4.3 Extinction Correction

The extinction towards W75N(OH) may be estimated from the emission measure at radio and visible wavelengths (Soifer and Pipher, 1975). The intensity of optically thin free-free emission is proportional to the emission measure, defined as $E = N_e^2\sigma$, where N_e is the electron density and σ is the thickness of the nebula along the line of sight.

The emission measure of HII(B) at 4.9 GHz is 9×10^{-5} cm^{-6} pc (Haschick et al., 1981). The Palomar E plate (λ = 6500 Å) of the region shows it to be heavily obscured. No nebulosity is visible at the position of HII(B). The minimum emission measure detectable on the Palomar plates is ~ 50 cm^{-6} pc (Soifer and Pipher, 1975). Hence a lower limit of $A_v > 12.8$ mag. is obtained, which corresponds to $A_K > 1.1$ mag. if the van de Hulst curve No.15 (Johnson, 1968) is adopted.

A better estimate of the extinction may be obtained from the far-IR observations of Harvey et al. (1977). These authors measured the optical depth at 53 μm to be unity at the position of the 53 μm emission peak in W75N. This peak roughly coincides with the position of the OH maser source.

A λ^{-1} extinction law over the range 0.5 μm - 100 μm is observed in a variety of sources (e.g. Harvey et al., 1979). Assuming such a law in this case gives A_K = 4.5 mag.

If the S(1) line is reddened by this amount, the dereddened flux is 3×10^{-18} W cm^{-2}.

4.4.4 Comparison with Model

In the model proposed by Hartquist and Dalgarno (1982), both the H_2 emission and OH maser emission are produced in a thin shell of material, ahead of the shock front, where the H and H_2 number densities are approximately equal. The thickness of the shell is comparable to the diameter of individual maser spots, while the shell radius can be obtained from the extent of the region over which the maser spots are distributed.

The flux f in an H_2 emission line received from a uniform thin shell of density n, thickness r and radius R, where $R \gg r$, is given by

$$f = n\epsilon r (R/d)^2 \qquad (4.4.1)$$

where d is the distance of the observer and ϵ is the emissivity of the transition (Black and Hartquist, 1979).

For W75N, the data of Haschick et al. (1981) suggest that the shell of H_2 emission would be 1.4 arcsec in diameter and 0.01 arcsec thick.

Using the relation above to scale the W3(OH) results of Black and Hartquist (1979) to W75N(OH), an S(1) flux from the shell of $F = 1.7 \times 10^{-19}$ W cm^{-2} at T = 700 K or $F = 1.8 \times 10^{-18}$ W cm^{-2} at T = 1100 K is predicted.

Comparison with the observed flux suggests that, if the H_2 emission does originate in the maser source, the gas temperature is probably about 1000 K.

From observations of Λ-doublet lines in absorption towards W75N, Guilloteau et al. (1984) estimated the temperature of the gas in front of the compact HII region to be of the order of 100 K. Some of these observations were made with a very large (65 arcsec) beam, however, and some line ratios were found to differ from the expected LTE optically thin ratio, possibly because the lines are at different excitation temperatures. This casts doubt on the rotation temperatures between doublets derived from those observations.

While the observations presented in this chapter are not sufficient to show that the maser spots are in the hot gas, they do illustrate the importance of shocks in the vicinity of OH masers. The observations show that in this star forming region, as in Orion, H_2 line emission is associated with an OH maser source.

The low surface brightness of H_2 line emission sources can make them difficult to find. OH masers may prove to be useful indicators of the activity that gives rise to H_2 emission in star-forming regions.

Subsequent more detailed observations of W75N have altered the conclusions presented above.

Broad band mapping at 2 μm, spectroscopic mapping of molecular hydrogen line emission (Moore et al., 1988), and 2 μm polarimetry (Yamashita et al., 1988) have revealed that the W75N region contains an infrared reflection nebula illuminated by the ultracompact HII region W75N(B). The molecular hydrogen emission is not confined to the vicinity of the HII region, but has a bipolar spatial distribution almost two arcminutes in extent. The H_2 emission is associated with the CO outflow source, and H_2 line ratio measurements suggest the emission may be powered in part by shock heating and in part by fluorescence, the illuminating sources probably being the two compact HII regions HII(A) and HII(B).

References

Abgrall, H., Roueff, E. and Viala, Y. 1982. *Astr. Astrophys. Suppl. Ser.*, **50**, 505.

Allen, D.A. and Barton, J.R. 1981. *Pub. A.S.P.*, 93, 381.

Anantharamiah, K.R. and Radhakrishnan, V. 1979. *Astr. Astrophys.*, **79**, L9.

Andresen, P. 1986. *Astr. Astrophys.*, **154**, 42.

Baart, E.E., Cohen, R.J., Davies, R.D., Norris, R.P. and Rowland,P.R. 1986. *Mon. Not. R. astr. Soc.*, **219**, 145

Bailey, J.A. and Hough, J.H. 1982. *Pub. A.S.P.*, **94**, 618.

Beckman, J.E. 1985. Birth and Infancy of Stars, ed. R. Lucas, A. Omont and R. Stora, p 433.

Beckman, J.E., Watt, G.D., White, G.J., Phillips, J.P., Frost, R.L. and Davis, J.H. 1982. *Mon. Not. R. astr. Soc.*, **201**, 357.

Beckwith, S., Persson, S.E., Neugebauer, G. and Becklin, E.E. 1978. Astrophys. J., 223, 464. Beer, R. and Taylor, F.W. 1973. *Astrophys. J.*, **179**, 309.

Beichman, C.A., Neugebauer, G., Habing, H.J., Clegg, P.E., Chester, T.J. 1984. Explanatory Supplement to the IRAS Catalogs and Atlases. U.S. Govt. Print. Off.

Black, D.C. 1972. *Geochim. Cosmochim. Acta*, 36, 347.

Black, J.H. and Dalgarno, A. 1976. *Astrophys. J.*, **203**, 132.

Black, J.H. and Hartquist, T.W. 1979. *Astrophys. J.*, **323**, L179.

Blackwell, D.E., Petford, A.D. and Shallis, M.J. 1980. *Astr. Astrophys.*, **82**, 249.

Blitz, L. and Heiles, C. 1987. *Astrophys. J.*, **313**, L95.

Boggess, N.W. 1984. *Laser Focus/Electro-Optics*, June 1984, 116.

Cesarsky, D.A., Moffet, A.T. and Pasachoff, J.M. 1973. *Astrophys. J.*, **180**, L1.

Chelli, A. and Mariotti, J.-M. 1986. *Astr. Astrophys.*, **157**, 372.

Chernoff, D.F., Hollenbach, D.J. and McKee, C.F. 1982. *Astrophys. J. Lett.*, **259**, L97.

Colgate, S.A. 1974. *Astrophys. J.*, **187**, 321.

Combes, F., Boulanger, F., Encrenaz, P.J., Gerin, M., Bogey, M., Demuynck, C. and Destombes, J.L. 1985. *Astr. Astrophys.*, **147**, L25.

Cook, A.H. 1977. Celestial Masers. Cambridge University Press.

Dalgarno, A. and Lepp, S. 1984. *Astrophys. J.*, **287**, L47.

di Benedetto, G. and Conti, G. 1983. *Astrophys. J.*, **268**, 309.

Dickel, J.R., Dickel, H.R. and Bieritz, J.H. 1978. *Astrophys. J.*, **223**, 840.

Dickel, H.R., Wendker, H.J. and Bieritz, J.H. 1969. *Astr. Astrophys.*, **1**, 170.

Downes, D., Genzel, R., Becklin, E.E. and Wynn-Williams, C.G. 1981. *Astrophys. J.*, **244**, 869.

Draine, B.T. and Roberge, W.G. 1982. *Astrophys. J. Lett.*, **259**, L91.

Draine, B.T., Roberge, W.G. and Dalgarno, A. 1983. *Astrophys. J.*, **264**, 485.

Durie, R.A. and Herzberg, G. 1960. *Can. J. Phys.*, **38**, 806.

Elitzur, M. 1982. *Rev. Mod. Phys.*, **54**, 1225.

Elitzur, M. and de Jong, T. 1978. *Astr. Astrophys.*, **67**, 323.

Epchtein, N., Le Bertre, T., Lepine, J.R.D., Marques dos Santos, P., Matsuura, O.T. and Picazzio, E. 1987. *Astr. Astrophys. Supp. Ser.*, **71**, 39.

Fellgett., P. 1958. *J. Phys. Rad.*, **19**, 187.

Fischer, J., Righini-Cohen, G. and Simon, M. 1980. *Astrophys. J.*, **238**, L155.

Fowler, A.M., Probst, R.G., Britt, J.P., Joyce, R.R. and Gillett, F.C. 1987. *Opt. Eng.*, **26**, 232.

Gautier, T.N., Fink, U., Treffers, R.R. and Larson, H.P. 1976. *Astrophys. J.*, **207**, L129.

Geiss, J. and Reeves, H. 1972. *Astr. Astrophys.*, **18**, 126.

Genzel, R. and Downes, D. 1977. *Astr. Astrophys. Supp. Ser.*, **30**, 145.

Glass, I.S., Catchpole, R.M. and Whitelock, P.A. 1987. *Mon. Not. R. astr. Soc.*, **227**, 373.

Goldreich, P. and Kwan, J. 1974. *Astrophys. J.*, **191**, 93.

Grasdalen, G.L. and Joyce, R.R. 1976. *Bull. A.A.S.*, **8**, 349.

Grevesse, N. 1970. *Colloque de Liege*, **19**, 251.

Guelin, M., Langer, W.D. and Wilson, R.W. 1982. *Astr. Astrophys.*, **107**, 107.

Guilloteau, S., Baudry, A., Walmsley, C.M., Wilson, T.L. and Winnberg, A. 1984. *Astr. Astrophys.*, **131**, 45.

Guilloteau, S., Omont, A. and Lucas, R. 1987. *Astr. Astrophys.*, **176**, L24.

Habing, H.J., Goss, W.M., Matthews, H.E. and Winnberg, A. 1974. *Astr. Astrophys.*, **35**, 1.

Hall, D.N.B., Aikens, R.S., Joyce, R. and McCurnin, T.W. 1975. *Applied Optics*, **14**, 450.

Harris, S. 1974. *Mon. Not. R. astr. Soc.*, **166**, 29p.

Hartquist, T.W. 1979. *Astr. Astrophys.*, **77**, 361.

Hartquist, T.W. and Dalgarno, A. 1982. ESA Spec. Publ. **192**, 29.

Hartquist, T.W., Mountain, C.M., Selby, M.J. and Roueff, E. 1982. ESA Spec. Publ. **192**, 45.

Harvey, P.J., Booth, R.S., Davies, R.D., Whittet, D.C.B. and McLaughlin, W. 1974. *Mon. Not. R. astr. Soc.*, **169**, 545.

Harvey, P.M., Campbell, M.F. and Hoffmann, W.F. 1977. *Astrophys. J.*, **211**, 786.

Harvey, P.M., Thronson, H.A. Jr. and Gatley, I. 1979. *Astrophys. J.*, **231**, 115.

Haschick, A.D., Reid, M.J., Burke, B.F., Moran, J.M. and Miller, G. 1981. *Astrophys. J.*, **244**, 76.

Hasegawa, T. 1986. *Astrophys. Space Science*, **118**, 421.

Herzberg, G. 1950. *Nature*, **166**, 563.

Ho, P.T.P., Haschick, A.D., Vogel, S.N. and Wright, M.C.H. 1983. *Astrophys. J.*, **265**, 295.

Howarth, I.D. and Maslen, D. 1983. SERC Starlink User Note No. 50.

Hoyle, F. and Fowler, W.A. 1973. *Nature*, **241**, 384.

Huber, K.P. and Herzberg, G. 1979. Constants of Diatomic Molecules.

IRAS Science Team. 1986. *Astr. Astrophys. Supp. Ser.*, **65**, 607.

Jennings, R.E. 1986. *Q. Jl. R. astr. Soc.*, **27**, 454.

Johnson, H.L. 1962. *Astrophys. J.*, **135**, 69.

Johnson, H.L. 1964. *Bol. Ton. y Tac.*, **3**, 305.

Johnson, H.L. 1965. *Comm Lunar and Planetary Lab.*, 3, 73.

Johnson, H.L. 1966. *Ann. Rev. Astron. Astrophys.*, **5**, 193.

Johnson, H.L. 1968. Nebulae and Interstellar Matter, ed. B.M. Middlehurst and L.H. Aller. University of Chicago Press.

Johnson, H.L., Mitchell, R.I., Iriarte, B. and Wisniewski, W.Z. 1966. *Comm. Lunar and Planetary Lab.*, **4**, 99.

Jorden, P.R. Long, J.F., MacGregor, A.D. and Selby, M.J. 1976. *Astr. Astrophys.*, **49**, 421.

Joseph, R.D., Wright, G. and Wade, R. 1984. *Nature*, **311**, 13.

Kleinmann, S.G., Gillett, F.C. and Joyce, R.R. 1981. *Ann. Rev. Astron. Astrophys.*, **19**, 411.

Labeyrie, A. 1970. *Astr. Astrophys.*, **6**, 87.

Langer, W.D., Schloerb, F.P., Snell, R.L. and Young, J.S. 1980. *Astrophys. J.*, **239**, L125.

Lemke, D., Grewing, M., Offermann, D., Drapatz, S. and Klipping, G. 1983. *Adv. Space Res.*, **2**, 123.

Loren, R.B. and Wootten, A. 1985. *Astrophys. J.*, **299**, 947.

Low, F.J. 1961. *J. Opt. Soc. Am.*, **51**, 1300.

Low, F.J., Beichman, C.A., Gillett, F.C., Houck, J.R., Neugebauer, G., Langford, D.E., Walker, R.G. and White, R.H. 1984. *Opt. Eng.*, **23**, 122.

Low, F.J. and Rieke, G.H. 1974. Methods of Experimental Physics, 12, ed. N. Carleton, p 415.

Lynn, P.A. 1982. An Introduction to the Analysis and Processing of Signals. Macmillan Education.

Mariotti, J.-M., Zadrozny, A., Monin, J.-L. and Vauglin, I. 1987. Ground-Based Astronomical Observations With Infrared Array Detectors. ed. Becklin, Univ. of Hawaii Press.

McClatchey, R.A., Benedict, W.S., Slough, S.A., Burch, D.E., Calfee, R.F., Fox, K., Rothman, L.S. and Garing, J.S. 1973. AFCRL Atmospheric Absorption Line Parameter Compilation. AFCRL-TR-73-0096.

Menten, K.M., Walmsley, C.M., Henkel,C. and Wilson, T.L. 1986. *Astr. Astrophys.*, **157**, 318.

Moore, T.J.T., Mountain. C.M., Yamashita, T. and Selby, M.J. 1988. *Mon. Not. R. astr. Soc.*, **234**, 95.

Moran, J.M., Burke, B.F., Barrett, A.H., Rogers, A.E.E., Carter, J.C., Ball, J.A. and Cudaback, D.D. 1968. *Astrophys. J. Lett.*, **152**, L97.

Moran, J.M., Reid, M.J., Lada, C.J., Yen, J.L., Johnston, K.J. and Spencer, J.H. 1978. *Astrophys. J. Lett.*, **224**, L67.

Mountain, C.M. 1983. Ph.D Thesis, Imperial College, London University.

Mountain, C.M., Selby, M.J., Zadrozny, A. and Hartquist, T.W. 1985. *Mon. Not. R. astr. Soc.*, **213**, 13p.

Neugebauer, G., Habing, H.J., van Duinen, R., Aumann, H.H., Baud, B., Beichman, C.A., Beintema, D.A., Boggess, N., Clegg, P.E., de Jong, T., Emerson, J.P., Gautier, T.N., Gillett, F.C., Harris, S., Hauser, M.G., Houck, J.R., Jennings, R.E., Low, F.J., Marsden, P.L., Miley, G., Olnon, F.M., Pottasch, S.R., Raimond, E., Rowan-Robinson, M., Soifer, B.T., Walker, R.G., Wesselius, P.R. and Young, E. 1984. *Astrophys. J.*, **278**, L1.

Neugebauer, G. and Leighton, R.B. 1969. Two Micron Sky Survey. NASA SP-3047.

Ostriker, J.P. and Tinsley, B.M. 1975. *Astrophys. J.*, **201**, L51.

Papoular, R. 1982. *Astr. Astrophys.*, **117**, 46.

Penzias, A.A., Wannier, P.G., Wilson, R.W. and Linke, R.A. 1977. *Astrophys. J.*, **211**, 108.

Penzias, A.A. 1979. *Astrophys. J.*, **228**, 430.

Phillips, T.G., Blake, G.A., Keene, J., Woods, R.C. and Churchwell, E. 1985. *Astrophys. J.*, **294**, L45.

Plambeck, R.L., Wright, M.C.H., Welch, W.J., Bieging, J.H., Baud, B., Ho, P.T.P and Vogel, S.N. 1982. *Astrophys. J.*, **259**, 617.

Price, S.D. and Murdock, T.L. 1983. The Revised AFGL Infrared Sky Survey Catalog. AFGL-TR-83-0161.

Price, S.D. and Walker, R.G. 1976. The AFGL Four Color Infrared Sky Survey. AFGL-TR-76-0208.

Reeves, H., Fowler, W.A. and Hoyle, F. 1970. *Nature*, **226**, 727.

Reid, M.J. and Moran, J.M. 1981. *Ann. Rev. Astron. Astrophys.*, **19**, 231.

Ridgeway, S.T. and Brault, J.W. 1984. *Ann. Rev. Astron. Astrophys.*, **22**, 291.

Ridgeway, S.T., Wells, D.C. and Joyce, R.R. 1977. *Astron. J.*, **82**, 414.

Roddier, F. and Lena, P. 1984. *J. Optics*, **15**, 363.

Rodriguez-Kuiper, E.N., Zuckerman, B. and Kuiper, T.B.H. 1978. *Astrophys. J. Lett.*, **219**, L49.

Rogerson, J.B. Jr. and York, D.G. 1973. *Astrophys. J. Lett.*, **186**, L95.

Rydbeck, O.E.H., Ellder, J. and Kollberg, E. 1969. *Astrophys. J. Lett.*, **156**, L141.

Sarma, N.V.G. and Mohanty, D.K. 1978. *Mon. Not. R. astr. Soc.*, **184**, 181.

Schloerb, F.P., Snell. R.L., Langer, W.D. and Young, J.S. 1981. *Astrophys. J. Lett.*, **251**, L37.

Schmidtke, P.C., Africano, J.L., Jacoby, G.H., Joyce, R.R. and Ridgeway, S.T. 1986. *Astron. J.*, **91**, 961.

Schramm, D.N. and Wagoner, R.V. Dec. 1974. *Physics Today*, p 41.

Seeley, J.S. (ed.) 1985. Proc. Soc. Photo-Opt. Instrum. Eng., 589, 166-249.

Selby, M.J., Wade, R. and Sanchez-Magro, C. 1979. *Mon. Not. R. astr. Soc.*, **187**, 553.

Shull, J.M. 1978. *Astrophys. J.*, **219**, 877.

Sibille, F., Chelli, A. and Lena, P. 1979. *Astr. Astrophys.*, **79**, 315.

Simon, T., Morrison, D. and Cruickshank, D.P. 1972. *Astrophys. J. Lett.*, **177**, L17.

Sinton, W.M. 1986. *Pub. A.S.P.*, **98**, 246.

Smith, D., Adams, N.G. and Alge, E. 1982. *Astrophys. J.*, **263**, 123.

Smith, D.S., Wimmers, J.T., Hermann, J.A. and Bailey, G.C. 1982. *Proc. SPIE*, **331**, 29.

Soifer, B.T.,Houck, J.R. and Harwit, M. 1971. *Astrophys. J. Lett.*, **168**, L73.

Soifer, B.T. and Pipher, J.L. 1975. *Astrophys. J.*, **199**, 663.

Soifer, B.T. and Pipher, J.L. 1978. *Ann. Rev. Astron. Astrophys.*, **16**, 335.

Spitzer, L., Drake, J.F., Jenkins, E.B., Morton, D.C., Rogerson, J.B. and York, D.G. 1973. *Astrophys. J.*, **181**, L116.

Stecher, T.P. and Williams, D.A. 1967. *Astrophys. J. Lett.*, **149**, L29.

Teich, M.C. 1968. *Proc. I.E.E.E.*, **56**, 37.

Tokunaga, A.T. 1984. *Astron. J.*, **89**, 172.

Townes, C.H. 1984. *J. Astrophys. Astr.*, **5**, 111.

Traub, W.A. and Stier, M.T. 1976. *Appl. Optics*, **15**, 364.

Trauger, J.T., Roesler, F.L., Carleton, N.P. and Traub, W.A. 1973. *Astrophys. J.*, **184**, L137.

Trefler, M. and Gush, H.P. 1968. *Phys. Rev. Lett.*, **20**, 703.

Turner, J., Kirby-Docken, K. and Dalgarno, A. 1977. *Astrophys. J. Supp. Ser.*, **35**, 281.

Turner, B.E. and Zuckerman, B. 1978. *Astrophys. J.*, **225**, L75.

van Dishoeck, E.F. and Dalgarno, A. 1983. *J. Chem. Phys.*, **79**, 873.

Vidal-Madjar, A., Laurent, C., Gry, C., Bruston, P., Ferlet, R. and York, D.G. 1983. *Astr. Astrophys.*, **120**, 58.

Walker, R.G. and Price, S.D. 1975. AFCRL Infrared Sky Survey, Vol. 1. AFCRL-TR-75-0373.

Wagoner, R.V., 1973. *Astrophys. J.*, **179**, 361.

Watson, D.M. 1982. Sub-Millimetre Wave Astronomy. ed. J.Beckman and J.P.Phillips, p 91.

Watson, W.D., 1976. *Rev. Mod. Phys.*, **48**, 513.

Watson, W.D. and Walmsley, C.M. 1982. Regions of Recent Star Formation, ed. R.S.Roger and P.E.Dewdney, p 357.

Weinreb, S. 1962. *Nature*, **195**, 367.

Wick, G.C. 1935. *Atti. R. Accad. Naz. Lincei (Ser. 6)*, **21**, 708.

Wright, M.C.H., Plambeck, R.L., Vogel, S.N., Ho, P.T.P and Welch, W.J. 1983. *Astrophys. J. Lett.*, **267**, L41.

Wynn-Williams, C.G., Becklin, E.E. and Neugebauer, G. 1974. *Astrophys. J.*, **187**, 473.

York, D.G. and Rogerson, J.B. 1976. *Astrophys. J.*, **203**, 378.

Yamashita, T., Sato, S., Tamura, M., Suzuki, H., Gatley, I., Hough, J.H., Mountain, C.M. and Moore, T.J.T. 1988. *Mon. Not. R. astr. Soc.*, **233**, 899.

Zuckerman, B., Ball, J.A., Dickinson, C.F. and Penfield, H. 1969. *Astrophys. J. Lett.*, **3**, L97.

Appendix A:
The Four-Channel Photometer Photometry Program

This Appendix contains a listing of the photometry data reduction software used with the Imperial College Four Channel Photometer. The software is written in BASIC for an AIM 65 microcomputer. It provides real-time reduction of photometric data by the procedures described in Chapter Two.

```
LIST

 1 REM PHOT6 23.2.83
 2 REM PROGRAM OCCUPIES 15084 BYTES
 7 REM
 8 REM
 9 REM
10 REM  ****   *   *   ***   *****
11 REM  *   *  *   *  *   *    *
12 REM  *   *  *   *  *   *    *
13 REM  ****   *****  *   *    *
14 REM  *      *   *  *   *    *
15 REM  *      *   *  *   *    *
16 REM  *      *   *   ***     *
17 REM
18 REM  **************************
19 REM
20 REM
21 REM
30 REM  THIS PROGRAM IS INTENDED FOR USE WITH THE IMPERIAL COLLEGE
31 REM  FOUR-CHANNEL PHOTOMETER.
32 REM  IT USES THAT INSTRUMENT IN ITS PHOTOMETRIC MODE,INTEGRATING
33 REM  FLUX FROM A SINGLE SOURCE.
34 REM  (SEE PROGRAM "SCAN" FOR SCANNING MODE)
48 REM
49 REM
50 REM  OVER-VIEW OF PROGRAM:
51 REM
52 REM  LINE 100-999:  SET CONSTANTS,'DIM'STATEMENTS
53 REM  LINE 1000-1100:INITIALISE HARDWARE,SET DEFAULT VALUES
54 REM  LINE 1100-2000:INPUT PSD PARAMETERS,CHECK THAT SENSIBLE.
55 REM  LINE 2000-2200:ALTER DEFAULTS TO THESE INPUT VALUES.
56 REM  LINE 2200-3500:START PSDS,MOVE CAMERA,OPTIMISE PSD PHASE
58 REM  LINE 3500-4000:INPUT AND DISPLAY RUN PARAMETERS.
60 REM  LINE 4000-4400:READ DATA FROM COUNTERS
62 REM  LINE 4400-4500:NOD TELESCOPE,INVERT FLAGS,ETC
64 REM  LINE 4500-6000:  BLANK
66 REM  LINE 6000-6500:DO STATISTICS
68 REM  LINE 6500-6600:FINISH RUN: DO AIR-MASS ETC.
70 REM  LINE 6600-8000:  BLANK
71 REM                 SUBROUTINES:
72 REM  LINE 8000-8200:KEYBOARD HANDLER SUBROUTINE
73 REM  LINE 8200-8300:VIDEO DISPLAY SUBROUTINE
74 REM  LINE 8300-8400:DISPLAY INITIAL VIDEO TEXT
75 REM  LINE 8400-8500:CHECK WHETHER TO STOP NOW OR CONTINUE
76 REM  LINE 8500-8600:UPDATE DATA DISPLAY
77 REM  LINE 8600-8700:ENTER GAINS
78 REM  LINE 8700-8800:PRINT DATA ON REQUEST
80 REM  LINE 8800-9999:  BLANK
81 REM
82 REM
83 REM
100 DIM INIT(5),C(4),AV(4,3),SIG(4),S1D(4),S3NR(4),MAG(4),DMAG(4)
101 DIM GAIN(4),KINST(4),CD(6),N1(2)
102 DIM A1(4,2),SA(4,2),S2A(4,2),MEAN(4,2),SD(4,2)
104 CNTR=22016
105 PULSE=22055:LPULS=22044:SPULS=22042:S9PUL=22057
106 REM    HI          LO          LO          HI
107 REM 1ST OF 2 ADDRESSES IS GIVEN,AND WHETHER HI OR LO TIMER BYTE
108 KINST(1)=1:KINST(2)=1:KINST(3)=1:KINST(4)=1
110 RADEC=22061:GFLG=22062:PRNT=22063
112 BEAM=22064:DRDY=22053:NDFLG=22054:IMAX=22047
120 INIT(1)=0:INIT(2)=112:INIT(3)=148:INIT(4)=166:INIT(5)=240
```

```
143 REM
144 REM
145 REM-POKE X(1) TO 0005, X(2) TO 0004
146 REM
150 DIM PSD(2):PSD(1)=129:PSD(2)=40
155 DIM DEL(2):DEL(1)=129:DEL(2)=144
160 DIM TELND(2):TELND(1)=129:TELND(2)=185
165 DIM KEY(2):KEY(1)=129:KEY(2)=249
170 DIM PHASE(2):PHASE(1)=131:PHASE(2)=74
175 DIM TVMOV(2):TVMOV(1)=144:TVMOV(2)=0
180 DIM BLANK(2):BLANK(1)=130:BLANK(2)=96
181 REM
185 REM-INSERT ACTUAL GAINS HERE!!!
190 DATA 1,3,10,30,100,300,1000,0.01:REM-CH.1
193 DATA 1,3,10,30,100,300,1000,0.01:REM-CH.2
195 DATA 1,3,10,30,100,300,1000,0.01:REM-CH.3
197 DATA 1,3,10,30,100,300,1000,0.01:REM-CH.4
996 REM
997 REM
998 REM:INITIALISE HARDWARE AND SET DEFAULT VALUES
999 REM
1000 POKE 5,128
1005 FOR I=1 TO 5
1010 POKE 4,INIT(I)
1015 X=USR(I)
1020 NEXT I
1025 J=1: SN=10: SL=1: POKE BEAM,1: POKE GFLG,1
1030 REM
1031 REM
1049 REM
1050 GOSUB 8300:REM-SET UP VIDEO DISPLAY
1051 REM
1052 REM
1100 INPUT"DEFAULT VALUES(Y/N)";YN$
1105 IF YN$="Y" GOTO 2200
1110 IF YN$="N" GOTO 1200
1115 GOTO 1100
1116 REM
1117 REM
1118 REM
1180 REM-ENTER PSD AND CHOP PARAMETERS
1181 REM
1200 INPUT"CHOP FREQUENCY(HZ)";CHOP
1205 REM:CHECK IF SENSIBLE HERE
1210 IF CHOP>7.7 AND CHOP<40 GOTO 1230
1215 PRINT"THIS FREQUENCY IS OUT OF RANGE:RE-INPUT IT":GOTO 1200
1230 REM:MAINS SUBHARMONIC?ETC
1248 REM
1249 REM
1250 INPUT"SAMPLE LENGTHtM-CD NOW EVEN,EXACTLY 2 TIMES C1
2080 C2=RCHOP-CD:REM-ENABLE LONG PULSE
2085 REM
2086 REM
2089 RCHOP=RCHOP-2
2090 R1=INT(RCHOP/256):REM-CHOP PULSE MSB
2095 R2=RCHOP-(R1*256):REM-CHOP PULSE LSB
2100 POKE PULSE,R1:POKE PULSE+1,R2
2101 REM
2102 REM
2104 CD-CD-2
2105 R1=INT(CD/256):REM-SHORT ENABLE PULSE MSB
2110 R2=CD-(R1*256):REM-SHORT ENABLE PULSE LSB
2115 POKE SPULS+1,R1:POKE SPULS,R2
```

```
2116 REM
2117 REM
2119 C1=C1-2
2120 R1=INT(C1/256):REM-HALF SHORT PULSE MSB
2125 R2=C1-(R1*256):REM-HALF SHORT PULSE LSB
2130 POKE S9PUL,R1:POKE S9PUL+1,R2
2131 REM
2132 REM
2134 C2=C2-2
2135 R1=INT(C2/256):REM-LONG ENABLE PULSE MSB
2140 R2=C2-(R1*256):REM-LONG ENABLE PULSE LSB
2145 POKE LPULS+1,R1:POKE LPULS,R2
2146 REM
2147 REM
2148 REM
2150 POKE IMAX,INT(CHOP*SL*2):REM-SET INTEGRATION TIME
2151 REM
2155 POKE BEAM,1:IF UL$="L" THEN POKE BEAM,0
2156 REM
2200 REM
2205 POKE 5,PSD(1):POKE 4,PSD(2):X=USR(X):REM-START PSD S
2500 REM-SET UP TELESCOPE,TV,OPTIMISE PHASE HERE!!!
3498 REM
3499 REM
3500 REM-ENTER RUN PARAMETERS
3501 REM
3520 INPUT"OBJECT IDENTIFIER";OI$
3521 IF LEN(OI$)<30 GOTO 3525
3523 PRINT"TOO LONG":GOTO 3520
3524 REM
3525 INPUT"ENTER OBJECT'S R.A(HR,MIN,SEC)";H,M,S
3527 DEF FNA(X)=ABS(INT(X))
3530 IF H=FNA(H)AND M=FNA(M)AND H<24AND M<60AND S<60AND S>=0 GOTO 3550
3535 PRINT"THIS INPUT CONTAINS AN ERROR":GOTO 3525
3536 REM
3537 REM
3550 INPUT"ENTER OBJECT'S DEC.(DEG,MIN,SEC)";D1,M1,S1
3555 IF D1=INT(D1) AND ABS(D1)<90 AND M1=FNA(M1) AND M1<60 GOTO 3560
3557 PRINT"THIS INPUT CONTAINS AN ERROR":GOTO 3550
3560 IF S1>=0 AND S1<60 GOTO 3570
3567 PRINT"THIS INPUT CONTAINS AN ERROR":GOTO 3550
3570 IF D1>-42 AND D1<66 GOTO 3580
3575 PRINT"THIS DEC.VALUE NOT ACCESSIBLE FROM THIS SITE":GOTO 3550
3576 REM
3577 REM
3580 INPUT"'J' OR 'H' FILTER";JH$
3585 IF JH$<>"J" AND JH$<>"H" GOTO 3580
3586 REM
3587 REM
3590 GOSUB 8600:REM-ENTER GAINS
3600 INPUT"WHAT S/N RATIO DO YOU REQUIRE";SONR
3605 IF SONR>1 AND SONR<100 GOTO 3612
3610 PRINT"NOT REASONABLE;RE-ENTER(>1,<100)":GOTO 3600
3611 REM
3612 INPUT"ENTER DATE";DATE$
3614 REM
3615 INPUT"ENTER PRESENT SIDEREAL TIME(HR,MIN)";H2,M2
3620 IF H2=FNA(H2)AND M2=FNA(M2)AND H2<24AND M2<60 GOTO 3630
3625 PRINT"THIS INPUT CONTAINS AN ERROR":GOTO 3615
3626 REM
3627 REM
3628 REM:DISPLAY OBJ.IDENT.,CO-ORDS,U/L BEAM ON VIDEO
3629 REM
```

```
3630 VM$=OI$:PSN=20:GOSUB 8200
3635 VM$=STR$(H)+":"+STR$(M)+":"+STR$(S)
3640 PSN=54:GOSUB 8200
3645 VM$=STR$(D1)+":"+STR$(M1)+":"+STR$(S1)
3650 PSN=134:GOSUB 8200
3655 VM$="UPPER":IF UL$="L" THEN VM$="LOWER"
3660 PSN=82:GOSUB 8200
3661 REM
3665 IFJH$="J" GOTO 3680
3670 VM$="H"
3675 PSN=533:GOSUB 8200 :PSN=205:GOSUB 8200
3677 PSN=211:GOSUB 8200 :PSN=217:GOSUB 8200
3680 REM:CONTINUE
3685 FOR I1=1 TO 4
3690 VM$=MID$(STR$(GAIN(I1)),2,4)
3695 PSN=690 +(I1-1)*8:GOSUB 8200:NEXT I1
3700 REM
3996 REM
3997 REM
3998 REM:BEGIN OBSERVATION RUN
3999 REM
4000 INPUT"HIT A KEY TO START";DUMMY$
4040 REM
4045 FOR K=2 TO 3:REM:K=1 IS FROM THE PREVIOUS NOD.SEE LINE 6130
4046 REM
4047 VM$="         ":PSN=96:GOSUB 8200:REM:REMOVE DELAY ANNOUNCEMENT
4048 REM
4049 REM
4050 FOR L1=1 TO SN/2
4055 IF PEEK(DRDY)=1 GOTO 4070
4060 GOSUB 8000:REM-READ KEYBOARD
4065 GOTO 4055
4067 REM
4068 REM:THE EXECUTION TIME OF THE FOLLOWING LOOP LIMITS THE FSDS' MIN
4069 REM INTEGRATION TIME.IT TAKES 0.28 SEC TO RUN.
4070 FOR I=1 TO 4
4075 C(I)=PEEK(CNTR+I-1)*65536
4080 C(I)=C(I)+PEEK(CNTR+I+3)*256
4085 C(I)=C(I)+PEEK(CNTR+I+7)
4090 IF C(I)>8388608 THEN C(I)=C(I)-16777216
4095 AV(I,K)=AV(I,K)+C(I)*2/SN
4100 NEXT I
4101 REM
4102 REM
4105 POKE DRDY,0  :REM-RESET DATA READY FLAG
4106 REM
4110 NEXT L1
4111 REM
4120 NEXT K
4121 REM
4122 REM
4123 REM
4150 REM:REJECT THIS SET IF "R" HAS BEEN TYPED
4151 REM
4155 IF REJFL=0 GOTO 4200
4160 FOR K=2 TO 3:FOR I=1 TO 4:AV(I,K)=0:NEXT I:NEXT K
4162 PRINT"SET NO.";N1(1)+N1(2);"REJECTED"
4165 REJFL=0 :GOTO 4040 :REM:RE-DO THIS SET
4200 REM:CONTINUE
4398 REM
4399 REM
4400 POKE 5,TELND(1):POKE 4,TELND(2)
4405 X=USR(X): REM-NOD THE TELESCOPE
```

```
4406 REM
4407 REM
4410 POKE 5,DEL(1):POKE 4,DEL(2)
4415 X=USR(X): REM-START DELAY TIMER
4416 REM
4417 REM
4418 VM$="DELAY ON":PSN=96:GOSUB 8200:REM-ANNOUNCE DELAY
4419 REM
4420 REM-INVERT 'J' AND 'BEAM' FLAGS
4421 REM
4422 REM-J=1:ODD SET. J=2:EVEN SET
4430 IF J=1 GOTO 4440
4435 J=1:GOTO 4445
4440 J=2
4445 REM:CONTINUE
4446 REM
4447 REM
4448 REM: BEAM=0:NOD UP.  BEAM=1:NOD DOWN
4449 REM
4450 IF PEEK(BEAM)=1 GOTO 4460
4455 POKE BEAM,1:GOTO 4465
4460 POKE BEAM,0
4465 REM:CONTINUE
4466 REM
4467 REM
4470 VM$="UPPER":IF PEEK(BEAM)=0 THEN VM$="LOWER"
4475 PSN=82:GOSUB 8200
5998 REM
5999 REM
6025 IF N1(1)>0 OR J=1 THEN N1(J)=N1(J)+1
6026 REM:N1 IS NO.OF VALUES IN SET J.DON'T INCREMENT N1 ON 1ST.NOD
6027 REM
6029 REM
6030 FOR I=1 TO 4
6035 IF N1(1)=0 GOTO 6130:REM:DON'T DO CALCS ON 1ST NOD
6039 REM
6040 A1(I,J)=ABS(AV(I,1)-AV(I,2))
6045 SA(I,J)=SA(I,J)+A1(I,J)
6050 S2A(I,J)=S2A(I,J)+A1(I,J)*A1(I,J)
6051 REM
6054 REM
6055 MEAN(I,J)=SA(I,J)/N1(J)
6060 SD(I,J)=SQR(N1(J)*S2A(I,J)-SA(I,J)*SA(I,J))/N1(J)
6065 IF N1(1)<>N1(2) OR N1(2)<2 GOTO 6130:REM-DO STATS FROM NOD 5 ONLY
6066 REM
6067 REM COMBINE THE ODD AND EVEN SETS
6068 REM
6070 SIG(I)=(MEAN(I,1)+MEAN(I,2))/2
6075 S1D(I)=SQR(SD(I,1)*SD(I,1)+SD(I,2)*SD(I,2))/2
6080 S3NR(I)=SIG(I)/S1D(I)
6085 MAG(I)=KINST(I)*LOG(SL*GAIN(I)/SIG(I))
6090 REM
6094 DEF FNLA(X)=LOG(ABS(X))
6095 Y1=FNLA((SL*GAIN(I))/(SIG(I)+S1D(I)))
6100 Y2=FNLA((SL*GAIN(I))/(SIG(I)-S1D(I)))
6105 DMAG(I)=KINST(I)/2*(Y2-Y1)
6130 AV(I,1)=AV(I,3):AV(I,2)=0:AV(I,3)=0
6150 NEXT I
6151 REM
6155 IF N1(1)=0 GOTO 6300:REM:DON'T DO CALCS ON 1ST NOD
6200 REM DO COLOUR DIFFERENCES HERE
6210 FOR Q1=1 TO 3:COL(Q1)=MAG(1)-MAG(Q1+1):NEXT Q1
6215 FOR Q1=4 TO 5:COL(Q1)=MAG(2)-MAG(Q1-1):NEXT Q1
```

```
6220 COL(6)=MAG(3)-MAG(4)
6243 REM
6244 REM
6245 GOSUB 8500:REM-DISPLAY DATA ON VIDEO
6247 GOSUB 8700:REM PRINT DATA ON REQUEST
6250 GOSUB 8400:REM-STOP?
6255 IF SFLG=1 GOTO 6500
6260 REM
6300 IF PEEK(NDFLG)=1 GOTO 4040
6305 GOSUB 8000:REM READ KEYBOARD
6310 GOTO 6300
6497 REM
6498 REM
6499 REM
6500 REM-FINISH RUN
6501 REM
6510 INPUT"ENTER PRESENT SIDEREAL TIME(HR,MIN)";H3,M3
6515 INPUT"REL.HUMIDITY(%)";RH:INPUT"SEEING(ARCSEC)";SG
6516 REM
6517 REM
6520 REM-DO AIR MASS ETC
6521 REM
6525 RA=H+M/60 +S/3600
6530 THETA=(28+18/60)/57.29578
6535 DEC=(D1+M1/60+S1/3600)/57.29578
6536 REM
6540 ST(1)=H2+M2/60:REM:INITIAL SIDEREAL TIME
6545 ST(2)=H3+M3/60:REM:FINAL SIDEREAL TIME
6550 IF ST(2)<ST(1) THEN ST(2)=ST(2)+24:REM-PASSED THROUGH 2400 HRS?
6551 HST=INT(ST(2)-ST(1))
6552 MST=INT((ST(2)-ST(1)-HST)*60)
6555 Z1=0:IF ST(1)<RA AND ST(2)>RA THEN Z1=1
6556 REM-IE-Z1=1 IF THIS RUN PASSED THROUGH ZENITH
6557 REM
6558 REM
6560 Q1=COS(THETA)*COS(DEC):Q2=SIN(THETA)*SIN(DEC)
6565 DEF FNB(X)=1/((Q1*COS((X-RA)*.26179939))+Q2)
6566 REM
6567 REM: PRINT OUT DATA FOR THIS RUN
6568 REM
6569 REM
6570 PRINT"":PRINT"":PRINT"":PRINT"":PRINT"":PRINT""
6575 PRINT"  OBJECT IDENTIFIER: "+OI$:PRINT""
6577 PRINT"TODAY'S DATE:"+DATE$:PRINT"":PRINT""
6580 PRINT"     R.A:"+STR$(H)+":"+STR$(M)+":"+STR$(S)
6585 PRINT"     DEC:"+STR$(D1)+":"+STR$(M1)+":"+STR$(S1)
6590 PRINT"":PRINT""
6595 PRINT"INITIAL SIDEREAL TIME:"+STR$(H2)+" HR"+STR$(M2)+" MIN"
6597 PRINT"  FINAL SIDEREAL TIME:"+STR$(H3)+" HR"+STR$(M3)+" MIN"
6600 PRINT""
6605 PRINT"ELAPSED RUN TIME:"+STR$(HST)+" HRS."+STR$(MST)+" MIN"
6610 PRINT"":PRINT"":PRINT""
6615 PRINT"  INITIAL AIR MASS=";FNB(ST(1))
6620 PRINT"    FINAL AIR MASS=";FNB(ST(2))
6625 IF Z1=1 THEN PRINT"AIR MASS AT ZENITH:";FNB(RA)
6630 PRINT"":PRINT""
6635 PRINT"RELATIVE HUMIDITY="+STR$(RH)+"%"
6640 PRINT"           SEEING="+STR$(SG)+" ARCSEC."
6645 PRINT"":PRINT"":PRINT"":PRINT""
6650 IF JH$="J" THEN PRINT TAB(21);"CH.1(J)";
6655 IF JH$="H" THEN PRINT TAB(21);"CH.1(H)";
6660 PRINT "     CH.2(K)     CH.3(L)     CH.4(M)"
6665 PRINT TAB(4);"AMPLIFIER GAIN";
```

```
6670 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);GAIN(I);:NEXT I:PRINT''
6675 PRINT TAB(12);'SIGNAL';
6680 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);SIG(I);:NEXT I:PRINT''
6685 PRINT TAB(10);'ST. DEV.';
6690 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);SID(I);:NEXT I:PRINT''
6695 PRINT TAB(9);'S/N RATIO';
6700 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);S3NR(I);:NEXT I:PRINT''
6705 PRINT TAB(9);'MAGNITUDE';
6710 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);MAG(I);:NEXT I:PRINT''
6715 PRINT'ERROR IN MAGNITUDE';
6720 FOR I=1 TO 4:PRINT TAB(20+(I-1)*12);DMAG(I);:NEXT I:PRINT''
6725 PRINT'':PRINT TAB(10);'NO.OF SETS:';N1(1)+N1(2)
6730 REM:PRINT COLOUR DIFFERENCES
6740 PRINT'':PRINT'':PRINT''
6750 IF JH$='J' THEN PRINT TAB(23);'J-K        J-L        J-M';
6755 IF JH$='H' THEN PRINT TAB(23);'H-K        H-L        H-M';
6760 PRINT TAB(53)'K-L        K-M        L-M'
6765 PRINT'COLOUR DIFFERENCES';
6770 FOR I=1 TO 6
6775 COL(I)=VAL(LEFT$(STR$(COL(I)),7))
6780 PRINT TAB(20+(I-1)*10);COL(I);
6785 NEXT I:PRINT''
7992 REM
7993 REM
7994 REM
7995 STOP:REM-SUBROUTINES LISTED BELOW
7996 REM
7997 REM
7998 REM- KEYBOARD HANDLER ROUTINE
8000 GET KB$  :REM-CAN'T PAUSE HERE-MUST USE 'GET',NOT 'INPUT'
8005 IF KB$='G' THEN POKE GFLG,1
8010 IF KB$='S' THEN POKE GFLG,0
8015 IF KB$='P' THEN POKE PRNT,1
8020 IF KB$='N' THEN POKE PRNT,0
8025 IF KB$='L' THEN POKE BEAM,1
8030 IF KB$='U' THEN POKE BEAM,0
8035 IF KB$='R' THEN REJFL=1
8050 RETURN
8198 REM
8199 REM
8200 REM-VIDEO MESSAGE DISPLAY SUBROUTINE
8201 REM-REQUIRES VM$ AND PSN(=1 TO 1120) ON ENTRY
8202 REM-VM$ IS THE MESSAGE,PSN IS ITS START POSITION ON THE SCREEN.
8205 LVM=LEN(VM$):PSN=INT(PSN)
8210 IF(PSN+LVM<=1120) AND(PSN>=1) GOTO 8220
8215 PRINT'MESSAGE'',VM$,'' NOT DISPLAYED:'
8216 PRINT'POSITION IS OUT OF RANGE':RETURN
8220 FOR I=1 TO LVM
8224 REM-16622 IS THE START OF VIDEO RAM
8225 POKE 16622+PSN+I,ASC(MID$(VM$,I,1))
8230 NEXT I
8235 RETURN
8298 REM
8299 REM
8300 REM ENTER INITIAL VIDEO MESSAGES
8310 VM$='CH.1:J  CH.2:K  CH.3:L  CH.4:M':PSN=528:GOSUB 8200
8315 VM$='HEX.COUNT':PSN=598:GOSUB 8200
8320 VM$='ANALOG SCALE;GAIN':PSN=670:GOSUB 8200
8325 VM$='SIGNAL':PSN=761:GOSUB 8200
8330 VM$='ST.DEV':PSN=841:GOSUB 8200
8335 VM$='S/N RATIO':PSN=918:GOSUB 8200
8340 VM$='MAGNITUDE':PSN=998:GOSUB 8200
8342 VM$='MAG.ERROR':PSN=1078:GOSUB 8200
```

```
8344 VM$="R.A:":PSN=50:GOSUB 8200
8345 VM$="DEC:":PSN=130:GOSUB 8200
8350 VM$="OBJECT IDENTIFIER:":PSN=2:GOSUB 8200
8355 VM$="BEAM":PSN=88:GOSUB 8200
8360 VM$="COL.DIFF":PSN=275:GOSUB 8200
8365 VM$="J-K   J-L   J-M   K-L   K-M   L-M":PSN=205:GOSUB 8200
8390 RETURN
8391 REM
8392 REM
8400 REM-THIS ROUTINE CHECKS WHETHER TO STOP NOW
8410 X1=0:SFLG=0
8415 FOR I=1 TO 4
8420 IF S3NR(I)>SONR THEN X1=X1+1
8425 NEXT
8430 X2=PEEK(GFLG)
8435 IF X1<4 AND X2=1 THEN RETURN:REM-CONTINUE
8440 IF N1(1)=N1(2)+1 AND N1(1)>=3 GOTO 8450
8441 REM:IE ONE MORE U THAN L BEAM,AT LEAST THREE U BEAMS REQUIRED
8445 PRINT"WRONG BEAM OR NOT ENOUGH SETS;RUN WILL CONTINUE":RETURN
8450 INPUT"DO YOU WANT TO CONTINUE RUN;OR FINISH(C/F)";CF$
8455 IF CF$<>"C" AND CF$<>"F" GOTO 8550
8460 IF CF$="F" GOTO 8475
8465 INPUT"HIT A KEY TO CONTINUE";DUMMY$
8470 PRINT"PAUSE OCCURRED AFTER NOD";N1(1)+N1(2):RETURN
8475 SFLG=1:RETURN
8498 REM
8499 REM
8500 REM DATA DISPLAY ROUTINE
8501 REM
8510 IF N1(2)<2 THEN RETURN:REM-NO DISPLAY IF NOT ENOUGH SETS
8511 REM
8514 REM-DISPLAY DATA
8515 FOR I1=1 TO 4
8520 VM$=LEFT$(STR$(SIG(I1)),5):PSN=768+(I1-1)*8:GOSUB 8200
8525 VM$=LEFT$(STR$(S1D(I1)),5):PSN=848+(I1-1)*8:GOSUB 8200
8530 VM$=LEFT$(STR$(S3NR(I1)),5):PSN=928+(I1-1)*8:GOSUB 8200
8535 VM$=LEFT$(STR$(MAG(I1)),5):PSN=1008+(I1-1)*8:GOSUB 8200
8540 VM$=LEFT$(STR$(DMAG(I1)),5):PSN=1088+(I1-1)*8:GOSUB 8200
8545 NEXT I1
8550 VM$="":FOR I1=1 TO 6:VM$=VM$+LEFT$(STR$(COL(I1)),5)+" ":NEXT I1
8555 PSN=284:GOSUB 8200:REM:DISPLAY COLOUR DIFFERENCES
8560 RETURN
8598 REM
8599 REM
8600 REM
8601 REM-ENTER AMPLIFIER GAINS
8605 PRINT"ENTER GAINS AS TWO NUMBERS;FIRST THE GAIN SET ON"
8606 PRINT"THE DIAL,THEN WHETHER THIS IS DIVIDED BY 1 OR 100."
8607 PRINT"SEPARATE THE NUMBERS BY A COMMA."
8608 REM
8610 FOR I=1 TO 4
8615 PRINT"CHANNEL"+STR$(I);:INPUT G0,G1
8620 G9=0
8621 IF G0=1 THEN G9=1
8622 IF G0=3 THEN G9=2
8623 IF G0=10 THEN G9=3
8624 IF G0=30 THEN G9=4
8625 IF G0=100 THEN G9=5
8626 IF G0=300 THEN G9=6
8627 IF G0=1000 THEN G9=7
8630 IF G9=0 GOTO 8615:REM IE IF G0 IS NONE OF THE ABOVE,RE-ENTER IT
8635 IF G1<>1 AND G1<>100 GOTO 8615
8640 FOR Q1=1 TO G9:READ GAIN(I):NEXT Q1
```

```
 8643 IF G9=7 GOTO 8650
 8645 FOR Q1=G9+1 TO 7:READ DUMMY:NEXT Q1
 8650 READ DIV:IF G1=100 THEN GAIN(I)=GAIN(I)*DIV
 8655 NEXT I
 8660 RETURN
 8700 REM:PRINT DATA TO TELETYPE IF PRINT-FLAG=1
 8705 IF PEEK(PRNT)=0 THEN RETURN
 8710 PRINT "SET NO.";N1(1)+N1(2)
 8715 PRINT TAB(12);"CH.1        CH.2        CH.3        CH.4"
 8720 PRINT"   SIGNAL";
 8725 FOR I=1 TO 4:PRINT TAB(12+(I-1)*12);SIG(I);:NEXT I:PRINT""
 8730 PRINT"S/N RATIO";
 8735 FOR I=1 TO 4:PRINT TAB(12+(I-1)*12);S3NR(I);:NEXT I:PRINT""
 8740 POKE PRNT,0:RETURN
 9997 REM
 9998 REM
 9999 PRINT"END":END


LIST
```

Appendix B:

The Near Infrared Spectra of Seventeen Bright Stars

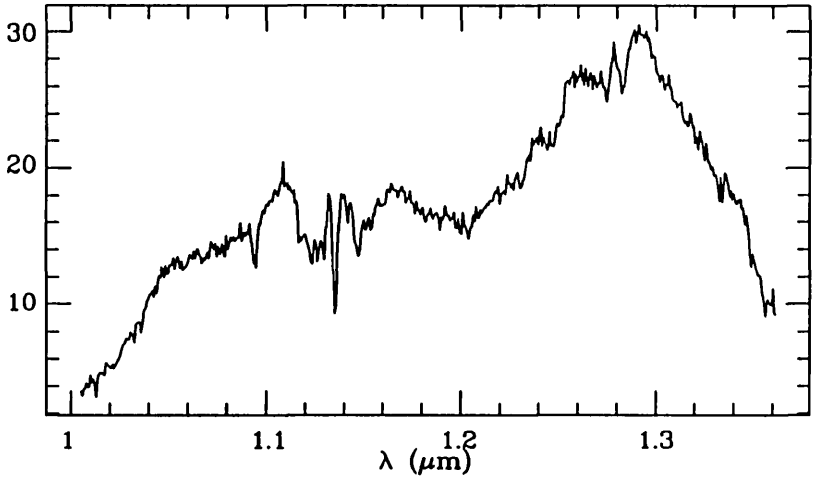
This Appendix presents the J, H, K and L' spectra of seventeen bright stars. The stars were observed with the Imperial College Cooled Grating Spectrometer at the 1.5 m Infrared Flux Collector, Tenerife, in June and July of 1985. The data were reduced by the methods described in Chapter Two.

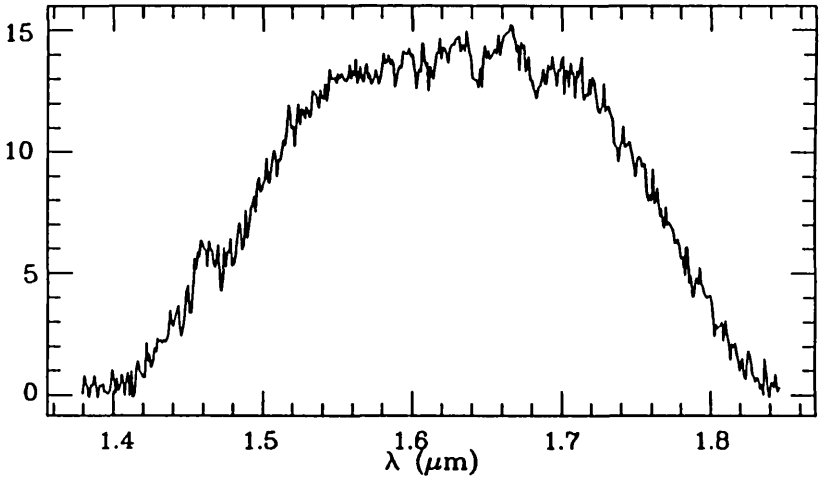
The fluxes of the spectra displayed here are expressed in instrumental units. The fluxes have been scaled to unit instrumental gain, and corrected for the effects of atmospheric extinction. They have been extrapolated to one airmass.

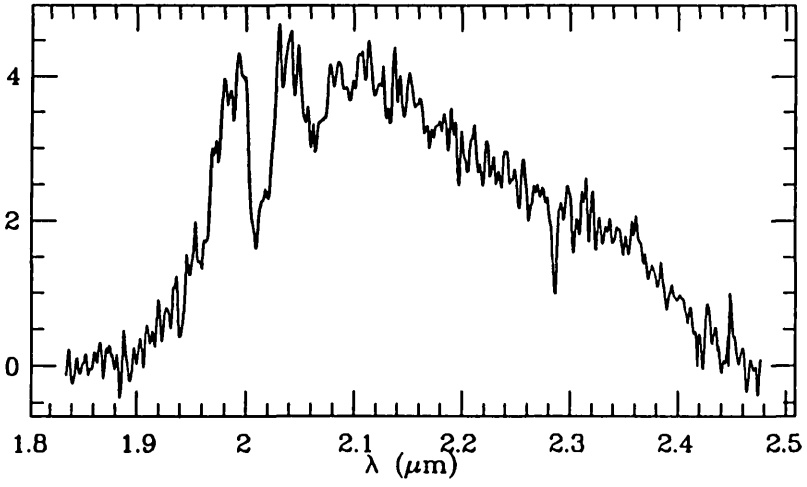
BS 15

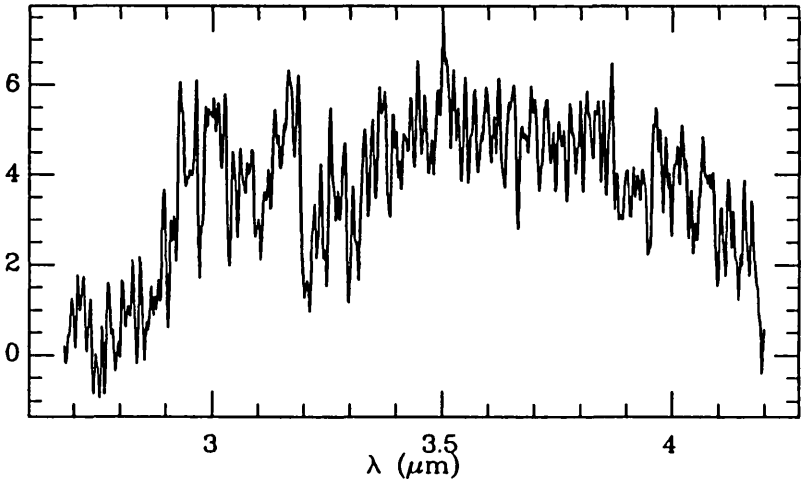
R.A.00^h03^m dec.+28°32′

B8 V m_v=2.1





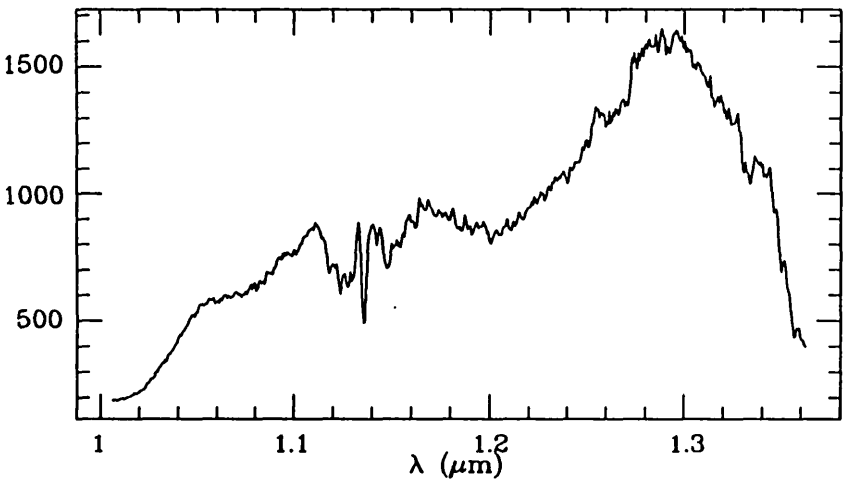


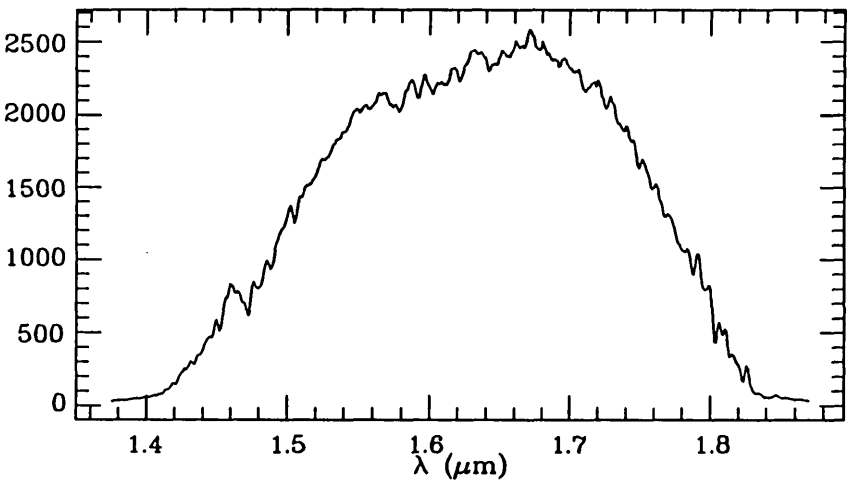


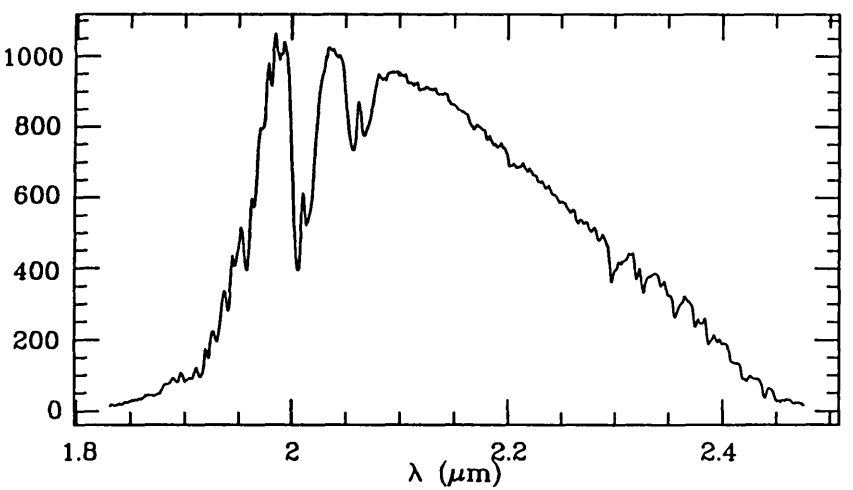
BS 5340

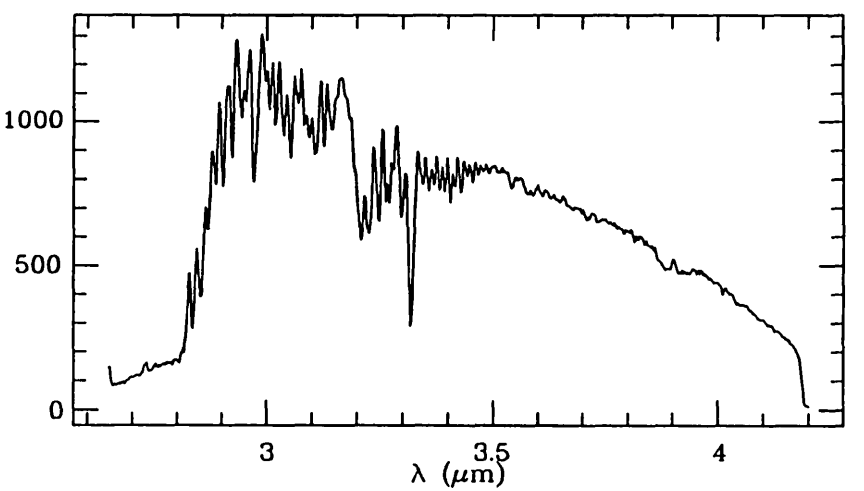
R.A.14^h16^m dec.+19°11′

K1 III m_v=0.0





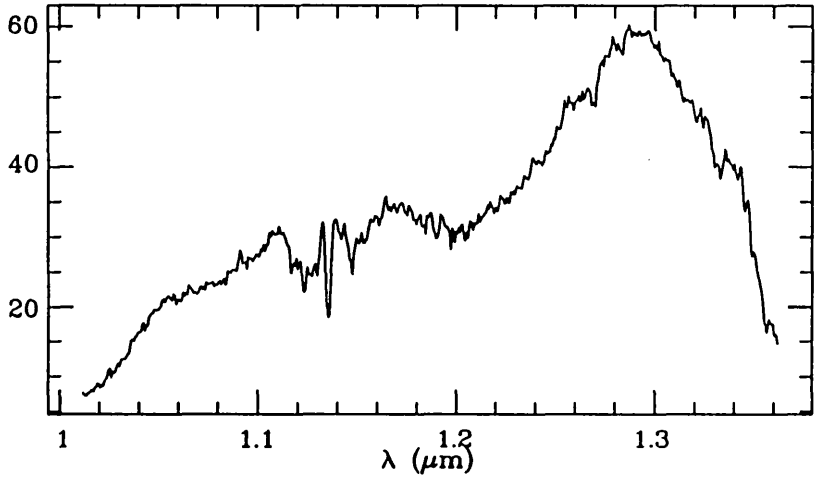


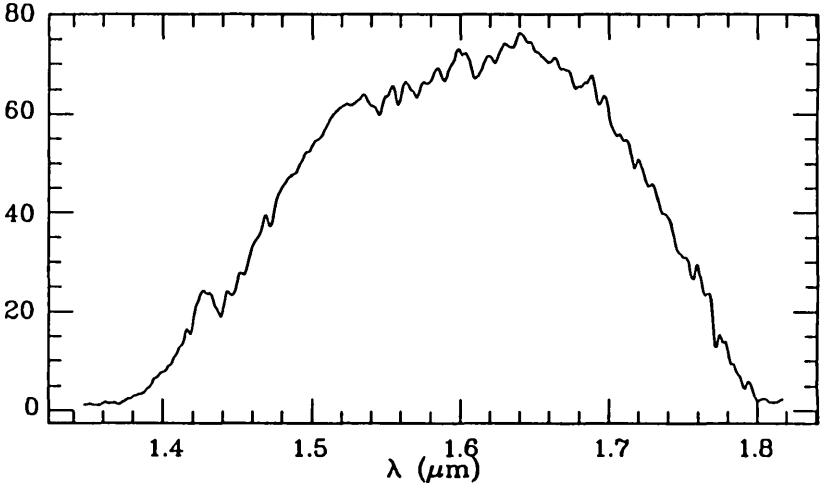


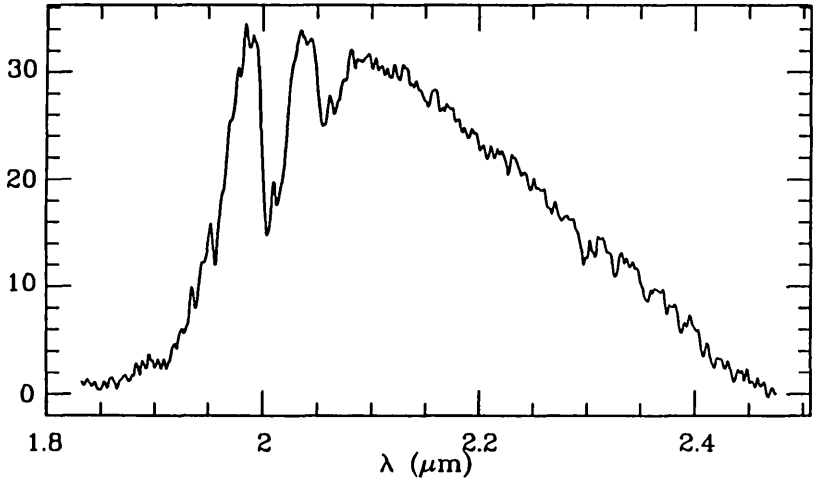
BS 5429

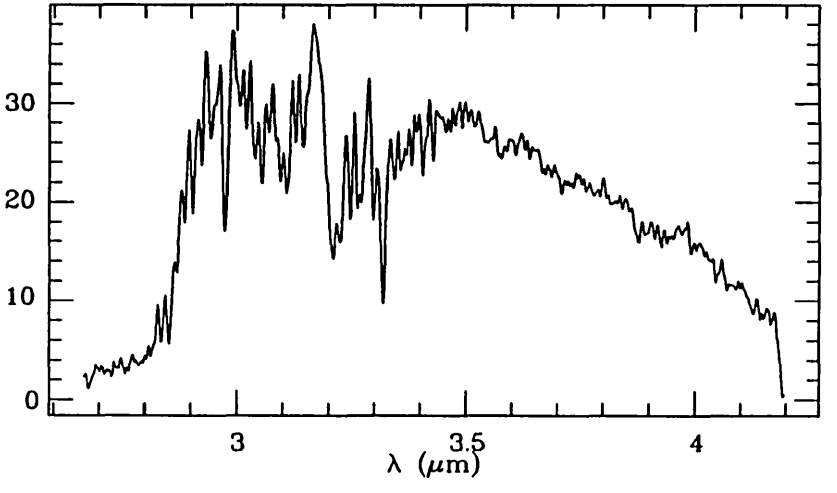
R.A.14^h32^m dec.+30°22′

K3 III m_v=3.6





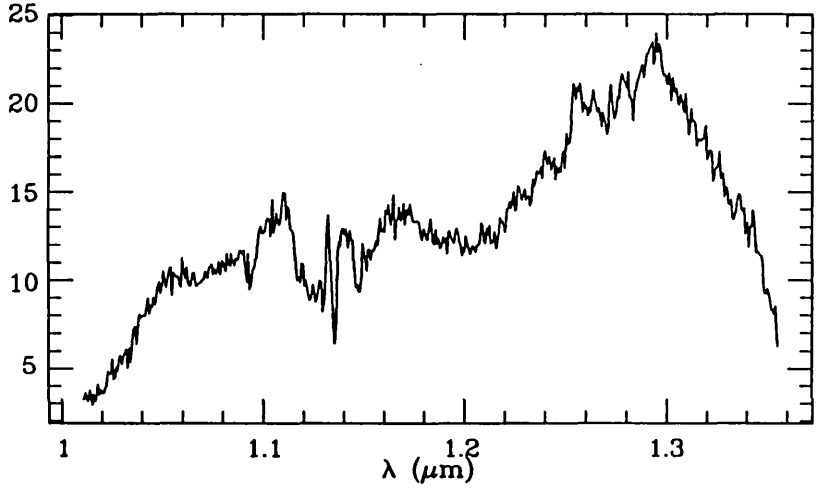


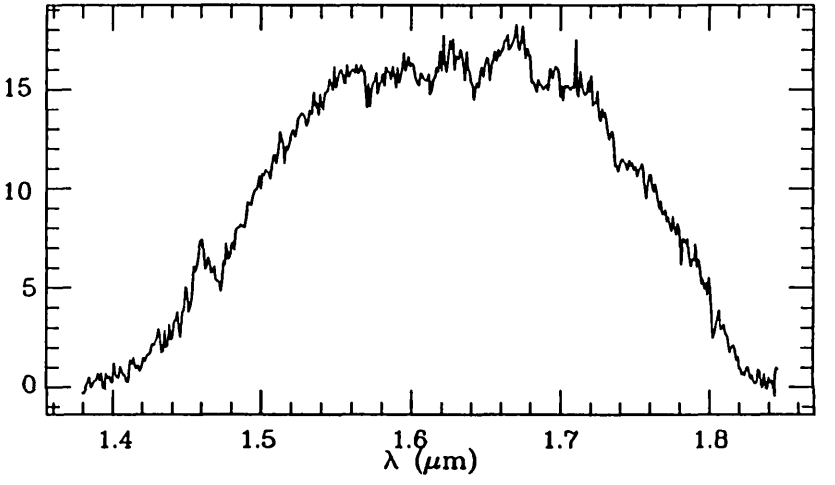


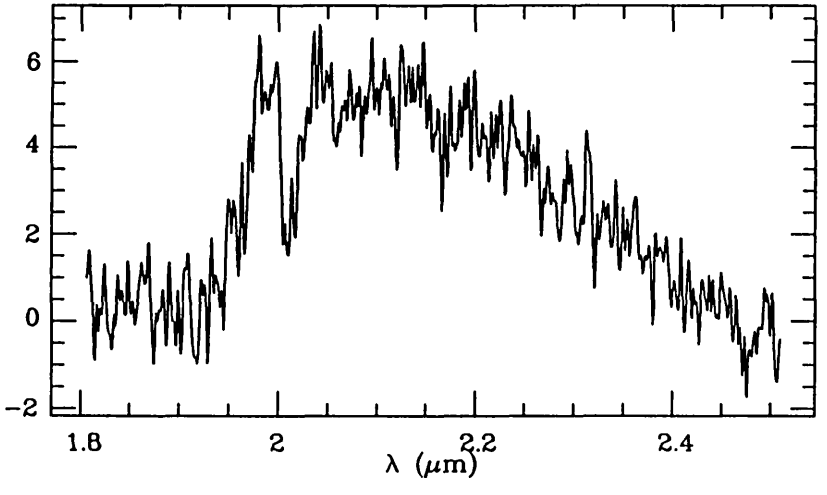
BS 5531

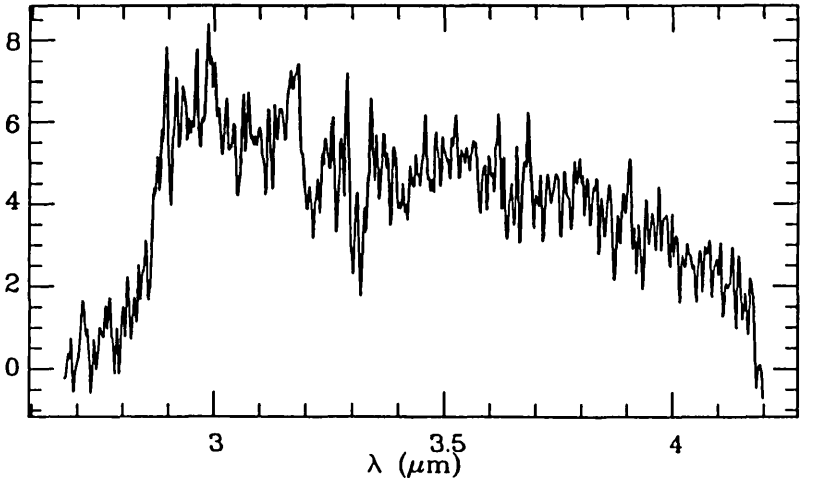
R.A.14^h50^m dec.$-15°58'$

A2 IV m_v=2.8



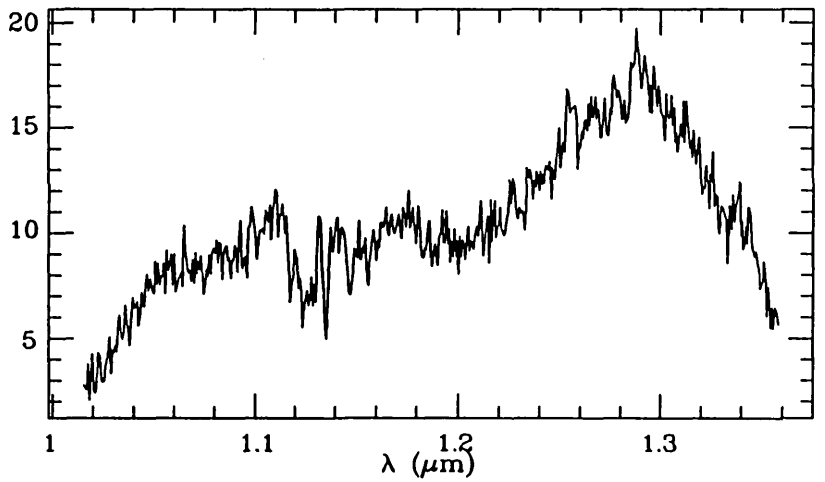


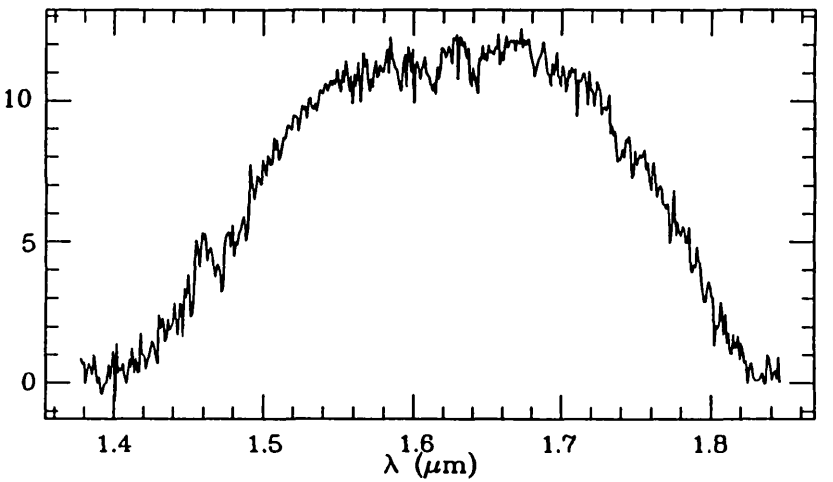


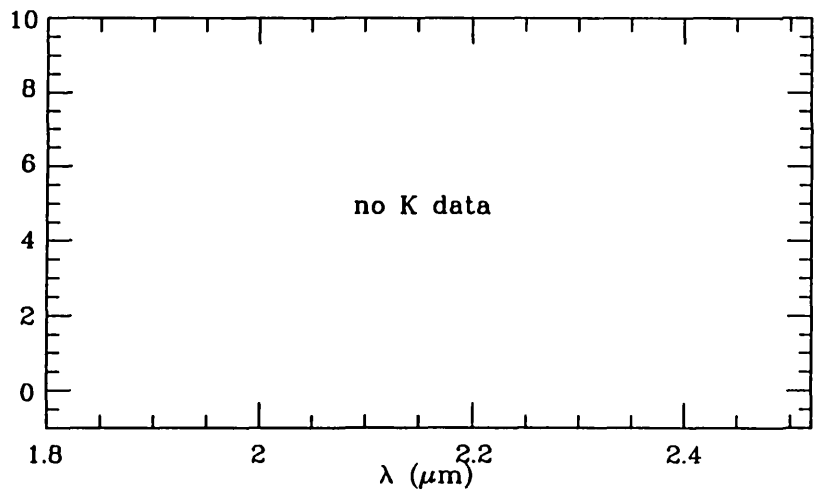


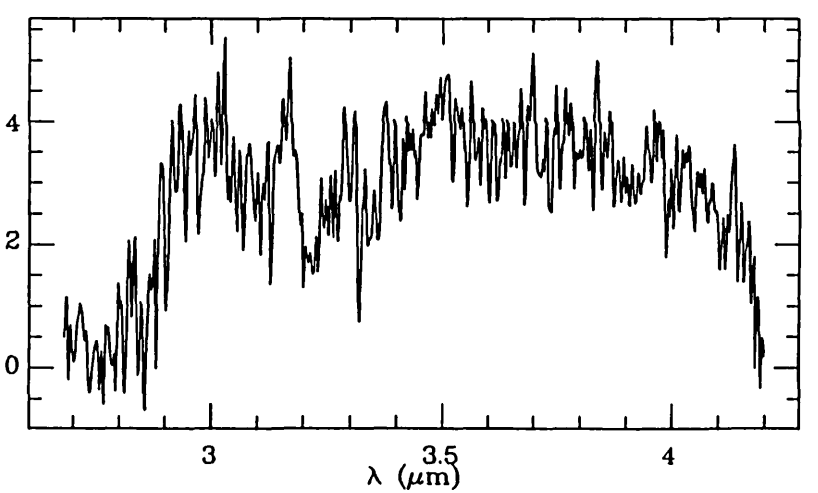
BS 5685

R.A.15^h16^m dec.– 9°19′
B8 V m_v=2.8



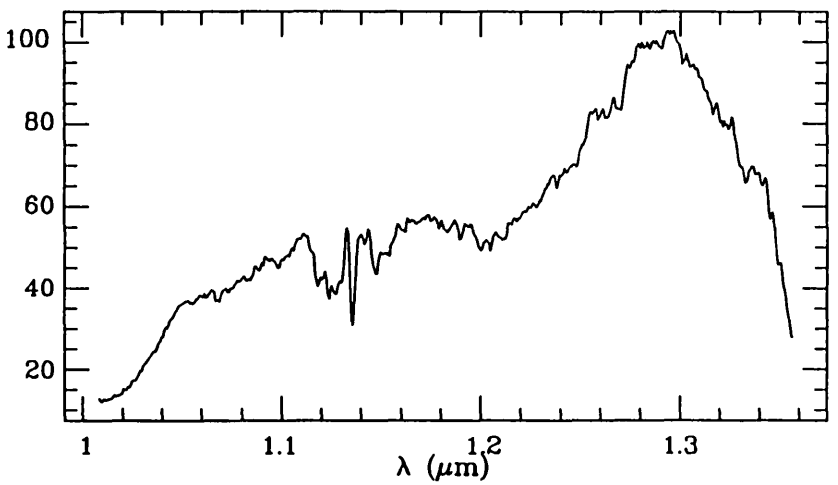


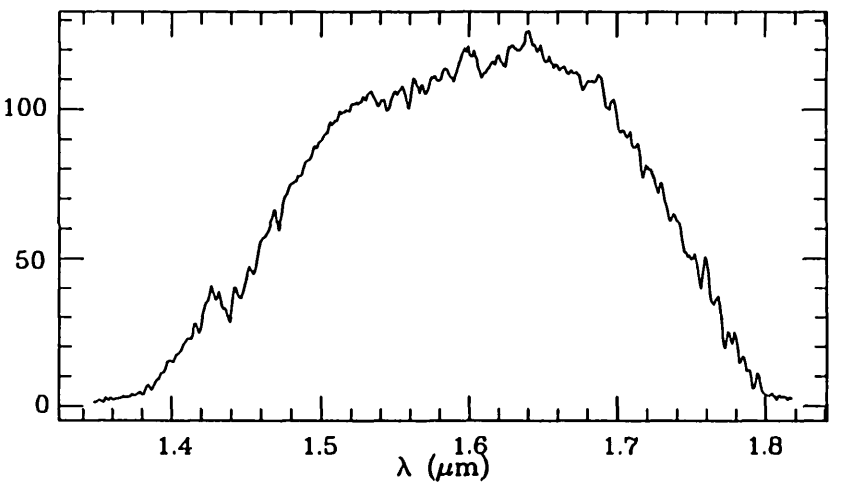


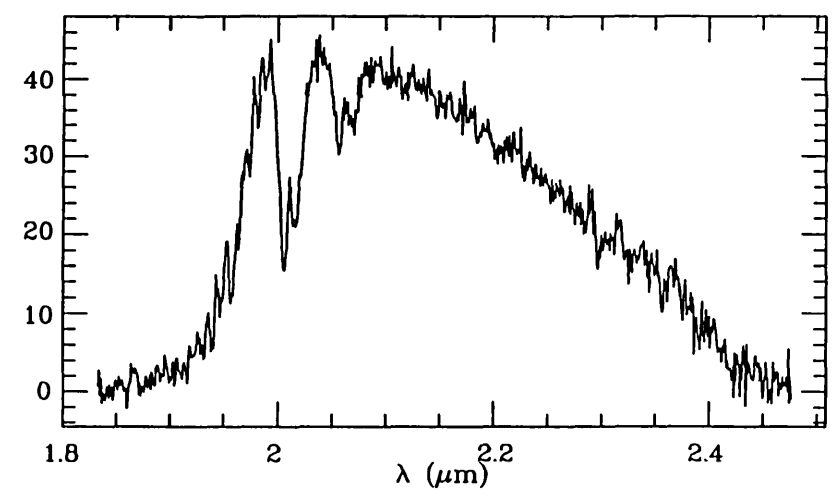


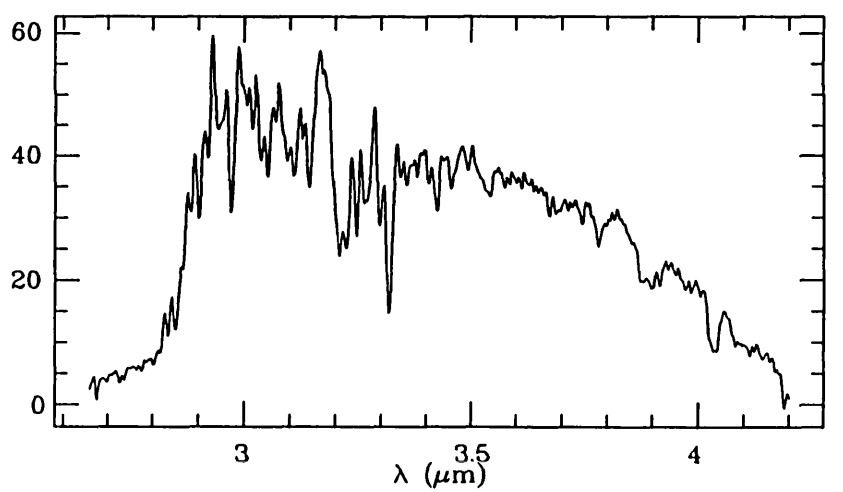
BS 5854

R.A.15^h44^m dec.+6°26′
K2 III m_v=2.7





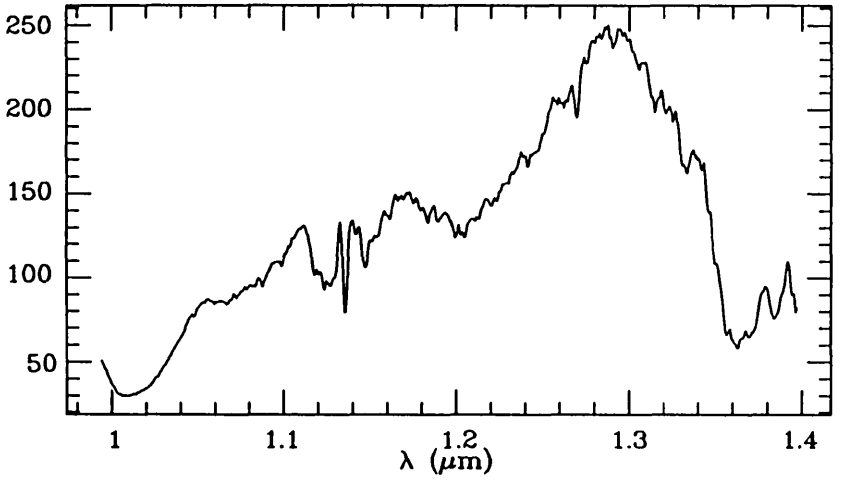


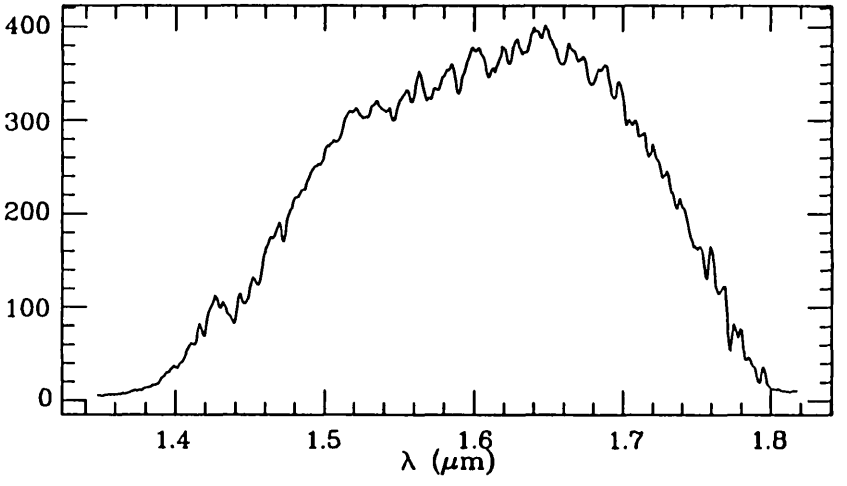


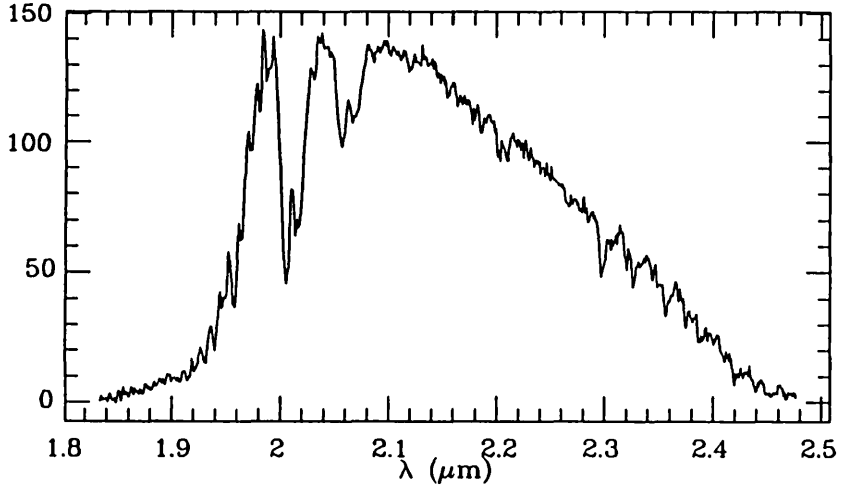
BS 6056

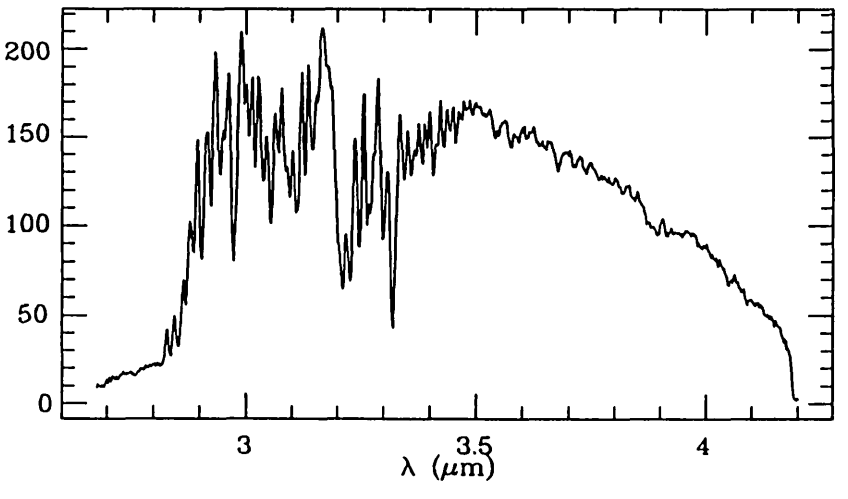
R.A.16^h14^m dec.$-3°42'$

M0 III $m_v=2.7$





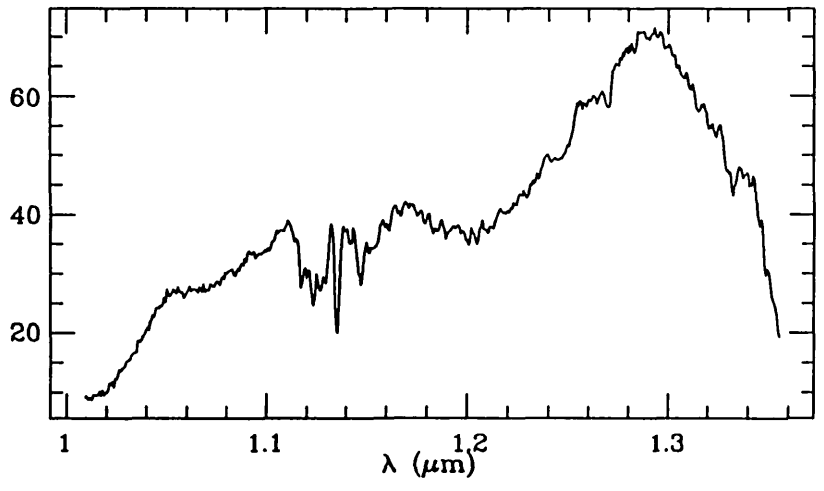


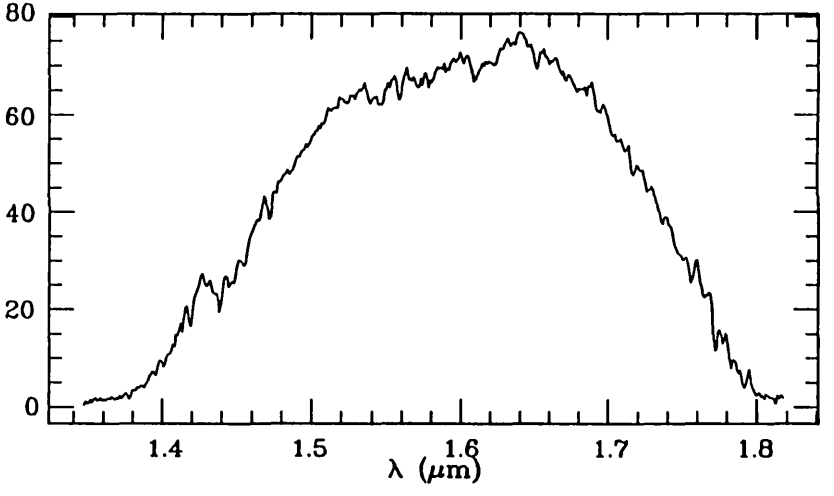


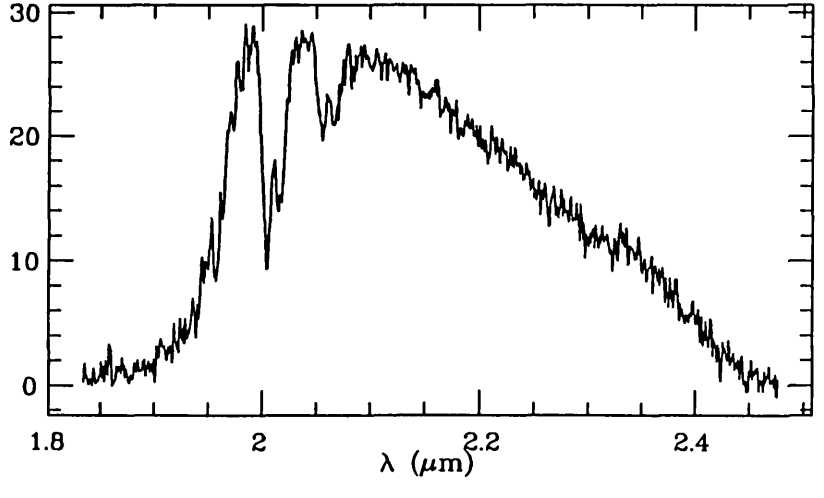
BS 6148

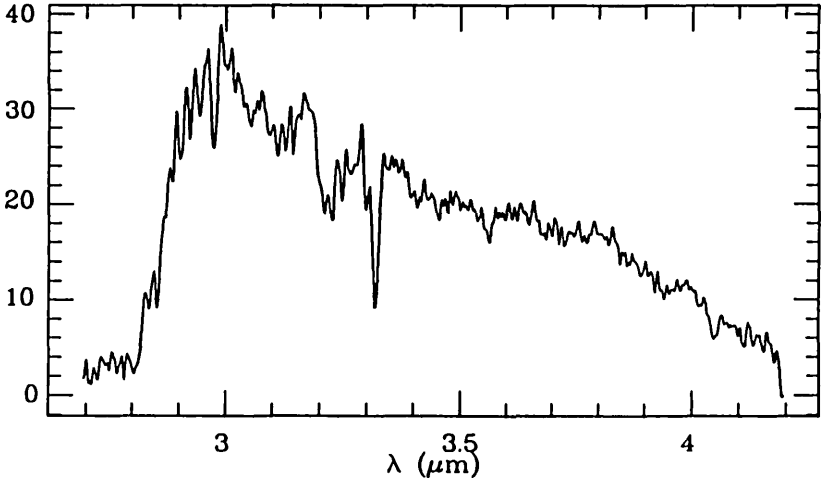
R.A.16^h30^m dec.21°29′

G7 III m_v=2.8





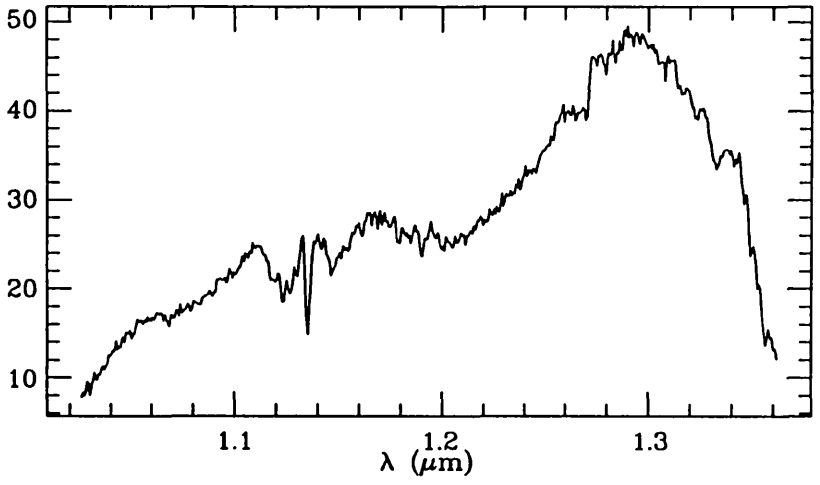


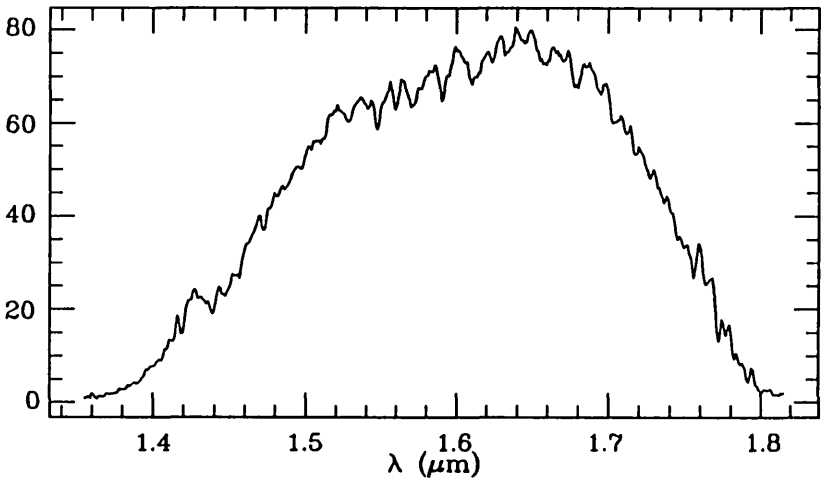


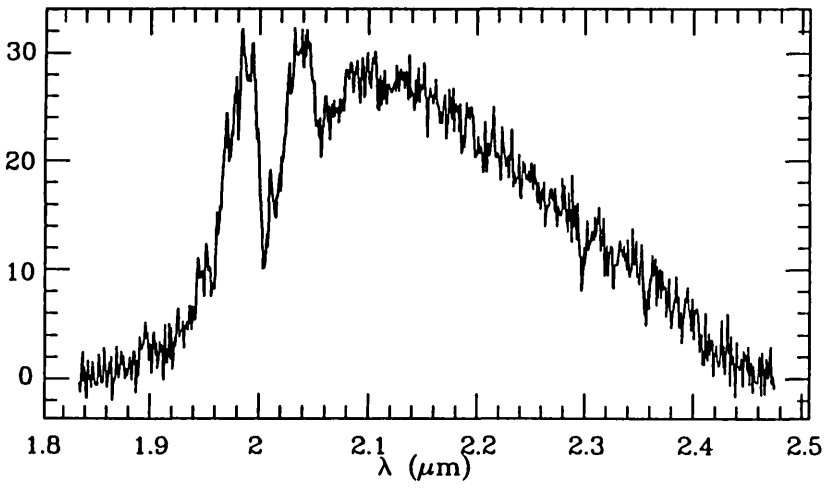
BS 6337

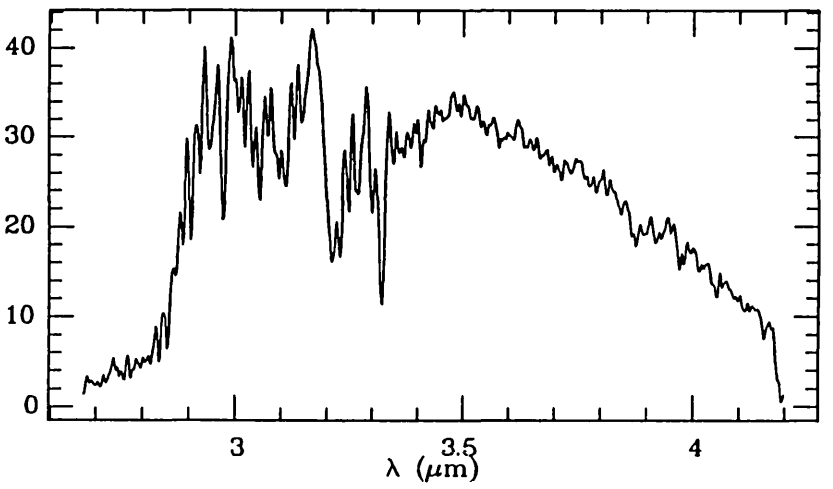
R.A.17^h03^m dec.14°05′

M3 III m_v=5.0





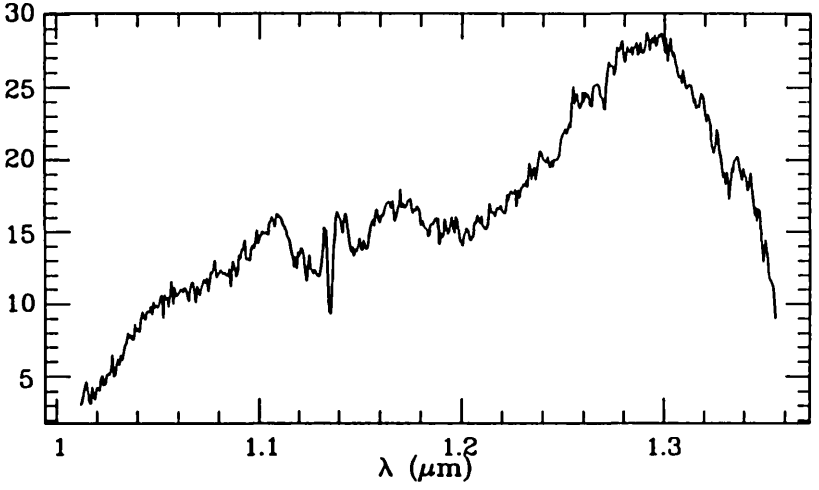


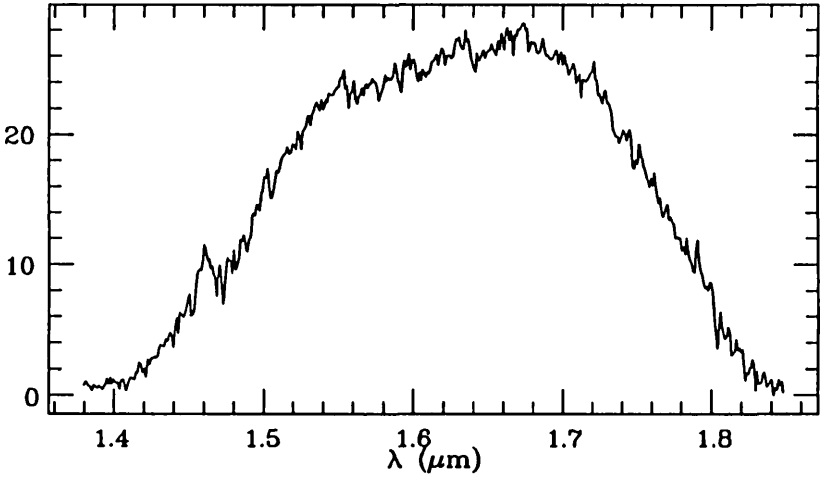


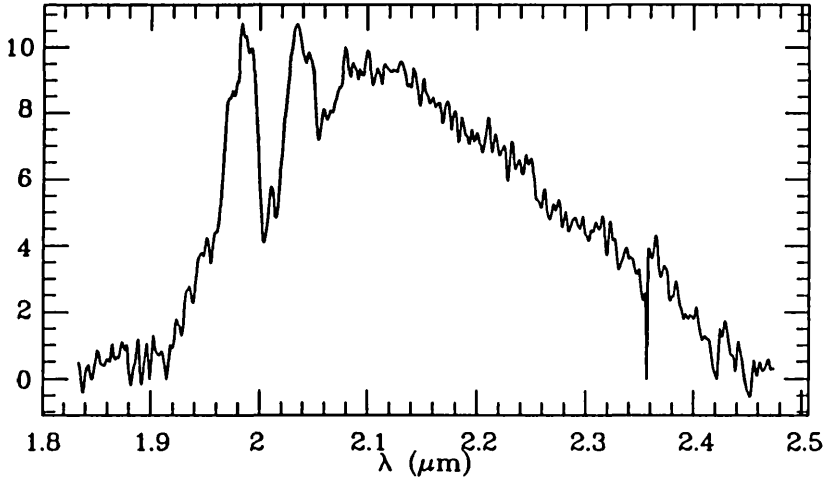
BS 6623

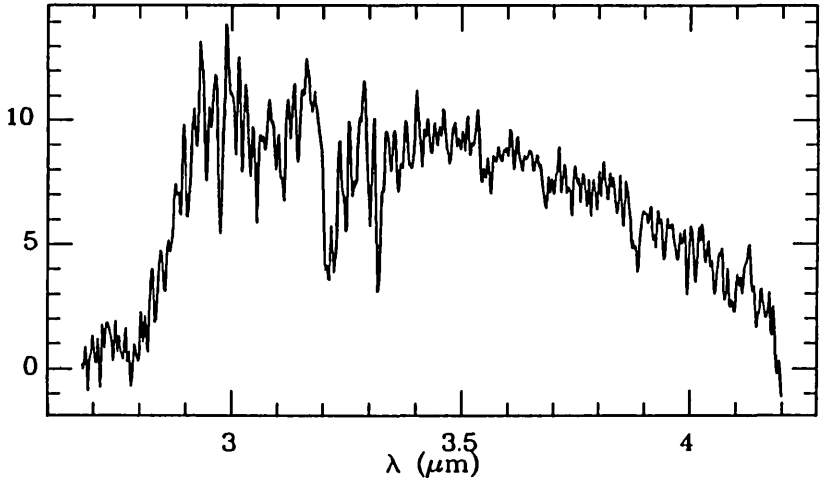
R.A.17^h46^m dec.27°43′

G5 IV m_v=3.4





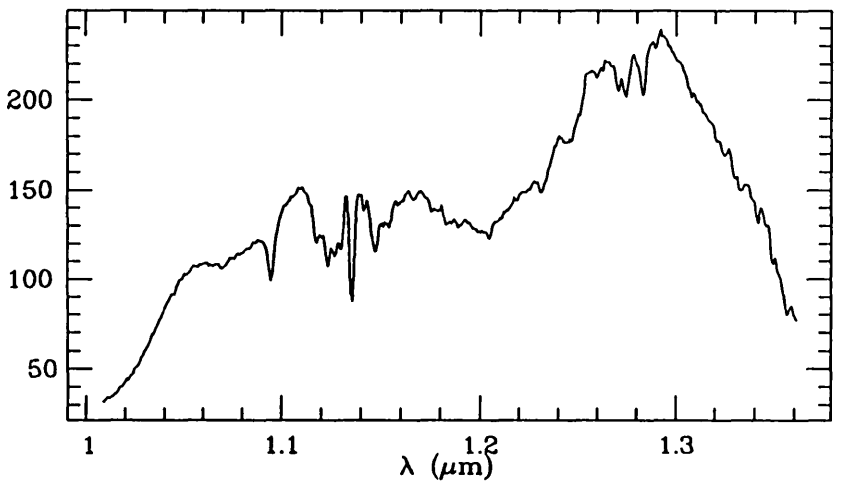


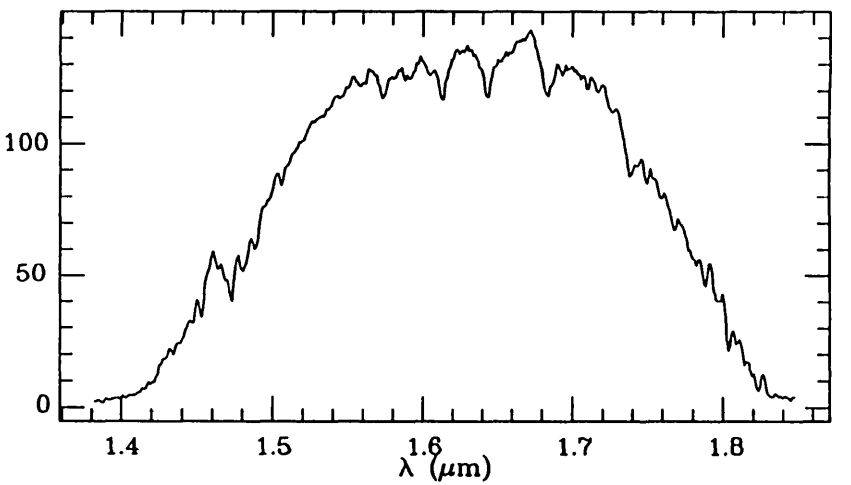


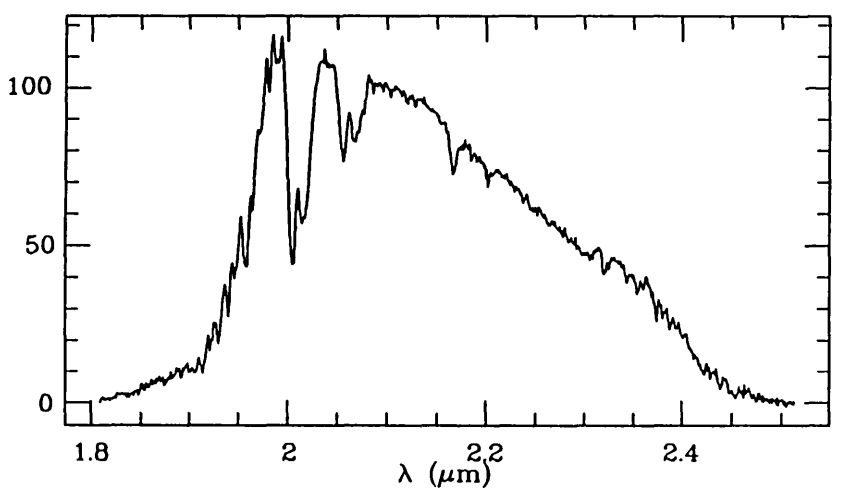
BS 7001

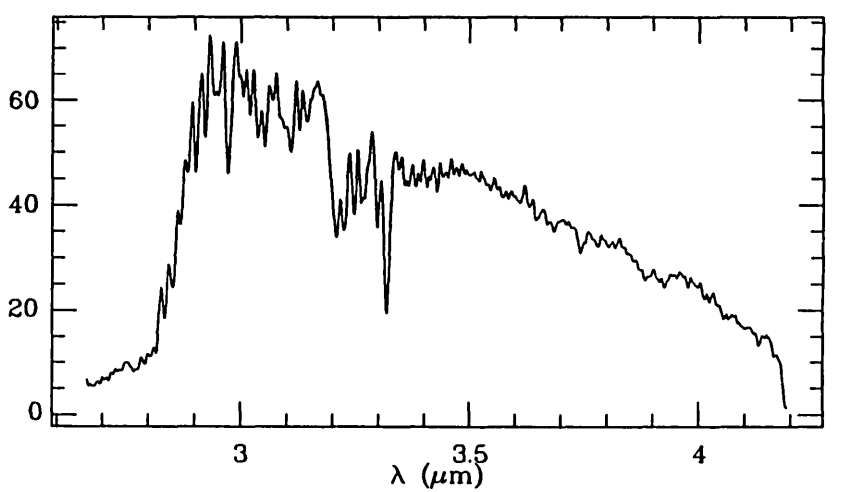
R.A.18^h36^m dec.+38°47′

A0 V m_v=0.0





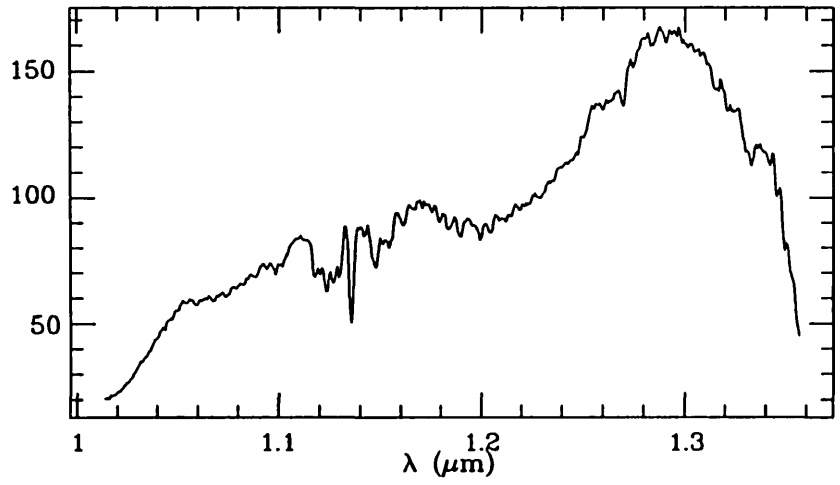


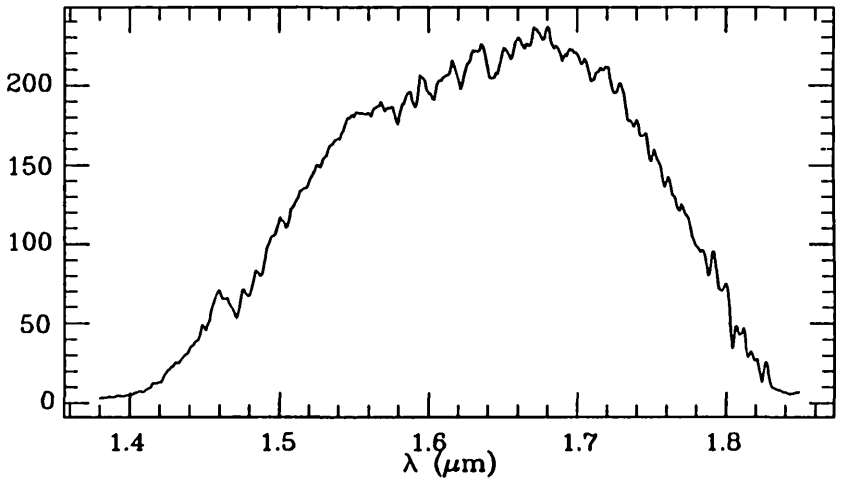


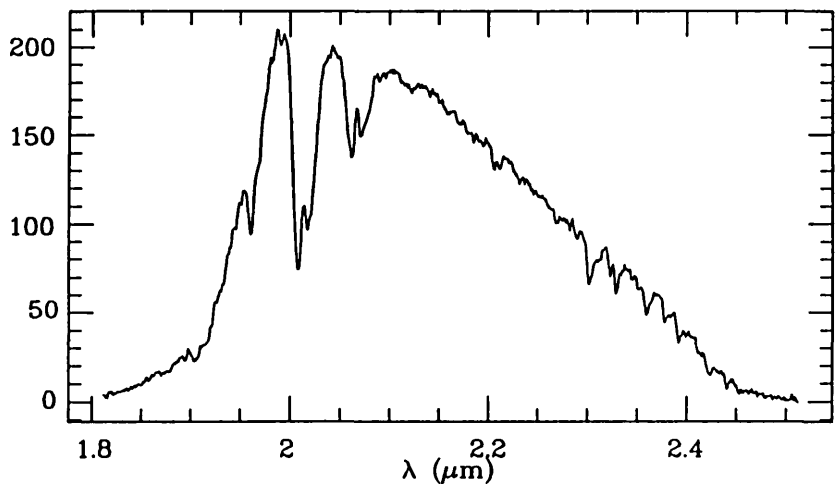
BS 7525

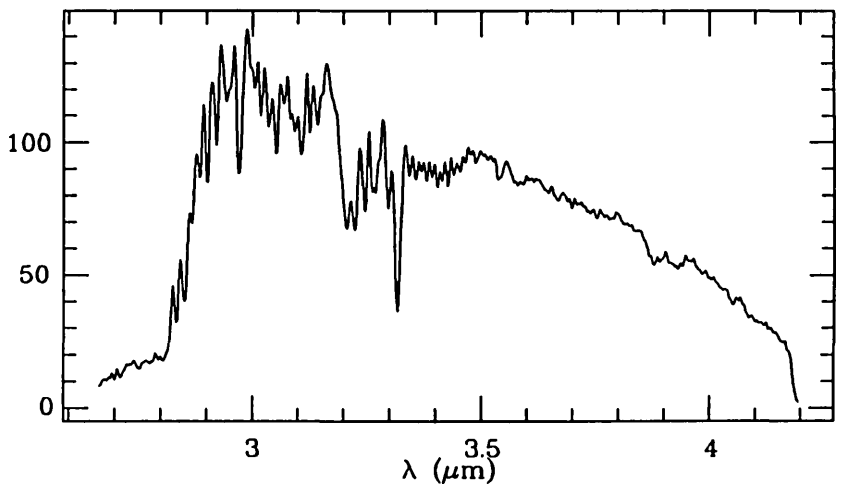
R.A.19^h46^m dec.10°37′

K3 III m_v=2.7





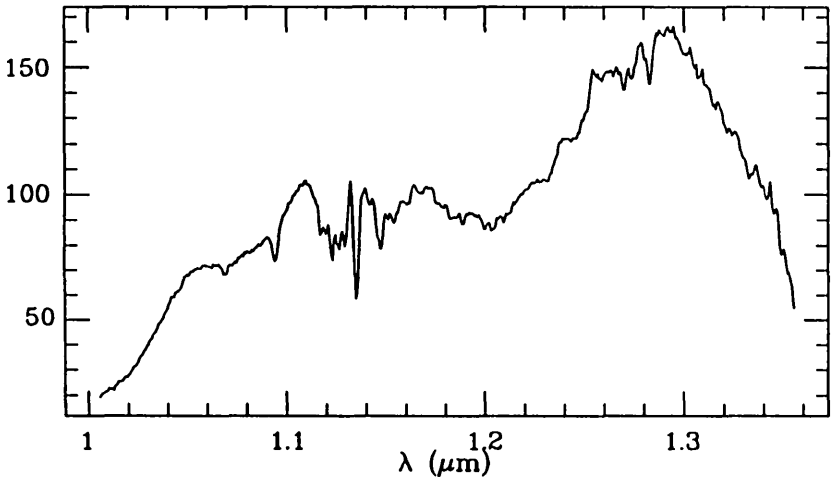


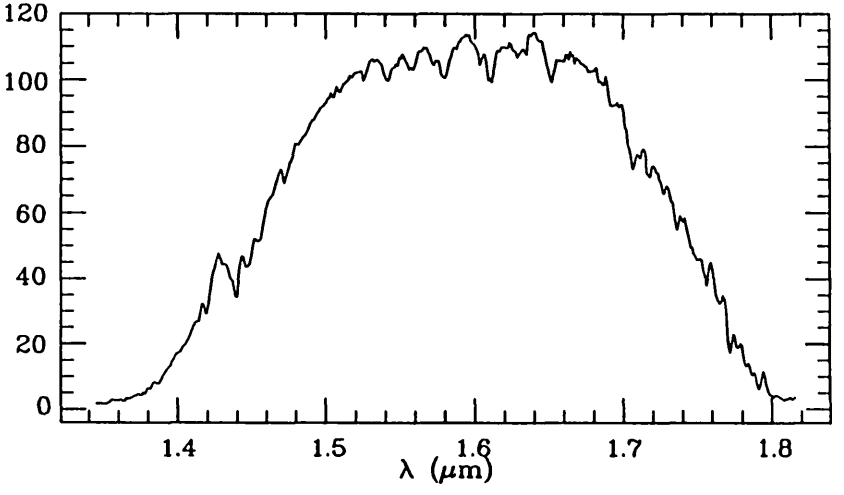


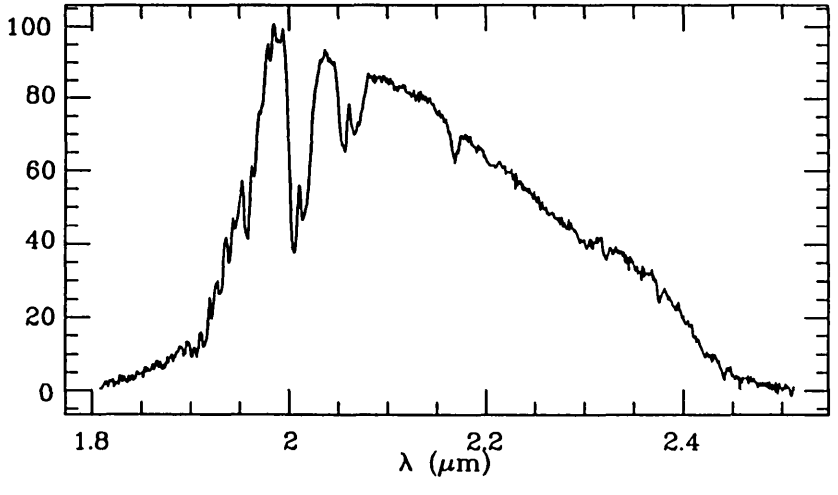
BS 7557

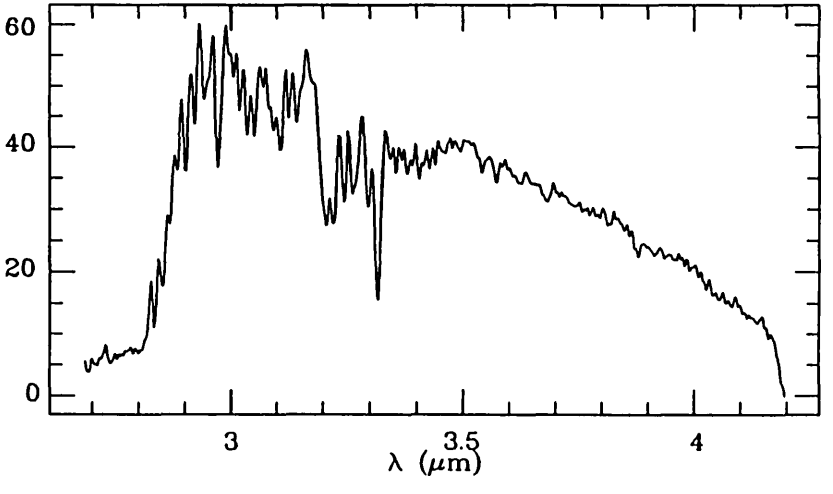
R.A.19^h50^m dec.8°52′

A7 V m_v=0.8





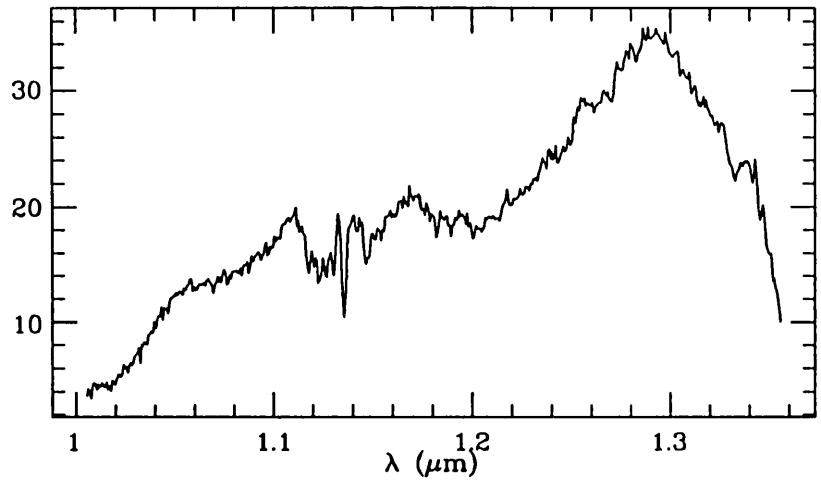


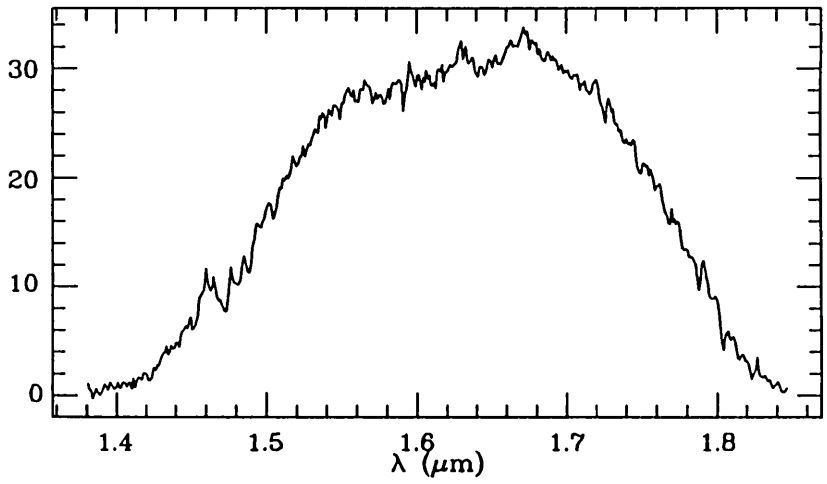


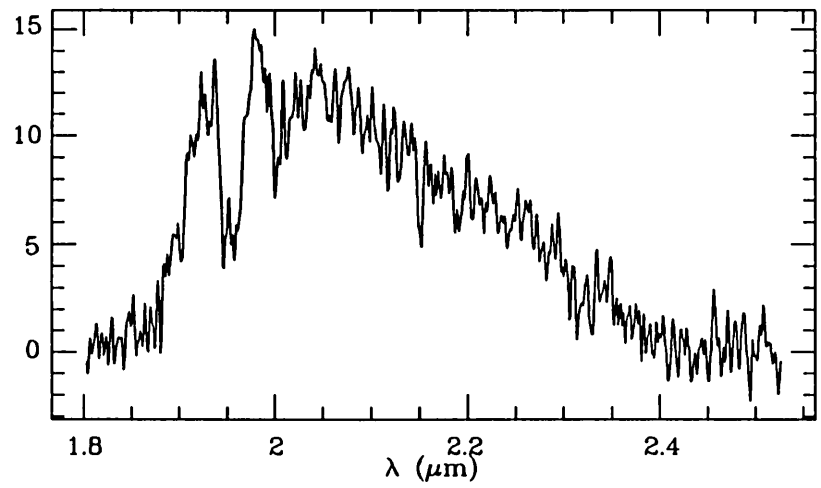
BS 7754

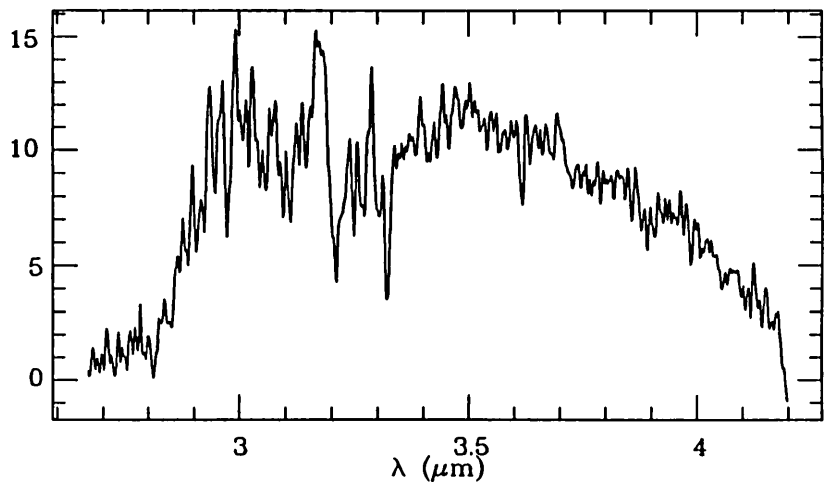
R.A.20^h18^m dec.$-12°33'$

G8 III m_v=3.6





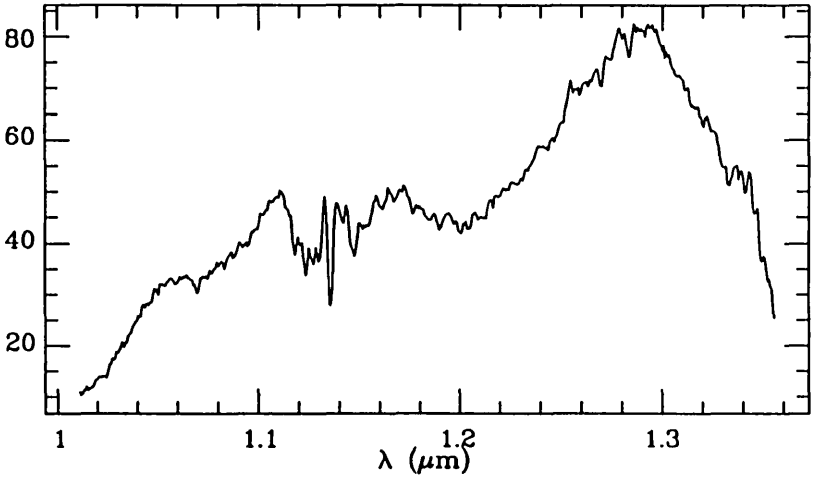


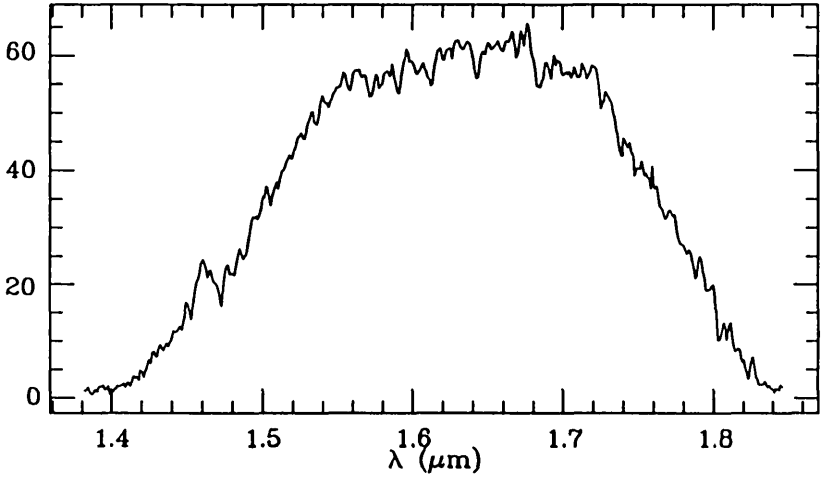


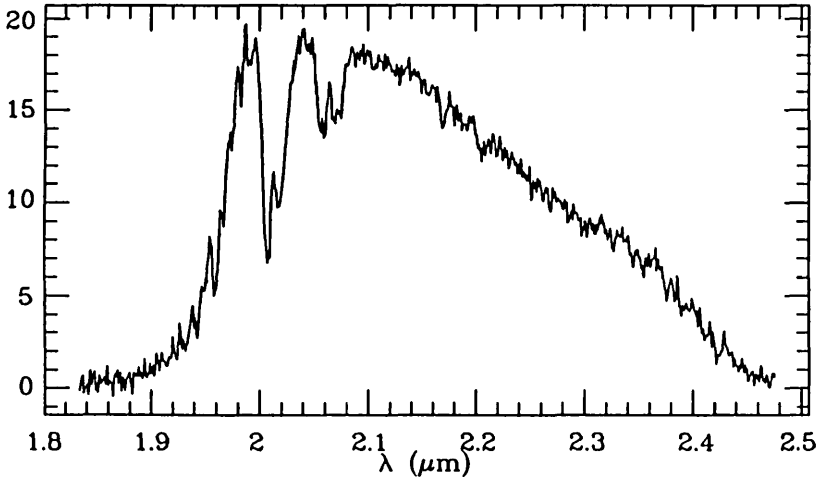
BS 7796

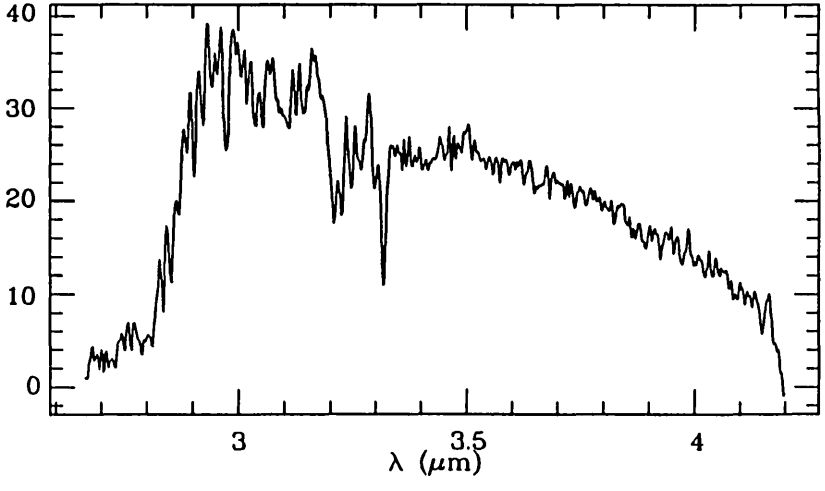
R.A.20^h22^m dec.+40°15′

F8 I m_v=2.2





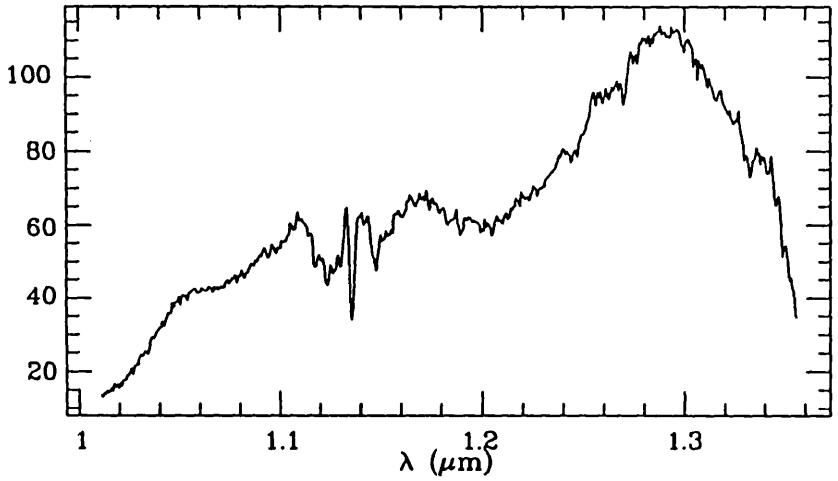


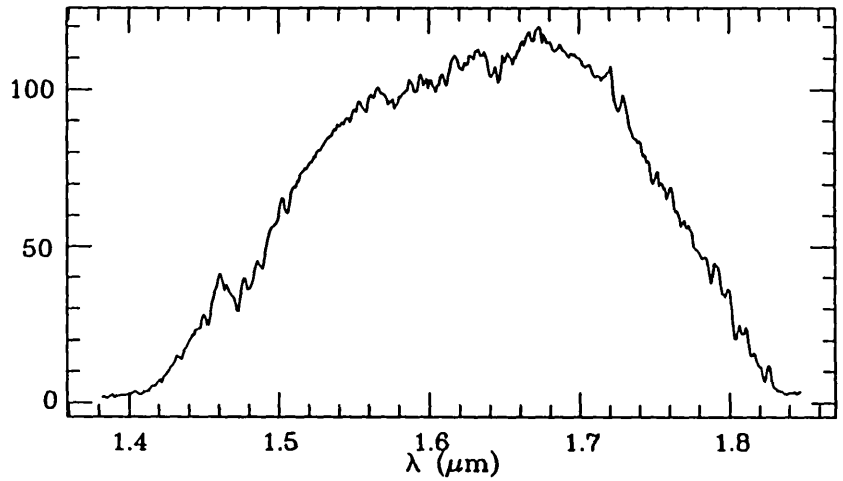


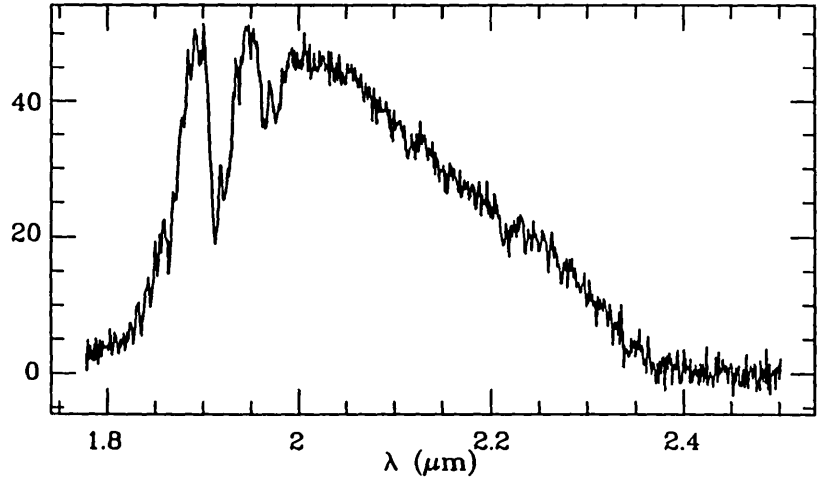
BS 7949

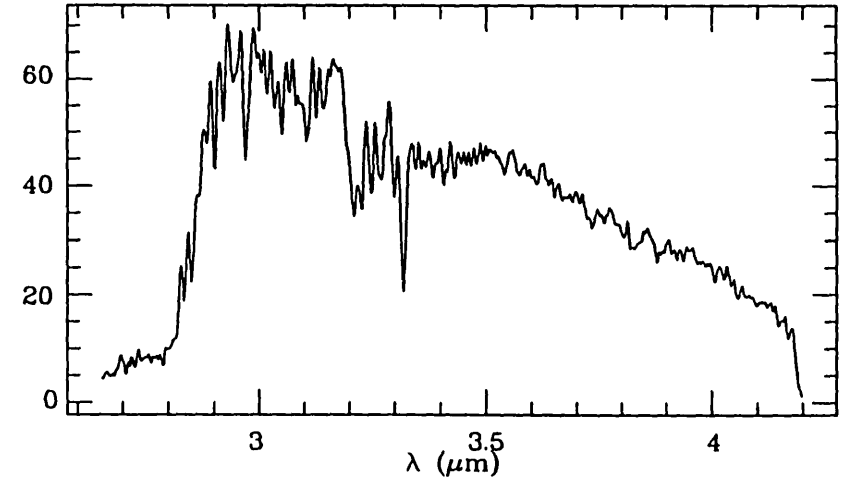
R.A.20^h46^m dec.+33°58′

K0 III m_v=2.5





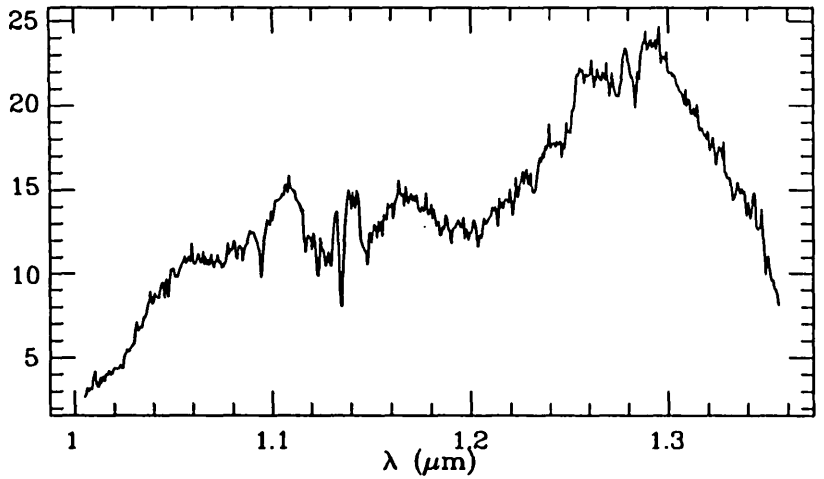


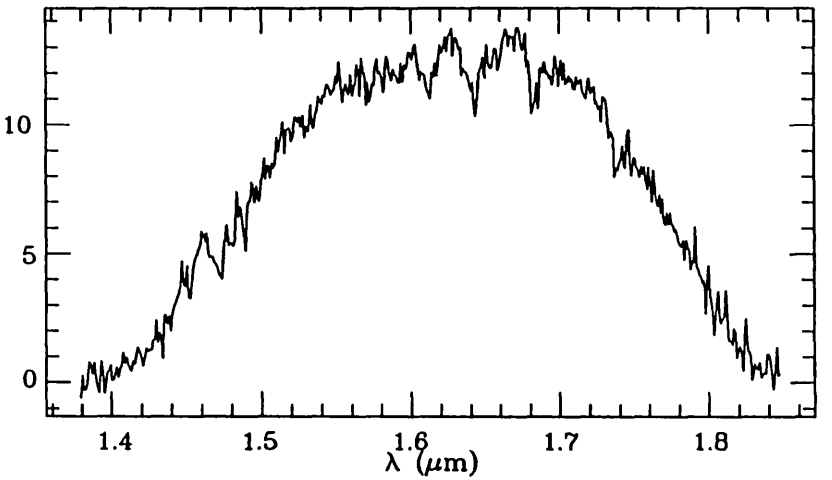


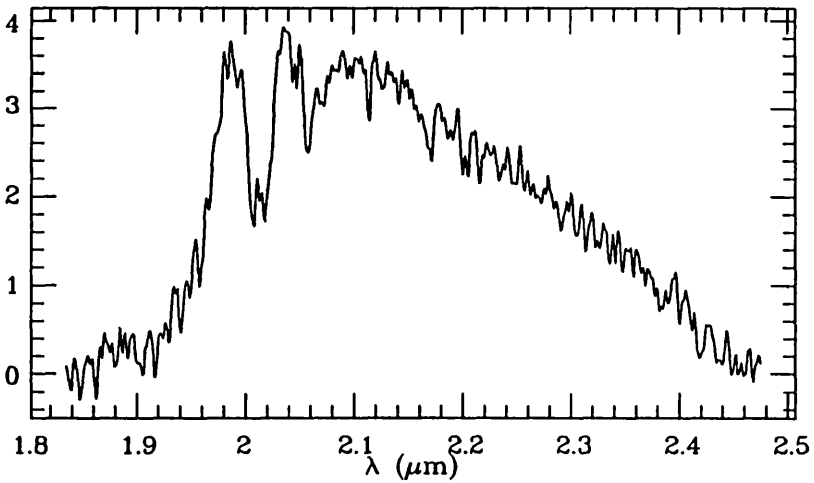
BS 8781

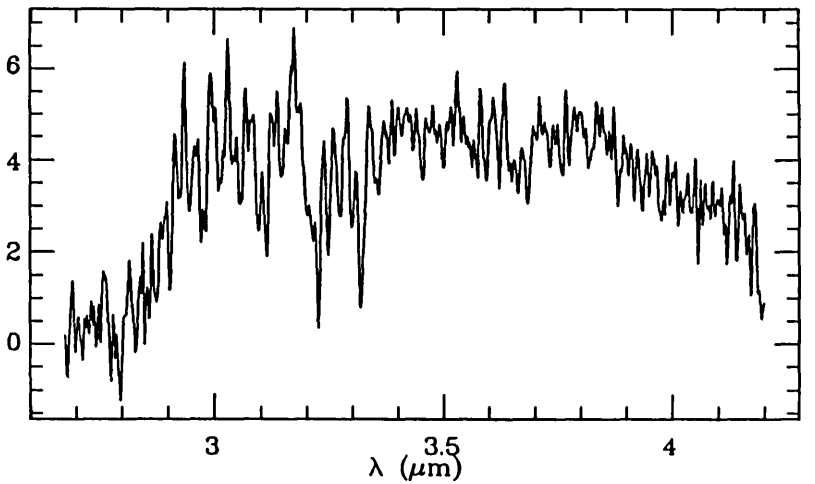
R.A.22^h59^m dec.+14°40′

B9 V m_v=2.5









Acknowledgements

The work described in this thesis was carried out under the supervision of Dr. Mike Selby, in the Astrophysics Group at Imperial College.

Dr. Selby led the two instrumentation projects described in Chapter Two, the Four Channel Photometer and the Cooled Grating Spectrometer. Dr. Matt Mountain was responsible for the development of the spectrometer up to its commissioning, and he conducted early observations with it.

Both instruments were built in the Astrophysics Group Mechanical Workshop, in large part by Mr. Bill Stannard. The data acquisition system for the Four Channel Photometer was designed and built by the staff of the Astrophysics Group Electronics Lab, Mr. Mick Bartholomew, Mr. Victor de Menezes and Mr. Mark Hooker. The Four Channel Photometer's preamplifier circuit was designed by Dr. Xin Dian Xiang of Peking University. The fine encoder for the Cooled Grating Spectrometer was developed by Mr. Roger Sidey of the Imperial College Astrophysics Group.

Many staff of the Instituto de Astrofisica de Canarias gave valuable assistance during observing trips to Izaña. I am particularly indebted to the late Dr. Carlos Sanchez-Magro for his vigorous support, and to Dr. Carlos Lazaro, who helped to obtain the spectrometric data presented in Chapter Two.

Dr. Tom Hartquist provided helpful advice on a part (section 3.5.6) of the HD data analysis. The UKIRT service observation of H_2 emission from W75N(OH), described in Chapter Four, was obtained by Dr. Andy Longmore.

Financial support was provided by the Science and Engineering Research Council.

The author's own contributions to the work described in this thesis were principally in the following areas.

The on-line photometry software for the Four Channel Photometer, the control and data acquisition software for the Cooled Grating Spectrometer, and a part of the spectrometer's control electronics were produced by the author, who also carried out the optical and electrical testing of the instruments described in Chapter Two, and took part in the commissioning of the instruments on the 1.5 m telescope on Tenerife.

The author was responsible for obtaining the observations of bright stars presented in Chapter Two, and he participated in the observations of HD described in Chapter Three. The reduction of the bright star and HD data, as described in Chapters Two and Three, was his work alone, and he contributed to the reduction of the H_2 data in Chapter Four.