Durham University

Durham E-Theses

Measurements of stellar spectra using a silicon photodiode array

A. Humrich

How to cite:

Humrich, A. (1980) Measurements of stellar spectra using a silicon photodiode array. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a https://etheses.durham.ac.uk/id/eprint/7897/ is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
https://etheses.durham.ac.uk

MEASUREMENTS OF STELLAR SPECTRA USING A SILICON PHOTODIODE ARRAY

A. Humrich, B.Sc., M.Sc.

Abstract

This work describes the preparation and use of the Plessey linear photodiode array in the observation of some stellar spectral features. An outline of the principles of the construction and operation of the diode array, along with some preliminary laboratory test results, form the initial part of this thesis. A suitable xy-movement control was designed and constructed, as well as the cryostat which was used to house the array at the coudé focus of the telescope. Constraints at the telescope imposed several limitations, all of which had to be taken into consideration. The general experimental operation, signal processing, and control logic are all described, along with the data acquisition and storage techniques. Data reduction and analysis methods were developed to deal with difficulties peculiar to this type of photodiode array system. A theoretical investigation into the effects of the array geometry on the collection of spectra is described.

On the astronomical side, observations of the neutral oxygen triplet at λ7774 form the major set of results obtained. The equivalent width of λ7774 can give information about the temperature absolute magnitude, and luminosity of many types of stars. The rôle of non-LTE is discussed, as well as its effect on the measurements of line widths.

Finally, the phenomenon of stellar mass loss is outlined. Unfortunately, few results were obtained in this investigation. However, a description of the reasons behind the loss of matter from stars is given, along with probable mass loss mechanisms, features, and detection. A brief word on mass loss rates is included, but the lack of sufficient results in this region pre-empted any quantitative comparisons.

MEASUREMENTS OF STELLAR SPECTRA USING A SILICON PHOTODIODE ARRAY

by

A. Humrich, B.Sc., M.Sc.

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

Being an account of work carried out at the University of Durham during the period October 1976 - February 1980.

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

DURHAM UNIVERSITY
12 MAY 1983
SCIENCE LIBRARY

ACKNOWLEDGEMENTS

I wish to express my gratitude to my supervisor, Dr. J.M.Breare, for his guidance and assistance during my period of study at Durham University. Also, thanks must be given to Prof. A.W.Wolfendale for his interest in the project.

The other members of the Nuclear Instrumentation Group receive my appreciation as well. In particular, Drs. A.R.Hedge, A.W.Campbell, and G.R.Hopkinson, being the three other members directly involved with this area of research, formed the rest of the team that launched the group into astronomical research. Involved with every project the group deals with, it's research technician, Mr. J.Webster, gave much support, both with his technical expertise and his good humour. Two workshop technicians, Messrs. R.McDermott and H.Davison, exhibited their great skills in creating working machinery out of the far from ideal drawings they were supplied with. I must also give thanks to Lindy Luke, Maureen Smith, and Vicky Orson-Wright for their effort put into the typing of this thesis.

Most of all, I appreciate the support and drive given by my wife, Lynne, during the times when I most needed it, and the love of our daughter, Katie, who will always mean so much to both of us.

I also acknowledge the financial support of the Science Research Council (now the Science and Engineering Research Council) during my three years of research at Durham University.

TABLE OF CONTENTS

CHAPTER ONE

INTRODUCTION

For many years now, computers have played an essential role in analysing the vast amounts of data that can be obtained from even the most simple of experiments. However, analog readings from meters (eg. pointer deflections measured in some quanta of degrees of arc) have had to be converted manually to a form suitable for computer input. One of the most significant advances in data acquisition was the development of the digital voltmeter, or DVM. In effect, this piece of equipment is an analog to digital converter (ADC). An infinitely variable signal (voltage) is the input, and this is converted to a form which can be displayed such that it can be transcribed to some permanent record (eg. log book or computer keyboard terminal) with no further conversion necessary by the experimenter. High accuracy devices can be obtained which can improve on the accuracy of the human eye watching a meter. Moreover, this DVM output signal can be tapped before it reaches the numerical display of the device where it still exists in its binary coded form. This is where one finds the use of the computer invaluable. As the computer only deals with binary signals, we now have available the possibility of continual monitoring of the output voltage of an experiment with no intermediate data conversion steps required by the experimenter. The output can be either stored in the computer or in some other storage medium, or it can be displayed to the experimenter. Therefore, any experiment where the output can be expressed in binary form is greatly suited to computer usage where it is possible to carry out immediate analysis of the data.

1.1 Early Observations

Throughout history, astronomers have been recording their observations

DURHAM UNIVERSITY
12 MAY 1983
SCIENCE LIBRARY

of the skies in many different ways. More than about 400 years ago, the only astronomical instruments available were the eyes of the observer. The first astronomers to use telescopes could still only record their observations with the materials available to them at the time - they had to make freehand drawings of what they saw through their telescopes. Each "recording session" could take a fairly long time to complete. By the end of the 19th century, the development of sufficiently sensitive photographic materials for astronomical purposes had taken place. With this new tool available to the astronomers, "exposure" times were greatly reduced. This meant that more observations could be made more accurately over a given period of time.

To investigate the actual constituents of the stars, astronomers and spectroscopists joined forces. Photographic plates seemed the ideal medium on which to record the information obtained from the spectra of the stars. A calibration spectrum (eg. from an elemental arc) can be recorded on the same plate as the stellar spectrum. The dispersion calculated from the calibration spectrum is used to determine which wavelengths of light are present or absent in the spectrum of the star. The missing wavelengths (or lines) "observed" indicate the presence of a particular element, either in a neutral or ionized state, in the atmospshere of the star. From these lines, we can calculate not only the presence of an element, but also its relative abundance in the stellar atmosphere.

1.2 Modern Devices

Once computers came onto the scene, it seemed that the photographic plate might not be as good a recording medium as originally believed. The problem here was how to transform an almost infinitely varying amount of darkness into a form suitable for input to a computer. One must have a machine that can transform shades of grey, varying from transparency to

opacity, into a form that accurately indicates the differences in intensities of light over a particular spectral range. The logarithmic response of the human eye is not an accurate measuring machine. Specially designed pieces of equipment, called microdensitometers, have been built to give a two-dimensional output of a stellar spectrum which has been recorded on a photographic plate. The output is in the form of a graph with the abscissa calibrated in units of wavelength and the absorption of light by the plate plotted along the ordinate. This is a hard copy output from a form of pen recorder, where the deflection of the pen is controlled by the voltage applied to it. As we cannot just feed a graph into a computer, an ADC can be used to receive the voltage signals of the pen at particular values of wavelength. This gives us an output in digital form for computer storage or analysis. The drawback of this system is that the necessary equipment is very large, slow, and expensive. Although it is believed to be a highly accurate method of plate analysis, the analysis can only proceed at a slow rate, and the equipment is far from being portable.

One of the most recent advances in the field of astronomy has been the introduction of the solid state diode. Effectively, it is a device which gives a voltage signal output directly proportional to the amount of light falling upon it. It is operated at low temperatures so that thermal leakages are reduced. An array of several hundreds of such diodes can be made on a single piece of silicon. The geometric stability of the individual devices is extremely high. By exposing this linear array of light detectors to a spectrum, different wavelengths are recorded on different pixcels, or picture elements of the array. The output from the array is input to an ADC, and the resultant signals can be fed into a computer for display and/or storage. Obviously, the detector system is more complicated than has been stated here, but the basic idea remains the same. A fuller explanation of the operation of a diode array is given in Chapter 2.

The work presented here describes the first observations obtained with a novel type of array of silicon photodiodes. The qualities which distinguish this type of array from others are the formation of both amplifiers and a shift register on the same array chip. A great deal of laboratory work was done with this diode array prior to its astronomical operation. A full description of the development and testing of the array is given in the Ph.D. thesis by A.R. Hedge (University of Durham, 1981). Although only the initial astronomical results are discussed here, work is still being carried out with the intention of investigating further into the infrared region of the spectrum.

One added feature of the diode array used in the experiments described is the non-destructive readout (NDRO) facility. To return to the photographic plate as a recording medium, here we find that the exposure time has to be discovered initially by a trial and error method. After a long exposure (~ several hours), the discovery of an under-exposed or over-exposed plate means that great amounts of observing time have been wasted. For very faint objects where there is only time for maybe one plate to be taken in one night, considerable amounts of time can be saved if the observer knows when his exposure is complete. On a diode array system with the NDRO facility, at any time during the exposure the observer can enable the computer to display the current values of the voltages on the diodes. At this time, the observer can decide whether to continue the exposure, or whether to terminate it and store the information temporarily in the computer for output onto paper tape or other storage media. This makes the diode array an extremely powerful, efficient tool in the observations of stellar spectra.

As mentioned above, this is a greatly simplified explanation of the detection system used for the observations described in this work. Chapter 3 describes the operation of the equipment more fully and explains the main functions of the computer and associated electronics employed during the collection of data.

1.3 Applicability of Diode Arrays

There are many fields of astronomy where the use of diode arrays can give equally good if not better results than those given by the photographic plate. By far, the largest area of research in which the greater efficiency of the diode array is realized is in the near infrared and infrared regions of astronomical spectroscopy. The sensitivity of photographic emulsions falls off drastically as investigations proceed further and further into the infrared region of the spectrum. Special emulsions have been developed to increase the infrared sensitivity, but most of the research done has been carried out in the visible and ultra-violet spectral regions.

The sensitivity of the solid state diode array extends well into the infrared, experiencing a fall off only at about 1μm. This fall off is due to the material of the diode being transparent to the lower energy, longer wavelength photons. The incident photon has insufficient energy to interact in the diode, and so passes straight through the diode material. The fall off is a temperature dependent effect. At higher temperatures, the energy gap in the silicon is reduced, so allowing photons of lesser energy to create electron-hole pairs. Infrared sensitivity can therefore be increased by operating the diode at a higher temperature. However, as the temperature of the diode increases, so also does the thermal noise. A compromise must be reached somewhere between a temperature high enough to give good infrared results, and low enough to keep the thermal noise to a minimum.

At the other end of the spectrum, the blue end, one finds that there is also a diode sensitivity reduction. At wavelengths of $\sim 0.4\mu m$, the photon penetration depth in silicon is only about $0.1\mu m$, three orders of magnitude less than for photons of $\sim 1\mu m$ wavelength. The cut-off here is due to surface effects in the diode. These surface effects are in the

form of recombination sites at the illuminated surface of the diode. Presently, there does not seem to be any way around this problem other than increasing the purity of the diode material, so not much is gained here over the photographic plate.

The combined effect of the dispersion of the spectrometer, the telescope, and the physical size of the 256-diode array used meant that at any one time a range of about 110Å was observed on any individual run. As the greatest advantages of the diode array are in the infrared region, the research carried out was in this area of the spectrum. Two specific wavelength regions were observed during the course of this research.

1.4 Spectral Regions Observed

The greatest number of runs were taken around the Balmer absorption feature, H_α, at 6562.808Å, incorporated in a spectral window approximately 110Å wide. A great deal of information can be gained from H_α measurements. This spectral feature can give insight into mass loss rates, luminosity, and possibly stellar temperatures. A closer investigation of the inferences to be made from H_α studies and a description of the results of H_α observation is given in Chapter 7.

The second wavelength region observed was centred on the 7774Å neutral oxygen triplet, also located in a 110Å window. An investigation into the resolution of the diode array system was carried out by attempting to resolve this feature into its three components. This oxygen triplet can be used as a stellar luminosity determinant in distinguishing supergiants from other classes of stars. Also, information can be obtained about the temperatures and absolute magnitudes of stars. Equivalent width values have been published by other astronomers for some stars with the triplet resolved into only two components, and for other stars into three components. Results obtained from the analysis of the data collected with the Plessey diode array are compared with published results. In addition

non-LTE effects are discussed for supergiants, and how these effects alter the strength of the absorption line are put forward.

Over these two wavelength regions, a range of different types of star were observed. Most of the stars were supergiants, with a few giants and sub-giants. The spectral types ranged from B2 (gamma Pegasi) to M0 (beta Andromadae), including stars from each of the luminosity classes. A full description of the observing program, including the stars observed and the analyses carried out, is given in Chapter 4. In summary, nineteen different stars were observed:- 10 supergiants, 2 bright giants, 3 giants, 1 beta Cepheid supergiant, 1 delta Scuti star, and 2 others. The spectral types of the supergiants ranged from A to K, with the emphasis being on the 7774Å region of the spectrum.

CHAPTER TWO

PHOTODIODE ARRAYS

2.1 Array Package

The integrated diode array devices used in this experiment were manufactured by the Plessey Company Limited, under the direction of Dr. D. McMullen of the Royal Greenwich Observatory. Four linear arrays, each consisting of 256 individual diodes, were constructed on a single piece of silicon. The four arrays are orientated into two parallel lines by pairing of the arrays end to end. The devices are packaged in a standard 24 pin DIL "chip". Also in the same package are the shift registers, switching transistors, and amplifiers necessary for the readout of information from the arrays. The size of each diode is 40μm x 200μm, with adjacent diodes at a 50μm centre to centre spacing. This gives a length of 12.8mm for each array of 256 diodes. Figures 2.1 and 2.2 show the physical characteristics of the chip, and an enlarged representation of one of the arrays, respectively. A schematic drawing of a single diode with its associated integrated electronic components is shown in Figure 2.3. For a complete description of the operation of the diode circuitry, see the Ph.D. thesis of A.R. Hedge (University of Durham, 1981). The following section briefly describes how the array chip handles the signals generated in the solid state diode. Refer to Figure 2.4 for the representation of the diode in use.

2.2 Diode Method of Operation.

Figure 2.4 shows a slice taken through the depth of the material of the diode. The bulk material is n-type silicon, with a thinner layer of p-type silicon on the surface. The diode is maintained at a reverse bias,

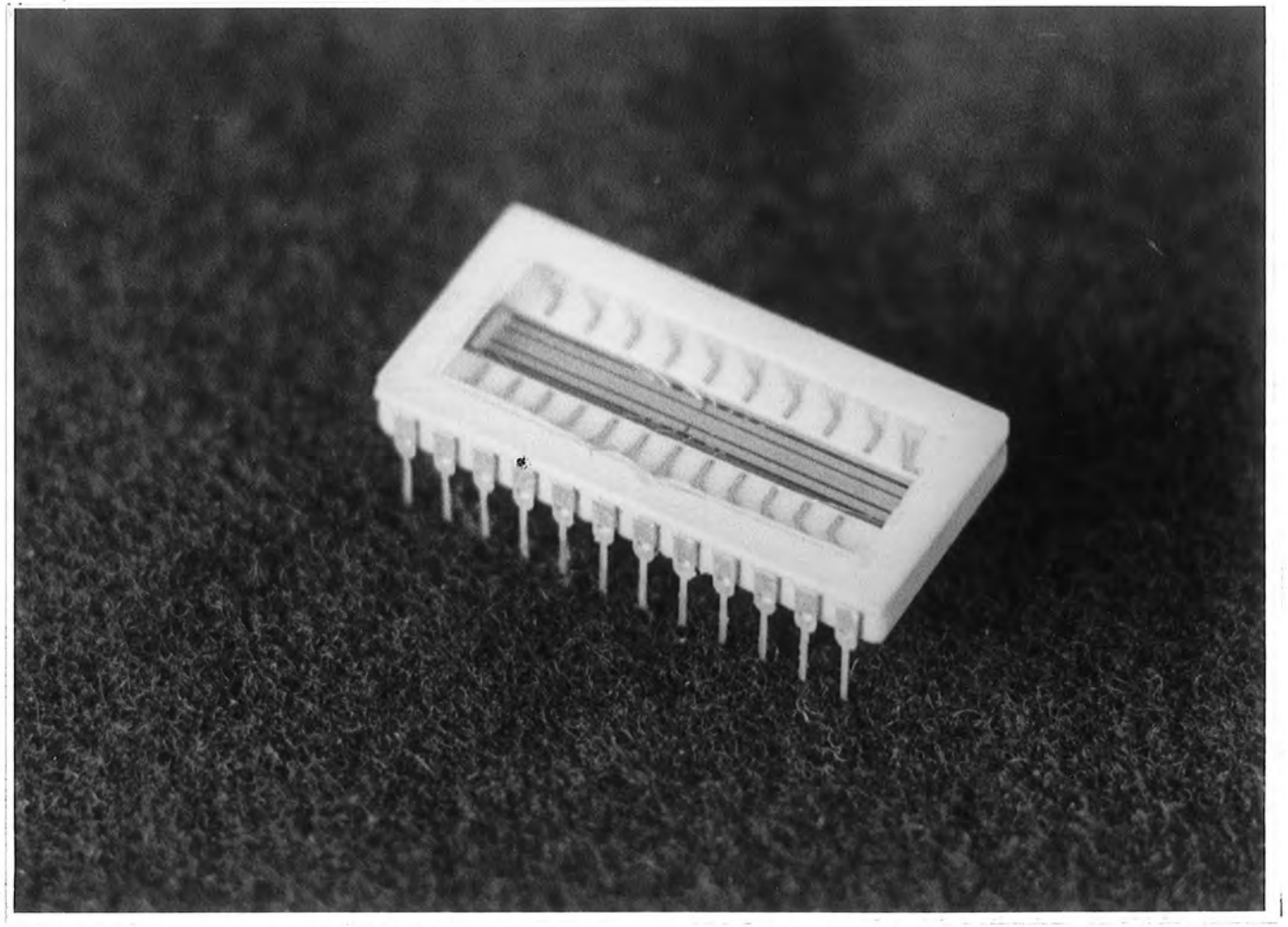


FIG. 2·1a: THE PLESSEY DIODE ARRAY CHIP

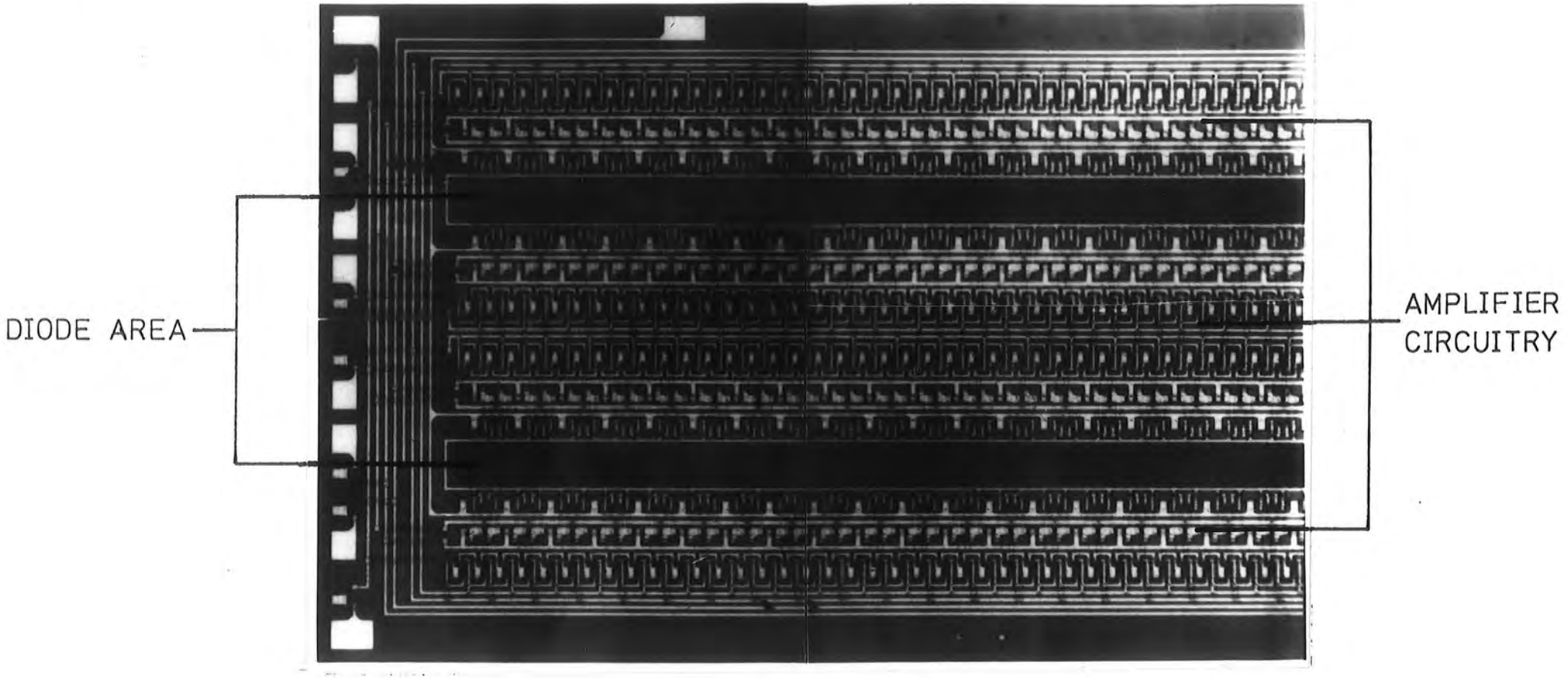


FIG. 2·1b: PHOTOMICROGRAPH OF THE DIODE ARRAY CHIP

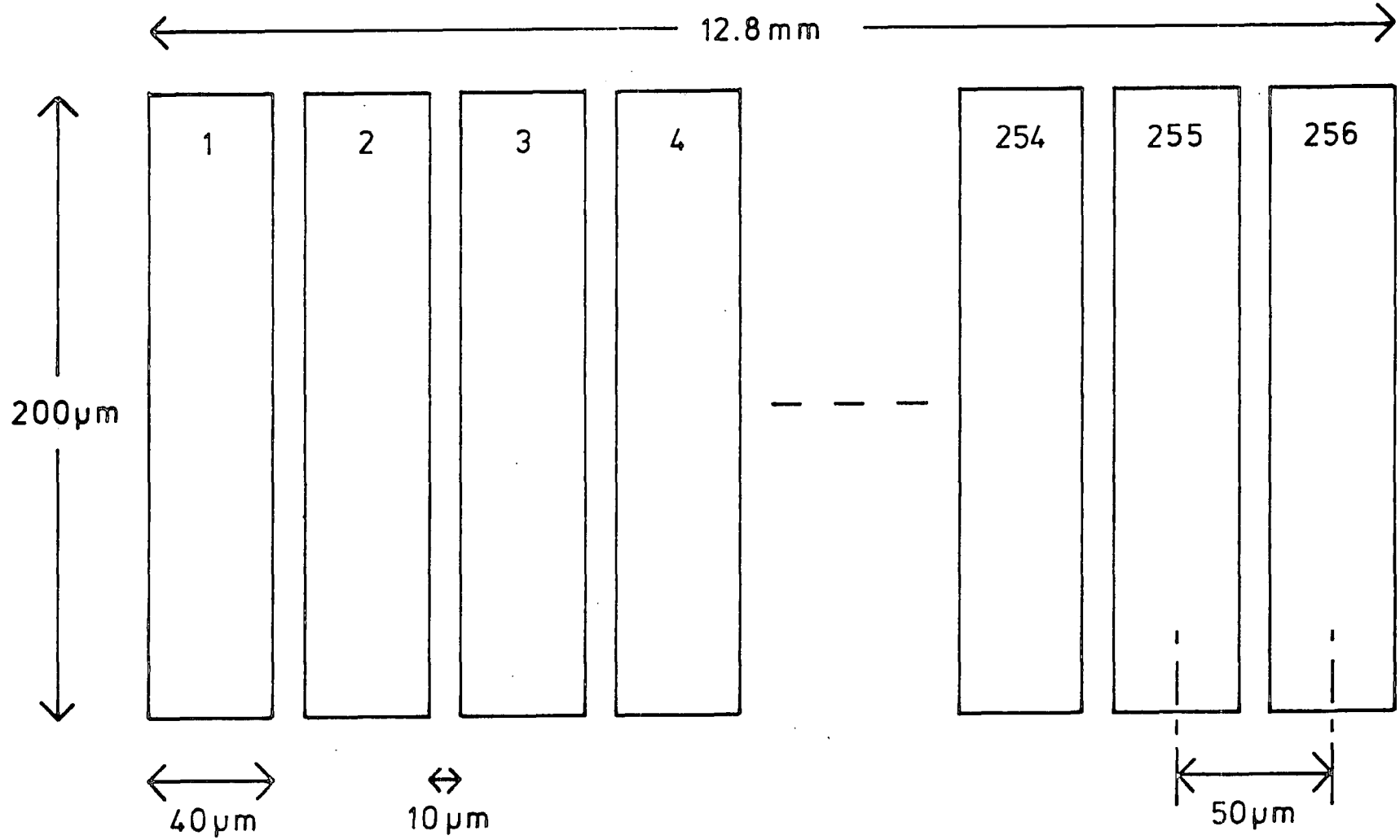


FIG. 2.2: DIMENSIONS AND GEOMETRY OF ONE ARRAY OF 256 DIODES

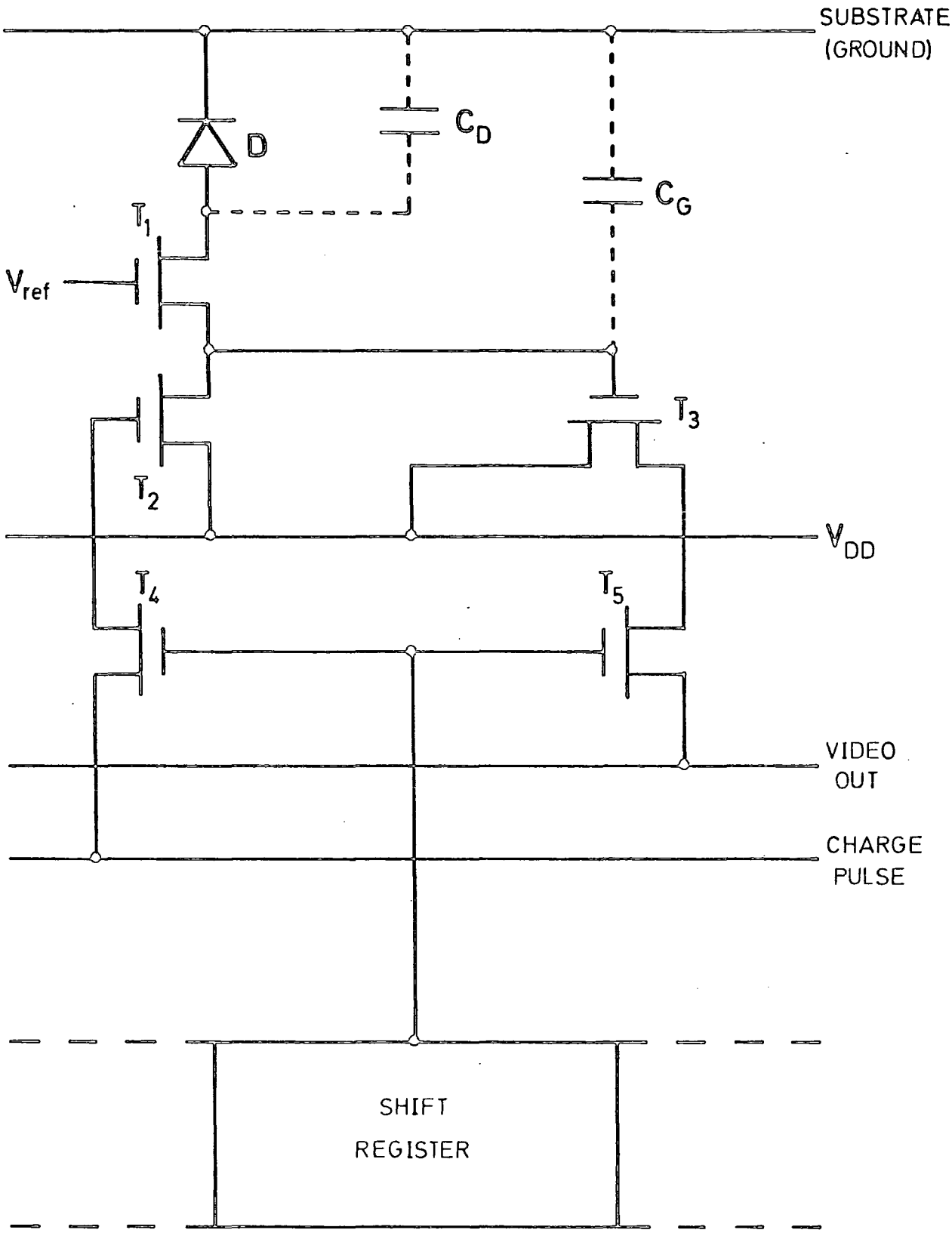


FIG. 2.3: SCHEMATIC REPRESENTATION OF ONE DIODE PLUS ASSOCIATED CIRCUITRY

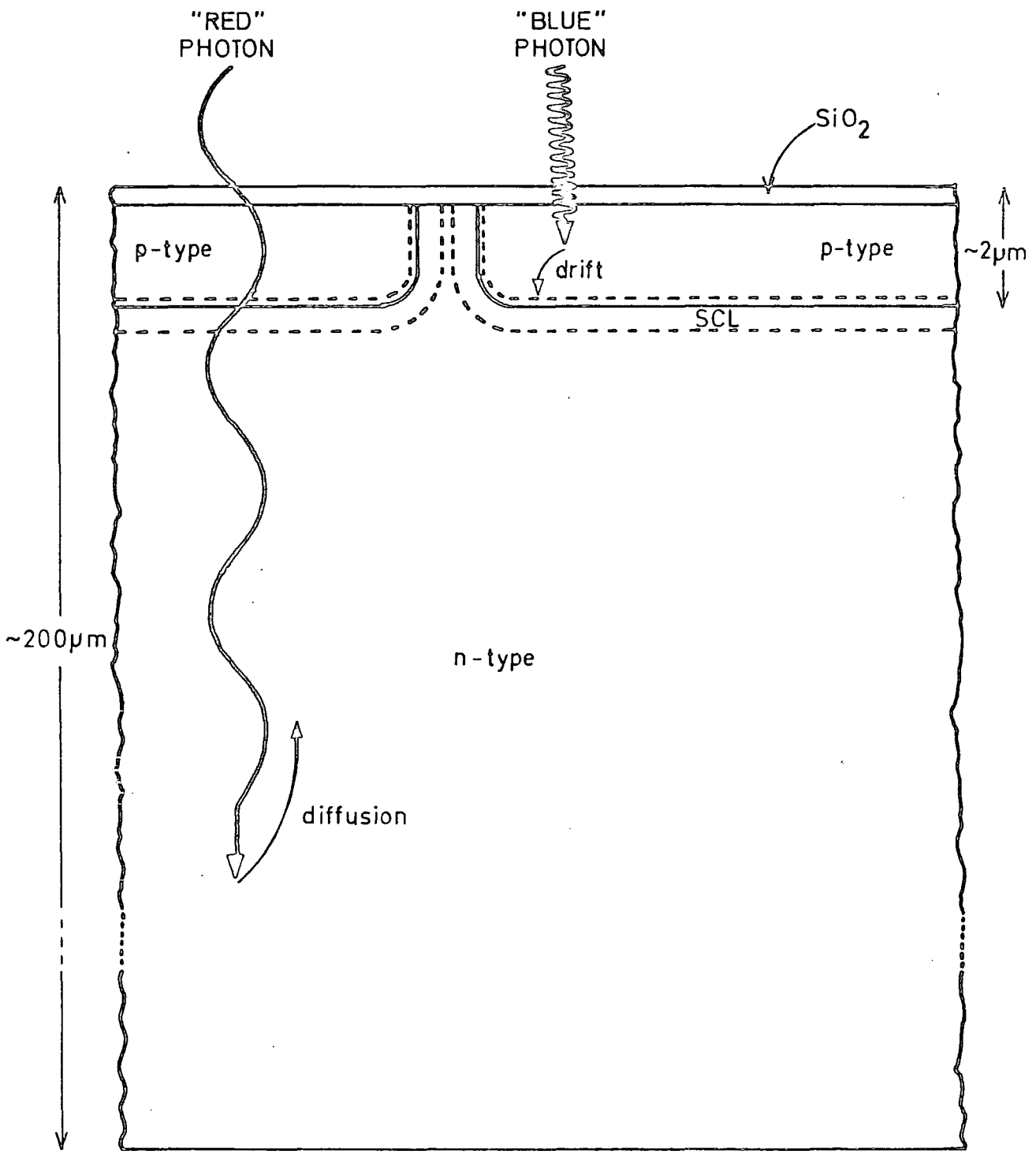


FIG. 2.4: SLICE TAKEN THROUGH DIODE MATERIAL (NOT TO SCALE)

forming a depletion region (or space change layer, SCL) at the n-type/p-type interface. Charge can now be stored in the capacitance of the diode. A typical value of this diode capacitance is of the order of 1pF. If the diode is exposed to light, a photon traversing the diode may give rise to the production of an electron-hole pair in the material. This can only occur if the energy of the photon is greater than the band gap energy, E_g, of silicon. The penetration depth of the photon is a very sensitive function of its energy (see Figure 2.5, from Geary, 1976). A photon from the blue end of the spectrum ($\lambda \sim 4000\text{Å}$), and a photon from the red-infrared end of the spectrum ($\lambda \sim 1\mu m$), will behave quite differently in the diode material.

The "red" photon will travel deep into the substrate and, provided it has sufficient energy, will create an electron-hole pair in the n-type material. The threshold energy of the photon for this electron-hole pair creation is equal to the energy gap of the silicon. The size of the energy gap is a function of the temperature at which the diode is operating, given by

$$\frac{dE_g}{dT} = -4.0 \times 10^{-4} \quad eV.K^{-1} \qquad (2.1)$$

from Geary (1975). This relationship is approximately true for silicon temperatures in the range $150<T<300K$. At temperatures below 150K, the value of E_g tends towards a constant energy, while increasing the temperature causes a quadratic falloff in E_g (see Smith (1961), Ch. 10). In the experiments to be described, only diode temperatures above 150K are encountered, so usage of the above relationship is considered valid.

From equation (2.1), and from the energy equation

$$E = h\nu = {}^{hc}/_{\lambda}$$

the change in the red end cutoff, $d\lambda$, between two temperatures can be

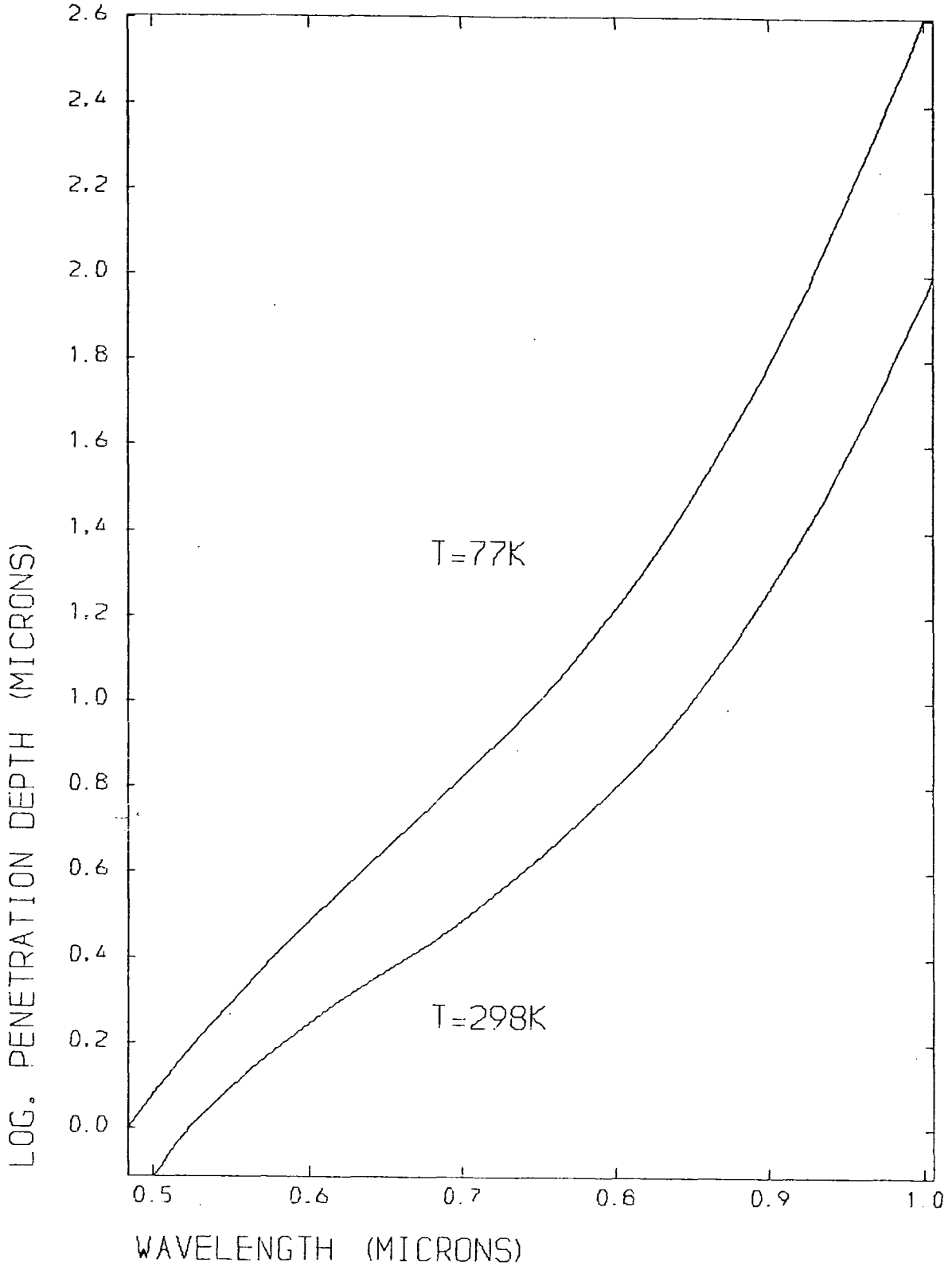


FIG. 2.5: PHOTON PENETRATION DEPTH AS A FUNCTION OF WAVELENGTH

calculated. For a change of diode temperature from room temperature ($20^{o}C$) to $-120^{o}C$, the value of $d\lambda$ can be estimated for an infrared photon as

$$d\lambda = \frac{-4.0 \times 10^{-4} \times 1.6 \times 10^{-19} \times 140 \times \lambda^{2}}{h \times c}$$

Therefore, a decrease in the temperature of the silicon would result in a decrease in the red sensitivity. For pure silicon at room temperature, the photon threshold wavelength is 1.1μm (Vogt, et al., 1978). At $-120^{o}C$, the calculated threshold wavelength is therefore only 1.06μm for infrared photons falling on the silicon. See section 2.5.2 for the measured spectral responsitivity as a function of diode temperature for the Plessey diode array.

Once the electron-hole pair has been created, the minority carriers diffuse towards the depletion region and tend to discharge the voltage across the diode. The amount of charge necessary to recharge the diode to its initial condition can be measured, and therefore the amount of light to which the diode has been exposed can be determined. Photons of energy less than that of E_g will either pass straight through the silicon, or they may possibly be reflected off the back face of the diode. These two effects give a gradual sloping off of the diode responsivity in the infrared region of the spectrum. If reflection does take place from the back of the diode, interference can occur in the bulk material.

Other than for measurements of the effect of temperature on the infrared sensitivity of the diodes, all of the data were collected with the array cooled down to $\sim -120^{o}C$, with liquid nitrogen. The method of cooling and the measurement of the array temperature is described in section 3.1.

If the photon entering the diode is a "blue" photon, the electron-hole pair will be produced in the p-type silicon, near to the diode surface. In this region, the minority carriers will drift towards the depletion

layer in an attempt to discharge the diode. Again, the amount of charge necessary to recharge the diode can be measured. The absorption coefficient of the silicon is a function of the wavelength of the incident photons. There is a blue wavelength cut-off at approximately 0.4μm. This cut-off is sharper than the red cut-off and is due to the fact that at these shorter wavelengths the absorption depth of the silicon is only $\sim 10^{-7}$m. Due to the larger density of recombination sites in the p-type material (mainly surface effects) than in the n-type, the photon produced carriers may recombine before drifting to the p-n junction. Geary (1976) reports that variations in the width of the depletion layer can give rise to non-linearities in the blue end response of the diode. Since the width of the depletion layer is a function of the square root of the junction potential, the nearness of the depletion layer to the surface of the diode will vary. This decrease in width of the depletion layer as the diode becomes more and more discharged will mean that the carriers produced will "see" more recombination sites before reaching the space charge region. This non-linearity can be as much as 20 - 30%.

In the preceding description of the detection of carriers produced by optical means, the same detection mechanism holds true for thermally produced electron-hole pairs. At room temperature, the diode will discharge in about a second due to these "thermal" carriers. This process is known as dark current or thermal leakage noise. In order to benefit from the solid state diode as a detector of photons, the diode array must be cooled down. Reducing the temperature effectively reduces the energy of the electrons, so keeping the dark current at a minimum, since the energy gap will be too large for the electrons to cross. Only in this way can we achieve the long exposure times necessary to observe stellar spectra with a good signal to noise ratio. Measurements of the

dark current noise in the Plessey diode array give a value of $3\mu V.s^{-1}$ at $-120^{o}C$.

Ideally, each incident photon, within the wavelength range to which the diode is sensitive, should produce one electron-hole pair. However, recombination sites do exist inside the bulk material and reflection surfaces are unavoidable. The surface of the array is coated with a layer of silicon dioxide to reduce reflections at the front surface of the chip. The quartz window will introduce some reflections, as will the back face of the diode, the latter in particular giving possible problems with infrared photons. Therefore, the efficiency of the diode array is not 100%, but, in general, has been measured to be above 80% (Geary, 1975).

2.3 The Plessey Diode Array Chip

The on-board amplification stage consists mainly of the two transistors, T_1 and T_3. Considering Figure 2.3, the stray capacitances of both the diode, C_D, and of the gate of T_3, C_G, are shown. Typical values of these capacitances are 0.1pF for C_G and 1pF for C_D. The gate of T_1 is kept at a constant negative voltage, V_{ref} (~-4V). This keeps the circuit closed so that there is a constant voltage across the diode. When events occur in the diode, either from photon interactions or thermal noise, the voltage across the diode is maintained by T_1 conducting charge from the gate capacitance of the source follower transistor, T_3. In this way, the diode is always fully charged. Since the width of the depletion layer in the diode is a function of the reverse bias voltage across the diode, variations in C_D would occur as the diode became more and more discharged. As mentioned earlier, non-linearities in the blue response have been observed. T_1 eliminates this problem. Also, since most of the thermal leakage takes place in the depletion region of these

diodes, this noise is kept constant, as well as small, by keeping the value of V_{ref} only slightly above the threshold voltage of T_1. Therefore, the volume of the depletion region is kept to a minimum to reduce the likelihood of thermally generated electron-hole pairs. The amplification is related to the presence of the stray capacitances C_D and C_G. The voltage gain of this negative feedback circuit is given by the ratio C_D/C_G. Using the values previously mentioned, one finds that there is a voltage gain of ~10 at the gate of T_3.

Transistor T_3 is operated in the source follower mode, with the drain voltage, V_{DD}, kept at a constant negative value (~-10V). The charge conducted by T_1 is deposited onto C_G, charging the gate voltage of T_3. Since V_{DD} is constant, the source voltage of T_3 will follow the gate voltage. By monitoring the source voltage, one can determine the change in voltage across C_G due to thermally or optically produced charges in D. When it is required to observe the source voltage, it is necessary to make use of the multiplexing transistor, T_5. Upon receipt of a clock pulse, T_5 switches on. This will then put the source voltage from T_3 onto the video output line. The video line is the final output line from the device, with one line for each of the four arrays on the chip.

T_2 and T_4 are switching transistors as well as T_5. They are used to recharge the gate capacitance of T_3 up to V_{DD} after the required integration time has elapsed. The remainder of the electronics on the chip is the shift register. This is the addressing section of the device and is used to enable the information from all of the 256 diodes to be output on to one video line.

2.4 Diode Array Output Monitoring

There are several methods of operation for diode arrays. Firstly,

there is the charge integration or recharge sampling method. Basically, a reverse bias is applied to the diode, and then the diode is left open circuit. Thermally and optically induced charges act to discharge the diode. When the integration time is completed, the original reverse biassing voltage is reapplied. The amount of current necessary to restore the diode to its original condition is then measured.

There is also the double sampling method. In order to conserve area in the design of the shift register, the outputs of adjacent pairs of diodes are connected together. When the integration time is complete, the output of the sum of the pair of diodes is put on to the video line (see Figure 2.6). Following this, one of the diodes is recharged and the remaining signal is placed on the line. Finally, the second diode is recharged and the voltage now on the video line is taken as the zero level for this scan. Simple subtraction of the appropriate voltage levels will give the signals from each individual diode. The Plessey diode array was designed to operate in this way. This method eliminates the fixed pattern noise of the array caused by the variation in the DC offset voltages of the individual diodes.

By far, the most advantageous characteristic of the Plessey diode array is the availability of a non-destructive readout (NDRO) facility. This is achieved by operating the device in the double sampling mode with the recharge pulses suppressed. In effect, during a NDRO it appears as if there is an array of 128 double width diodes. This means that the resolution is halved, but, as NDRO is only used to monitor the build-up of the signal, it is not considered a disadvantage. However, one drawback of double sampling is that the dynamic range of the device is reduced. Since the signals on two diodes are being summed together, the double sampling method will cause the ADC to receive its maximum input signal when in fact the signal on an individual diode will only be about half of this maximum value.

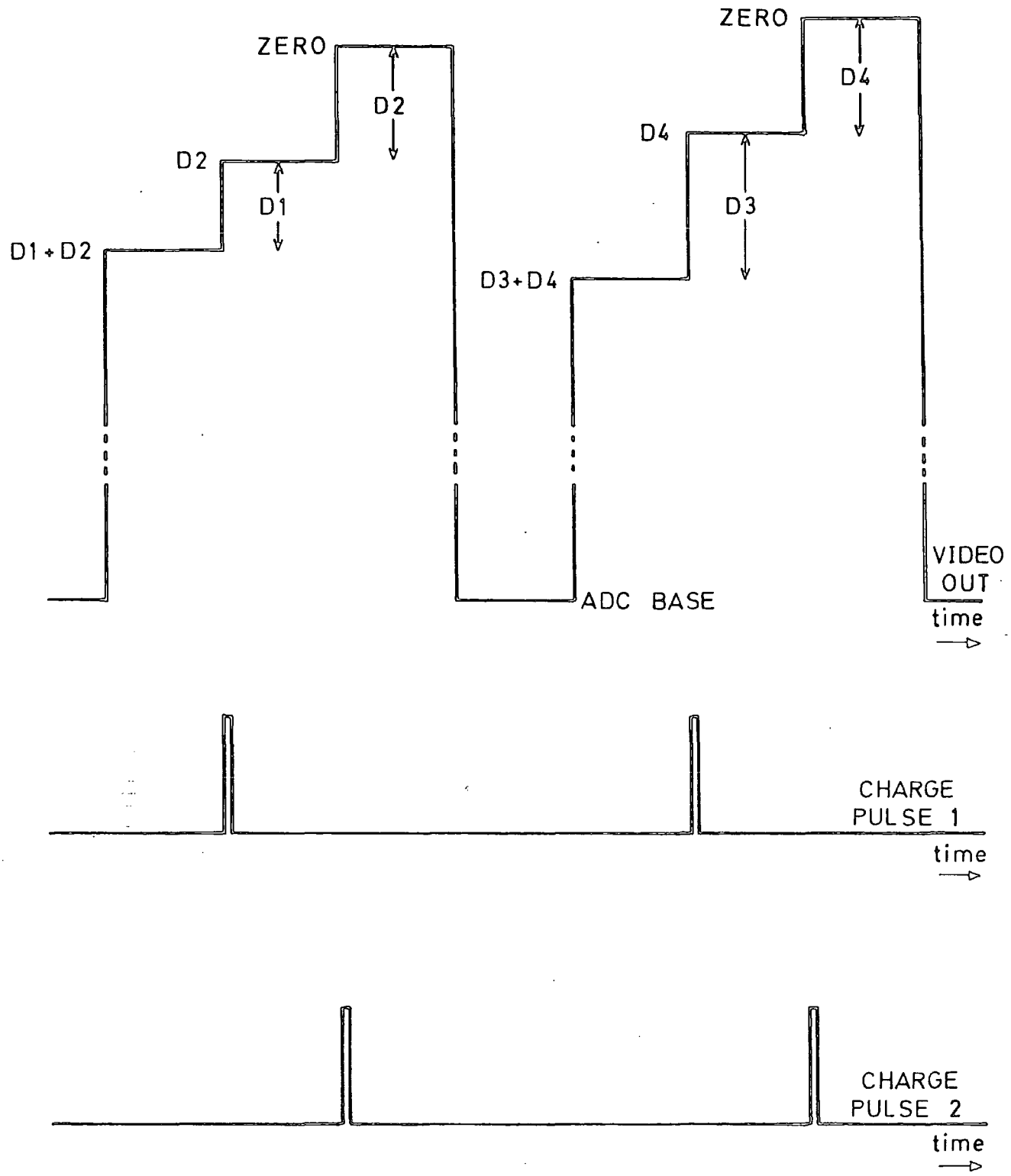


FIG. 2.6: EXAMPLE OF VIDEO OUTPUT USING DOUBLE SAMPLING METHOD

2.5 Laboratory Tests.

Laboratory tests were carried out on the Plessey diode array both before and after the system was used at the RGO. These tests included investigating the effect of temperature on the spectral response and observing the sensitivity of a particular diode. A random diode was chosen for these tests. The sensitivity test was carried out before the stellar observations took place. The spectral response tests were done after the observations, using a monochromator at the RGO.

2.5.1 Sensitivity.

Since a standard light source was not available, sensitivity measurements were carried out by first calibrating a non-standard light source using a Centronic OSI 5K hybrid photodetector. The OSI 5K was purchased already calibrated at three points in the spectrum (all measured at 25.5°C):

wavelength (μm)	responsivity ($mV.\mu W^{-1}.cm^{-2}$)
0.45	64.73
0.53	108.74
0.63	146.52

These three points were plotted on a graph, and, using the manufacturers relative response curve, an actual responsivity curve was approximated for wavelengths between 3500Å and 11000Å (see Figure 2.7). The circuit shown in Figure 2.8 was constructed for the OSI 5K and the device was mounted in a specially built "can" that was light-tight except for a glass window (see Figure 2.9). The window was first covered to make the can completely light proof, and the 10kΩ potentiometer was adjusted so that the dark current output could be set to an arbitrary DC level. Any

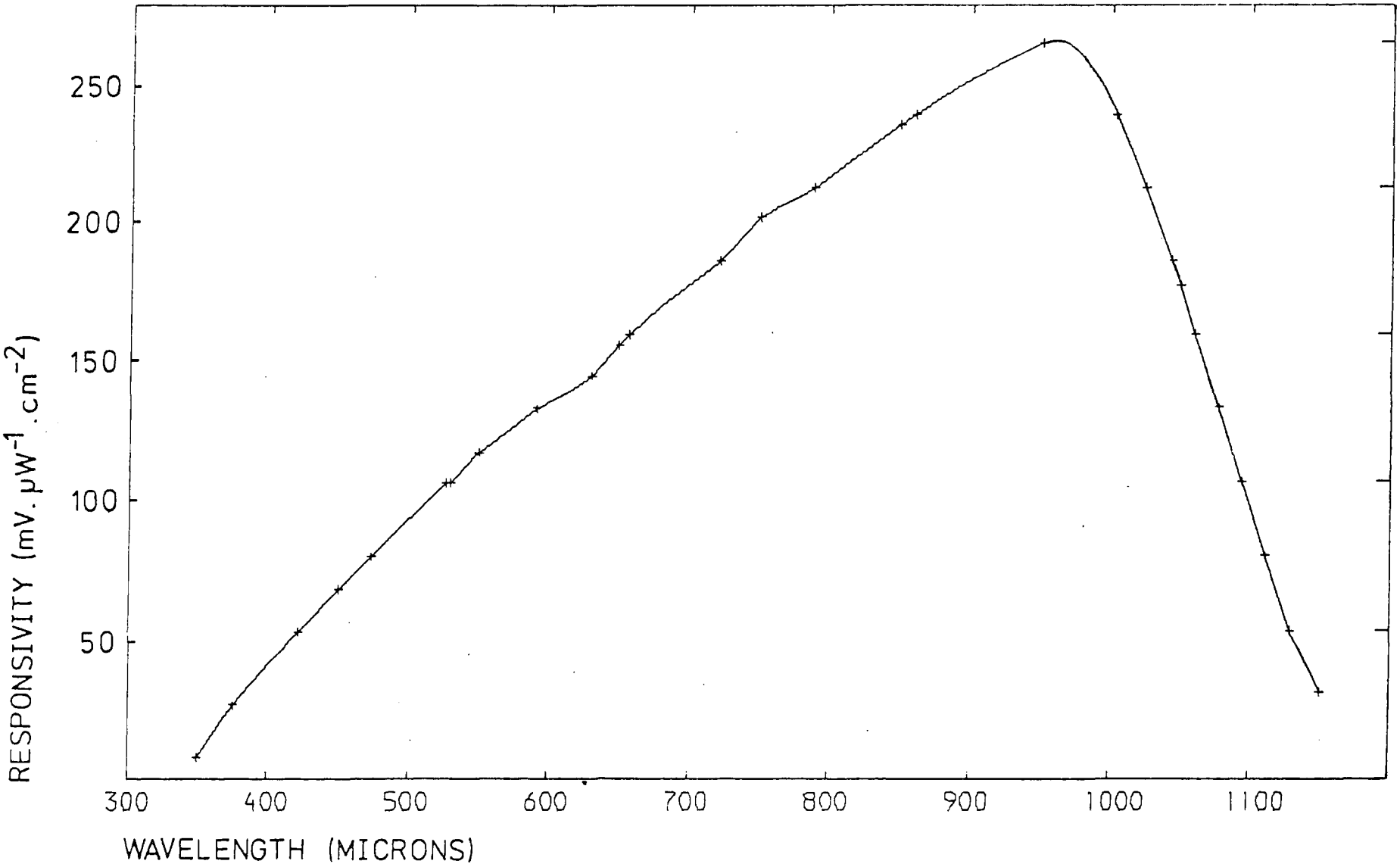


FIG. 2.7: SPECTRAL RESPONSE OF OSI 5K PHOTODETECTOR

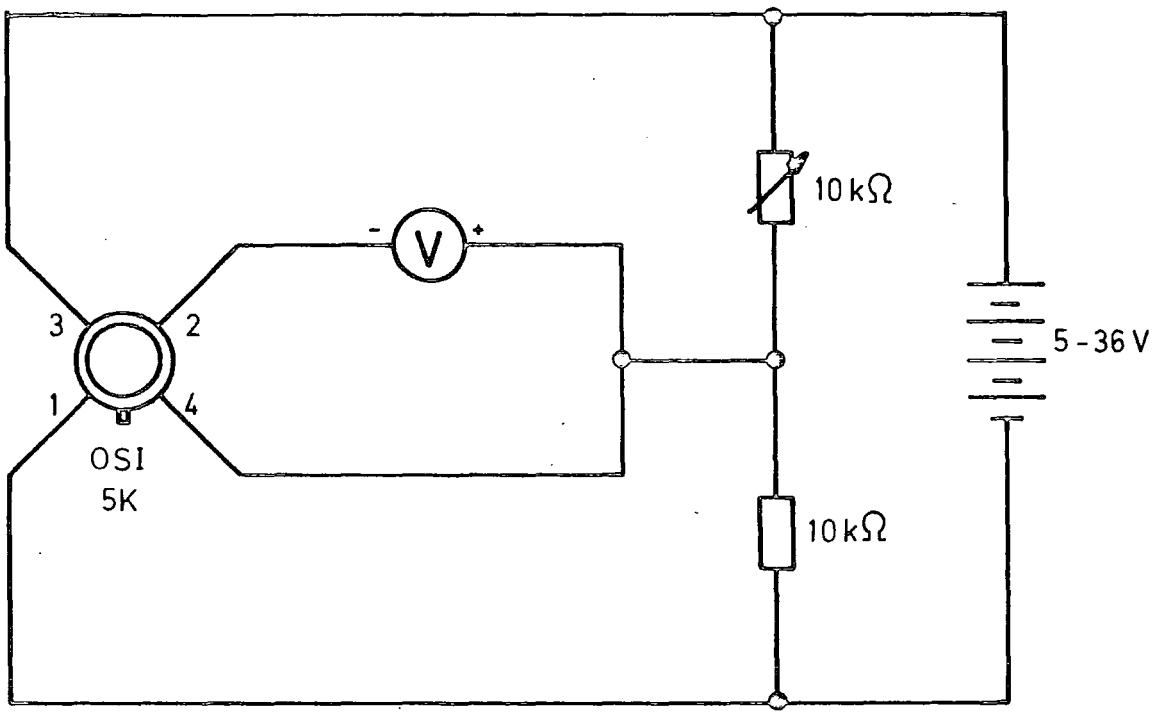


FIG. 2.8 : OPERATING CIRCUIT FOR THE OSI 5K PHOTODETECTOR

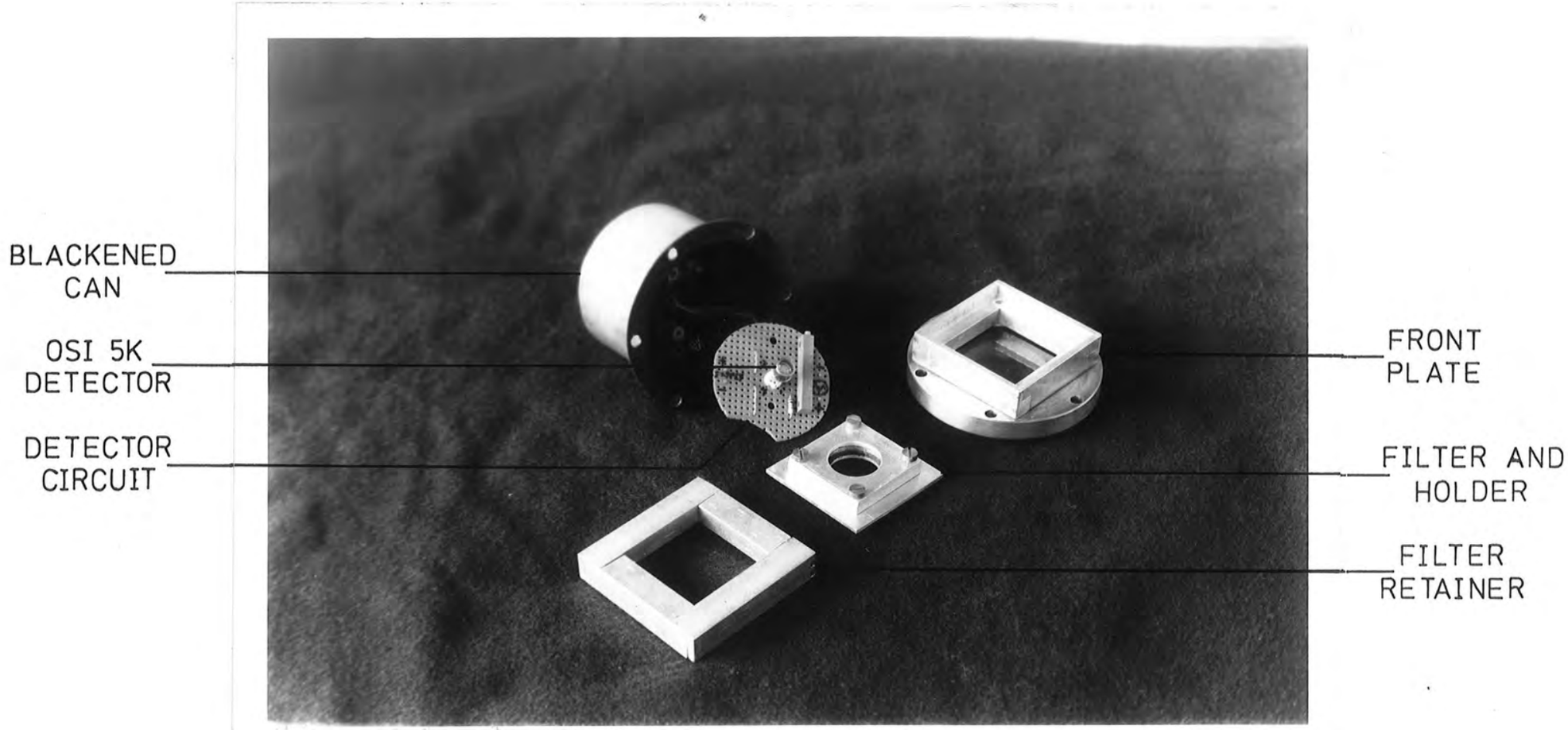


FIG. 2-9: CONTAINER ASSEMBLY FOR OSI 5K PHOTODETECTOR

output from the photodetector due to light falling on it was measured as a voltage above this "zero" level. Since the non-standard source was an ordinary 12V DC tungsten bulb, a narrow band optical filter was used to transmit only a small part of the white light spectrum onto the photodetector. The filter used was a Barr and Stroud type MD26 narrow band filter with the following characteristics:

serial number	MD26/2668
size	25mm diameter
wavelength at band centre	533.9nm (ie. mid-visible)
peak transmission	35%
transmission at 546nm	34%
band width at half peak	12.9nm
transmission >1075nm and <500nm	<0.05%

The light source calibration was carried out using the construction shown in Figure 2.10. The bulb was fixed in a brass can attached to the end of a brass tube approximately 0.65m long. The inside diameter of the tube was such that the can holding the OSI 5K fitted snugly inside. The tube was used both to eliminate stray light and to partially collimate the light from the bulb. The narrow band filter was located on the outside of the photodetector can, on the light source side of the window. The zero level of the OSI 5K was checked and the bulb operated at 6V, 0.3A. This gave an output from the OSI 5K of 4 mV. From this value, the flux of photons falling on the photodetector was calculated as being

$$\frac{(4/110) \times 10^{10}}{hc/\lambda} = 9.9 \times 10^{14} \text{ photons } .s^{-1}.m^{-2}$$

where λ= the band centre of the narrow band filter. The light source,

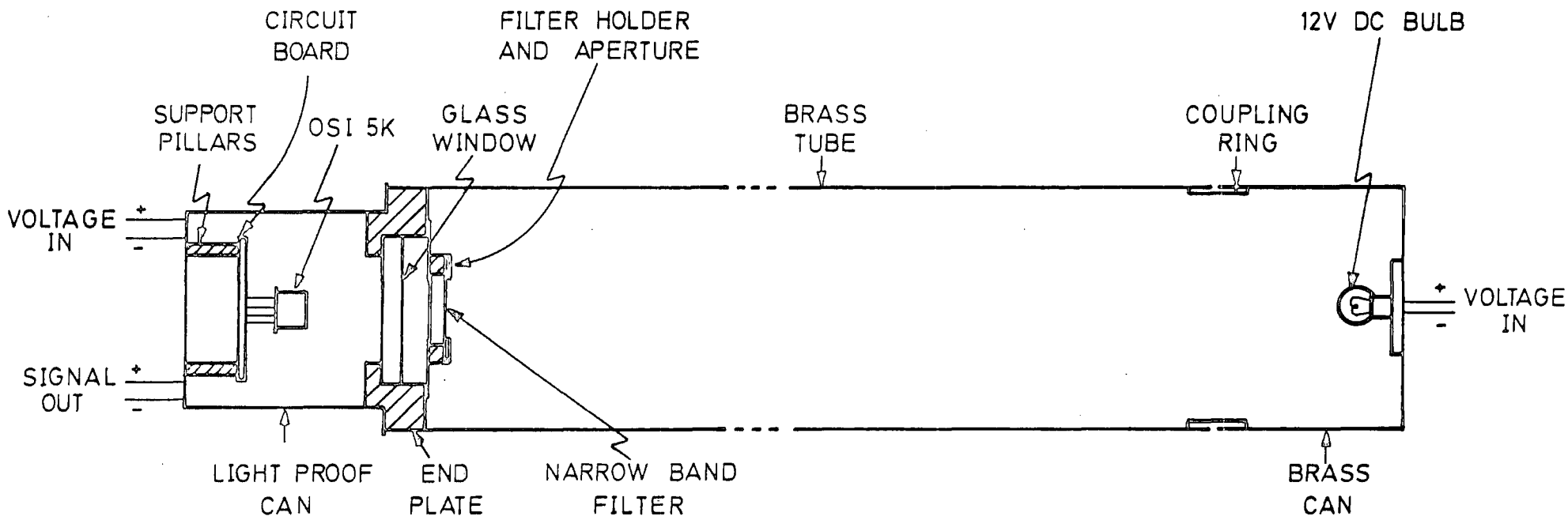


FIG. 2.10: CONSTRUCTION OF APPARATUS USED TO CALIBRATE A LIGHT SOURCE USING THE OSI 5K PHOTODETECTOR

collimator and filter were transferred to the part of the cryostat containing the Plessey diode array. For these responsivity measurements, the glass window in the cryostat was of the same type as that of the OSI 5K window. (Later, the cryostat window was replaced by an optical flat glass.) Under the same source conditions, the output of the array was measured as 3.4V for an integration time of thirty 11.2 ms frame time periods. Since the diode outputs were paired, the available sensitive area measured 1.6×10^{-8} m^2. The sensitivity of the array was calculated to be

$$\frac{3.4/30 \times 11.2}{9.9 \times 10^{14} \times 1.6 \times 10^{-8}} = 0.63\mu V. \text{ (incident photon)}^{-1}$$

Assuming an absolute quantum efficiency of 63% at 533.9nm (Geary, 1975), the true sensitivity can be calculated as being $1.0\mu V.\ (\text{detected photon})^{-1}$. The theoretical value of the sensitivity of the diode array can be determined from the gate capacitance of transistor T_3. If each photon generates an electron-hole pair, and C_G 0.1 pF, one would expect a voltage signal of

$$\frac{Q}{C} = \frac{1.6 \times 10^{-19}}{0.1 \times 10^{-12}} = 1.6\mu V.\ \text{photon}^{-1}$$

The head amplifier of the diode array has a voltage gain of approximately 0.94. Therefore, the expected sensitivity is

$$1.6 \times 0.94 = 1.5\mu V.\ (\text{detected photon})^{-1}.$$

This value does agree fairly well with the observed value. Errors occur in the value of C_G, the true responsivity of the OSI 5K at an uncalibrated

wavelength, the assumed quantum efficiency of the diode array, and variations in the output of the light source. Better agreement between the observed and theoretical values could not be expected.

2.5.2 Spectral response

The spectral response of a randomly chosen diode was investigated using a monochromator at the RGO. The array was cooled to $\sim -120^{o}C$ via the method described in section 3.1. The monochromator was adjusted to emit light at 400nm, and the resultant signal on the diode due to this light was recorded. The wavelength of the monochromator output was increased in steps of 50nm up to a value of 950nm, with the diode signal level being recorded at each step. A bolometric calibration of the monochromator was carried out at the same values of wavelength between 400nm and 950nm. A graph was drawn of the response of the diode as a function of the wavelength of the incident light (see Figure 2.11). The array was then allowed to warm up to room temperature, and the described procedure was repeated at the same wavelength intervals. The spectral response of the diode at room temperature is shown in Figure 2.12. Both of these graphs show the characteristic response of a solid state diode as a function of wavelength, namely a gradual fall-off in response as the wavelength of the incident photons approaches the infrared, and the sharper cut-off in response at the shorter wavelength, blue end of the spectrum. The reasons behind the spectral response effects have been described in section 2.2 (of this chapter). To show the effect of temperature on the spectral response of the diode, the two previous graphs have been superimposed (see Figure 2.13). This graph shows clear evidence for the increase in response in the infrared region of the spectrum as the temperature of the diode increases (as previously described). The results collected here indicate that increasing the temperature of the diode array from $\sim -120^{o}C$ to $\sim 20^{o}C$ means that one can penetrate approximately

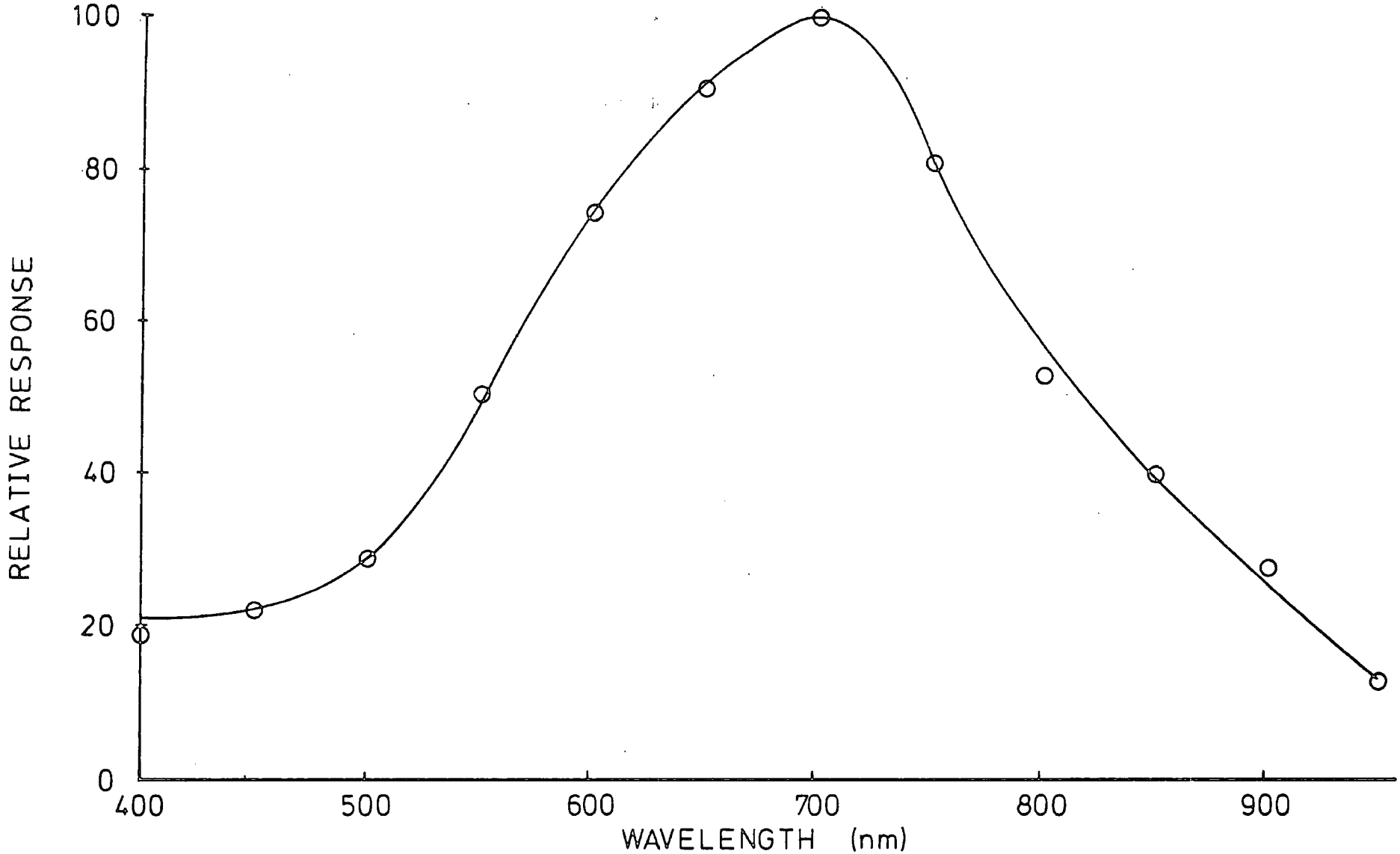


FIG. 2·11: SPECTRAL RESPONSE OF A TYPICAL ARRAY ELEMENT AT -120° C

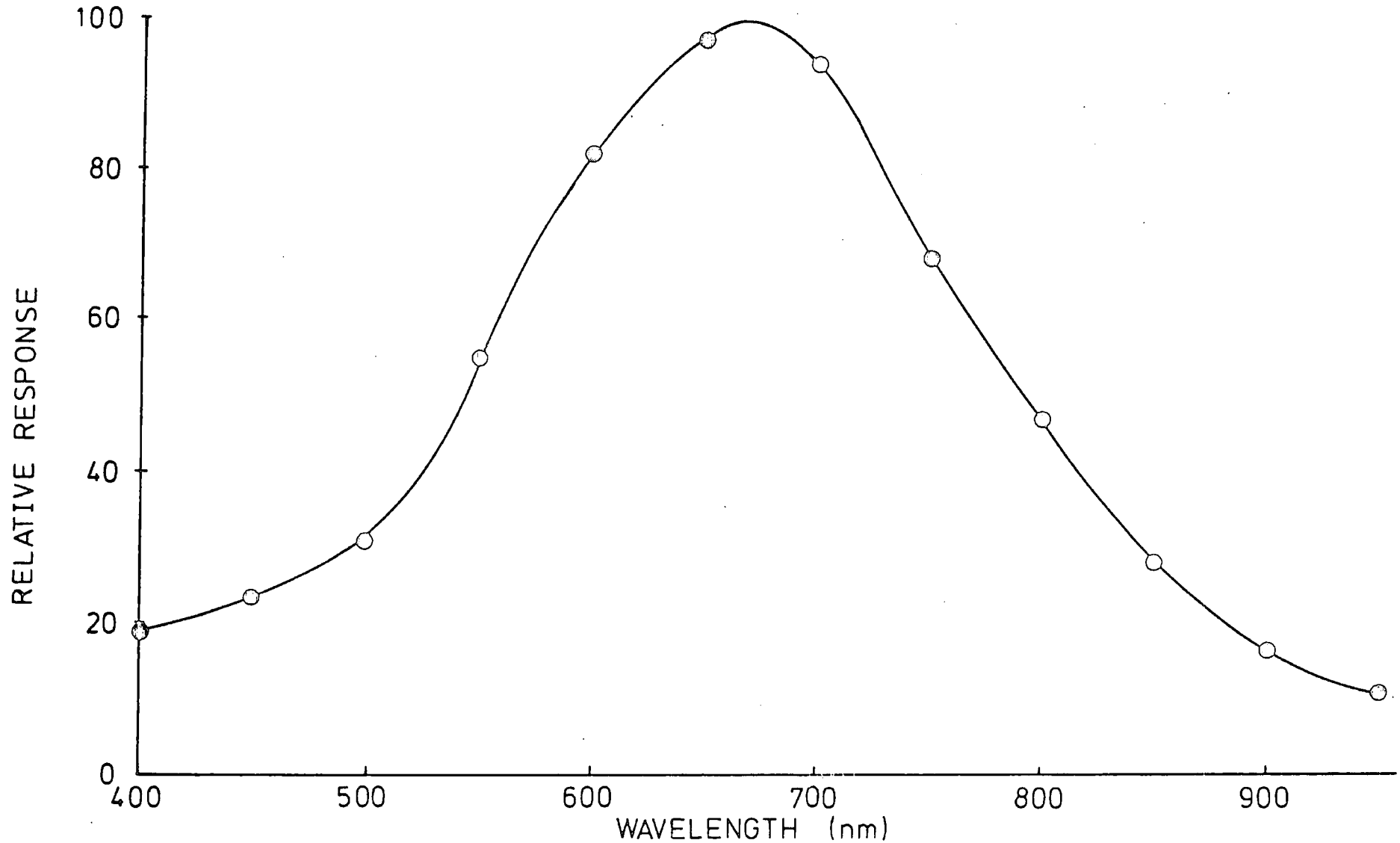


FIG. 2·12: SPECTRAL RESPONSE OF A TYPICAL ARRAY ELEMENT AT 20°C

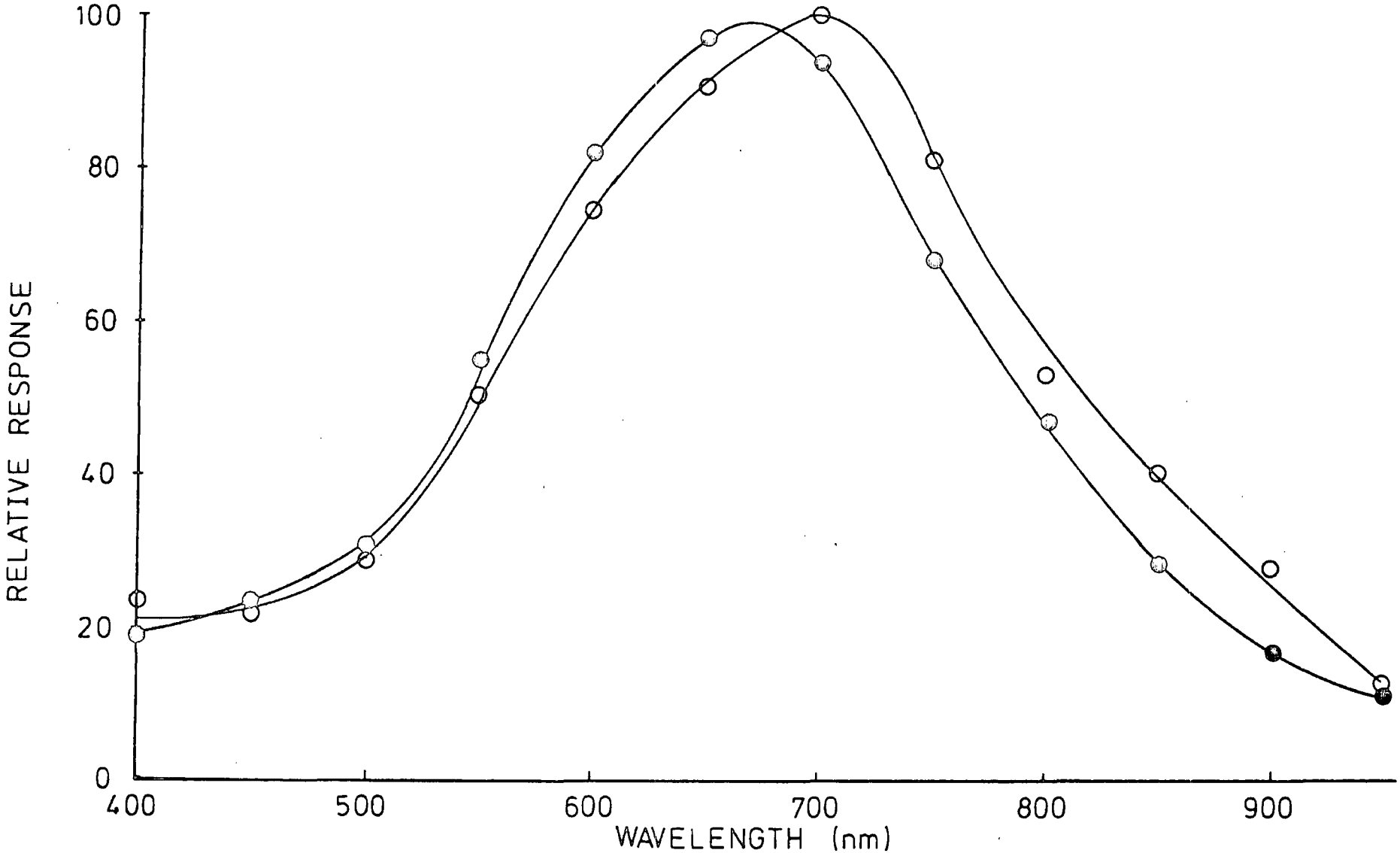


FIG. 2·13: CHANGE IN DIODE RESPONSE AS A FUNCTION OF TEMPERATURE

50nm further into the infrared region of the spectrum.

This value of 50nm compares favourably with the value of 40nm for the infrared response change calculated in section 2.2. The measured spectral response curves are prone to inaccuracies in the longer wavelength ranges. The maximum wavelength attainable by the monochromator is 1μm (i.e. full scale deflection), therefore wavelengths close to this value may be slightly incorrect. Another source of error is the calculation of the diode signal at room temperature. Noise constitutes a significant fraction of the total diode signal when it is not cooled. Moreover, the noise is not constant, and so a mean value of the observed noise was subtracted to give an approximate diode signal level.

REFERENCES

Geary, J.C.: 1975, Ph.D. Thesis, University of Arizona.

Geary, J.C.: 1976, Proceedings IAU Colloquium 40, 28-1.

Hedge, A.R.: 1981, Ph.D. Thesis, University of Durham.

Smith, R.A.: 1961, "Semiconductors", Cambridge University Press, Cambridge.

Vogt, S.S., Tull, R.G., Kelton, P.: 1978, Applied Optics, 17, 574.

CHAPTER THREE

INSTRUMENTAL CONSTRUCTION AND CONTROL

3.1 Temperature Measurement and Control

The construction of the original cryostat built for the Royal Greenwich Observatory is described here. The preliminary tests on the diode array were performed both at the RGO and at Durham using this cryostat. Temperature control was investigated using a copper-constantan thermocouple to measure the combined effects of a cold finger (with one end at liquid nitrogen temperature) and a 120kΩ resistor heater on the system. The temperature range investigated was from $\sim -110^{o}C$ to $\sim -20^{o}C$. Rates of cooling down and warming up of the cryostat without the use of the heater were also observed.

A new cryostat was designed and constructed for use at the coudé focus of the RGO 30" telescope. Physical dimensions were important here so as not to obstruct the light from the telescope in the limited available space.

For ease of temperature monitoring, an instrumentation amplifier was used in conjunction with the thermocouple and a digital voltmeter. With a gain of ~100, a convenient emf value of the thermocouple could be observed.

3.1.1 Original cryostat construction

The original cryostat construction can be seen in Figure 3.1. It consists of four parts, namely a cold finger section, a cryostat body section, and two end plates. The cold finger protrudes into the cryostat body on assembly. The detector array is then attached to the cold finger by means of stainless steel bolts and specially shaped pieces of copper. A 2-inch diameter glass disk is sealed into one of

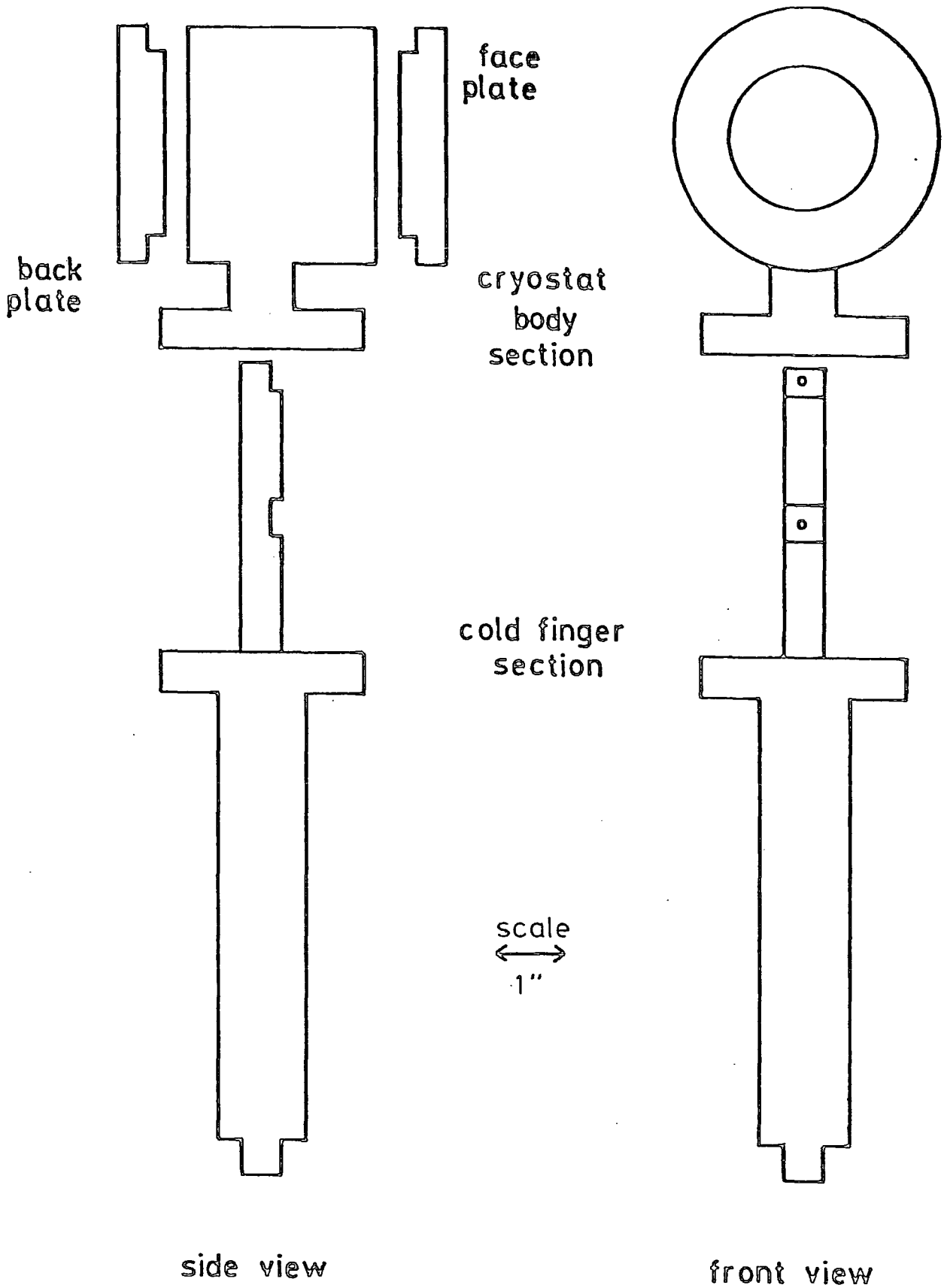


FIG. 3.1: CONSTRUCTION OF ORIGINAL CRYOSTAT

the end plates to allow light to enter the cryostat. Electrical connections from the inside to the outside of the cryostat are made via vacuum tight terminals on the back end plate.

The entire cryostat is connected to a vacuum pump assembly (see Figure 3.2). Evacuating the cryostat minimizes heat transfer from external sources and prevents "icing-up" of the cryostat interior at very low temperatures ($\sim -120^{o}C$). A vacuum of $\sim 4 \times 10^{-2}$ torr was achieved at room temperature with the assembly shown. After cooling, a pressure of $\sim 1.5 \times 10^{-2}$ torr was reached inside the cryostat. This improvement in reduced pressure is due to residual vapors being frozen out at low temperatures.

The vacuum coupling is located on the side of the cryostat body. A Leybold vacuum lock flange was fitted to allow the vacuum in the cryostat to be maintained after disconnecting the vacuum pump.

A detailed view of the diode array to cold finger mounting can be seen in Figure 3.3.

All heat conducting joints are improved using Radio Spares heat sink compound. With clean joints, the lowest temperature reached was only $\sim -90^{o}C$, 30 degrees warmer than when the heat sink compound had been used.

The temperature of the diode array can be altered by changing the length and/or the material of the mounting bolts on the cold finger. For the transfer of heat through a metal rod, the following equation holds:

$$\dot{Q} = -kA \frac{\theta_1 - \theta_2}{L}$$

where $\dot{Q}$ = rate of flow of heat through the metal rod

θ_1, θ_2 = temperatures at the ends of the rod

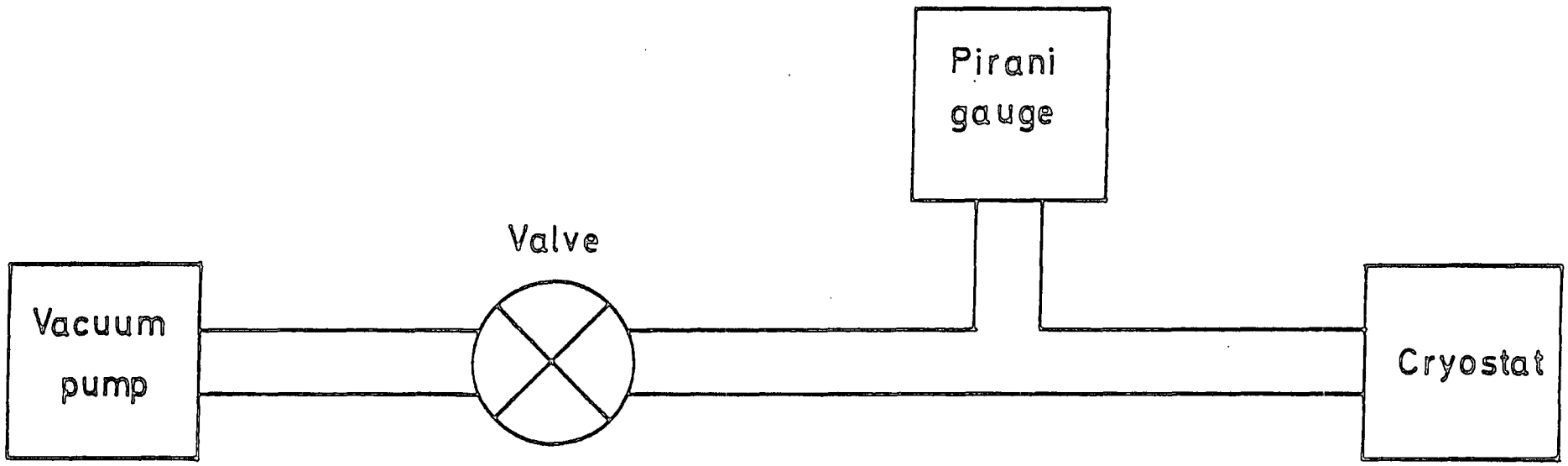


FIG. 3.2: BLOCK DIAGRAM OF VACUUM SYSTEM

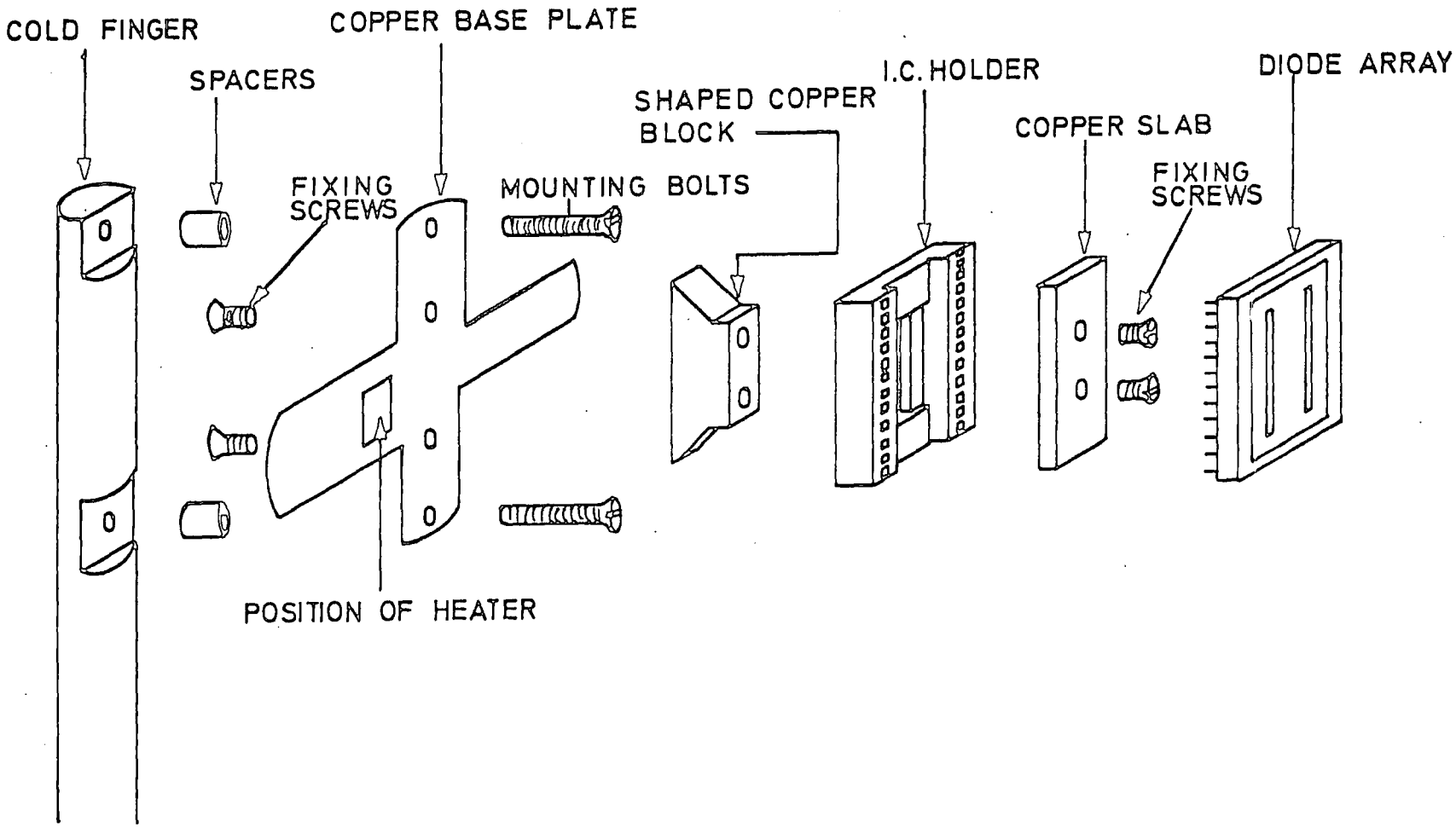


FIG. 3.3: EXPLODED VIEW OF DIODE ARRAY MOUNTING

k = mean coefficient of thermal conductivity between the temperatures θ_1 and θ_2

A = cross-sectional area of the rod

L = length of the rod

By altering the size or material of the mounting bolts, a coarse temperature variation can be obtained. Brass mounting bolts were tried at first (k(brass) = 0.52 Kcal/sec m ^{o}C), but it was found that stainless steel bolts (k(st.steel) = 0.08 Kcal/sec m ^{o}C) resulted in a more suitable diode array temperature.

Aside from any fixing screws (which are made of brass), all other parts of the heat conduction path are made of copper. Tracing the heat path from the coldest point, the lower end of the cold finger protrudes through the lower cryostat section and is immersed in a bath of liquid nitrogen contained in a dewar. At the top of the cold finger, two parallel stainless steel bolts are screwed in perpendicular to the length of the cold finger. The bolts hold a cross-shaped copper base plate onto which the heater, thermocouple and the earth connections are made. A specially shaped block of copper is screwed onto the base plate. The end of the shaped block passes through a 24 pin DIL IC socket. A rectangular slab of copper bolts onto the shaped block to hold the IC socket in place. Heat sink compound is spread onto the exposed face of the copper slab. This face is in contact with the back of the diode array chip so a good thermal contact is required to keep the chip cold. A copper-constantan thermocouple originally monitored the temperature at the base plate. However, the thermocouple was later transferred to a position sandwiched between the shaped copper block and the copper slab. In this position the thermocouple maintained direct thermal contact with the back of the chip. This gave a better indication of the true temperature of the chip (see section 3.1.3).

The reference junction of the thermocouple was originally at room temperature. However, thermocouple temperature calculations with varying reference junction temperatures proved to be tedious, so the reference junction was later kept at $0^{o}C$ in a melting ice bath. This method allowed the thermocouple temperature to be read directly from a graph, since copper-constantan thermocouple calibrations with $0^{o}C$ reference junctions have been accurately recorded (see Figures 3.4a and 3.4b).

3.1.2 Thermocouple calibration

Even though accurate thermocouple potential values are published for given temperatures (Kaye and Laby, 1960), each individual thermocouple will exhibit slight emf variations from the true values. These variations result in a small temperature discrepancy for a given measured temperature, the value of which need only be added to the observed temperature to give a true reading.

Table 3.1 shows the standard temperatures used as calibration points for the thermocouple in the cryostat.

The values of the temperature differences $(t_t - t_c)$ were plotted against the calculated temperature values (t_c) in Figure 3.5. A least squares fit drawn through these points indicates the value required to be added to a particular observed temperature to obtain the true temperature. It can be seen that in the region under observation ($-20^{o}C$ to $-110^{o}C$) the temperature difference ranged from $-1^{o}C$ to $-6^{o}C$. For all subsequent temperature measurements, the corresponding correction value has been added.

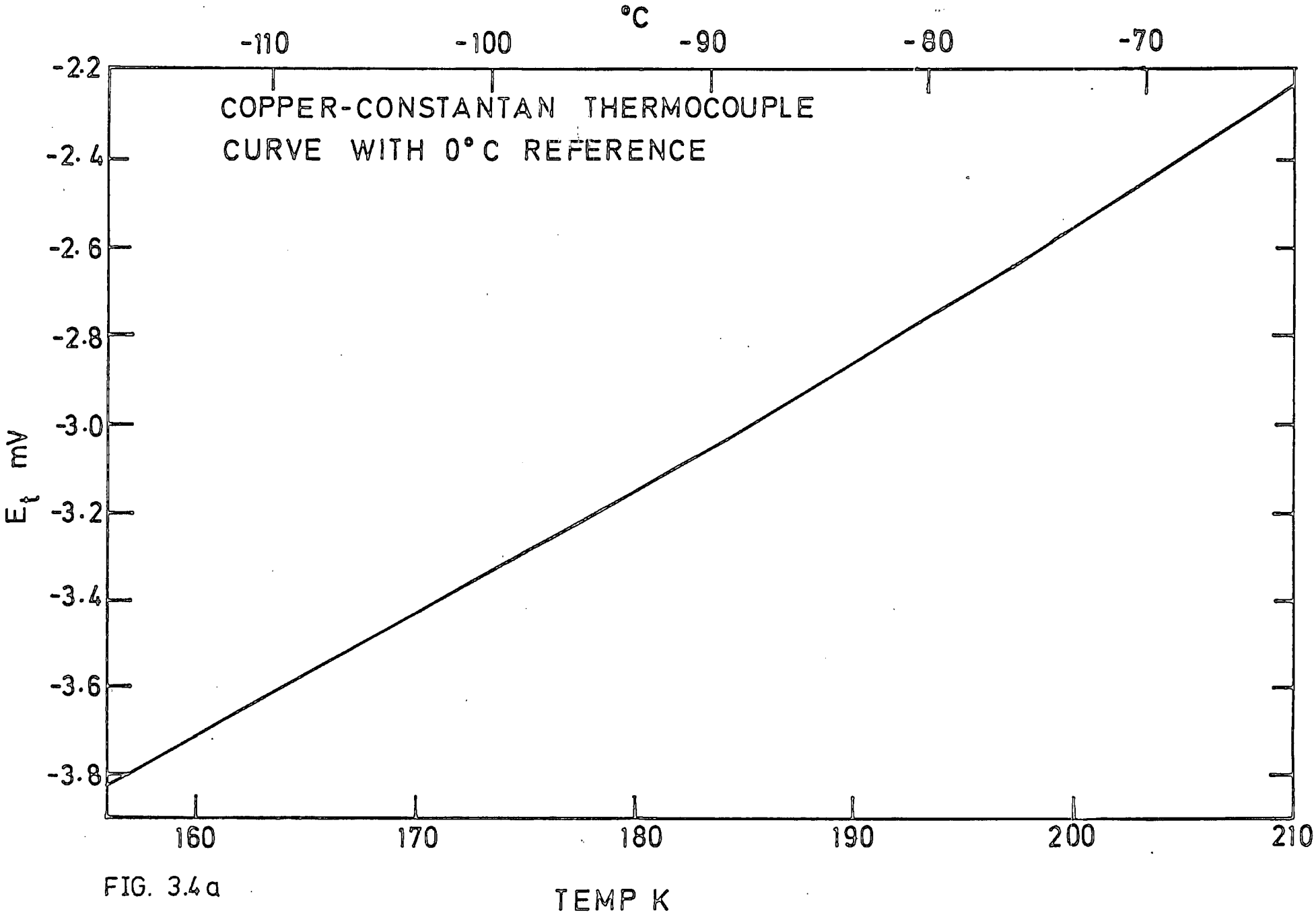

°C
-110
-100
-90
-80
-70
COPPER-CONSTANTAN THERMOCOUPLE
CURVE WITH 0° C REFERENCE
-2.2
-2.4
-2.6
-2.8
-3.0
-3.2
-3.4
-3.6
-3.8
E_t mV
160
170
180
190
200
210
TEMP K

FIG. 3.4a

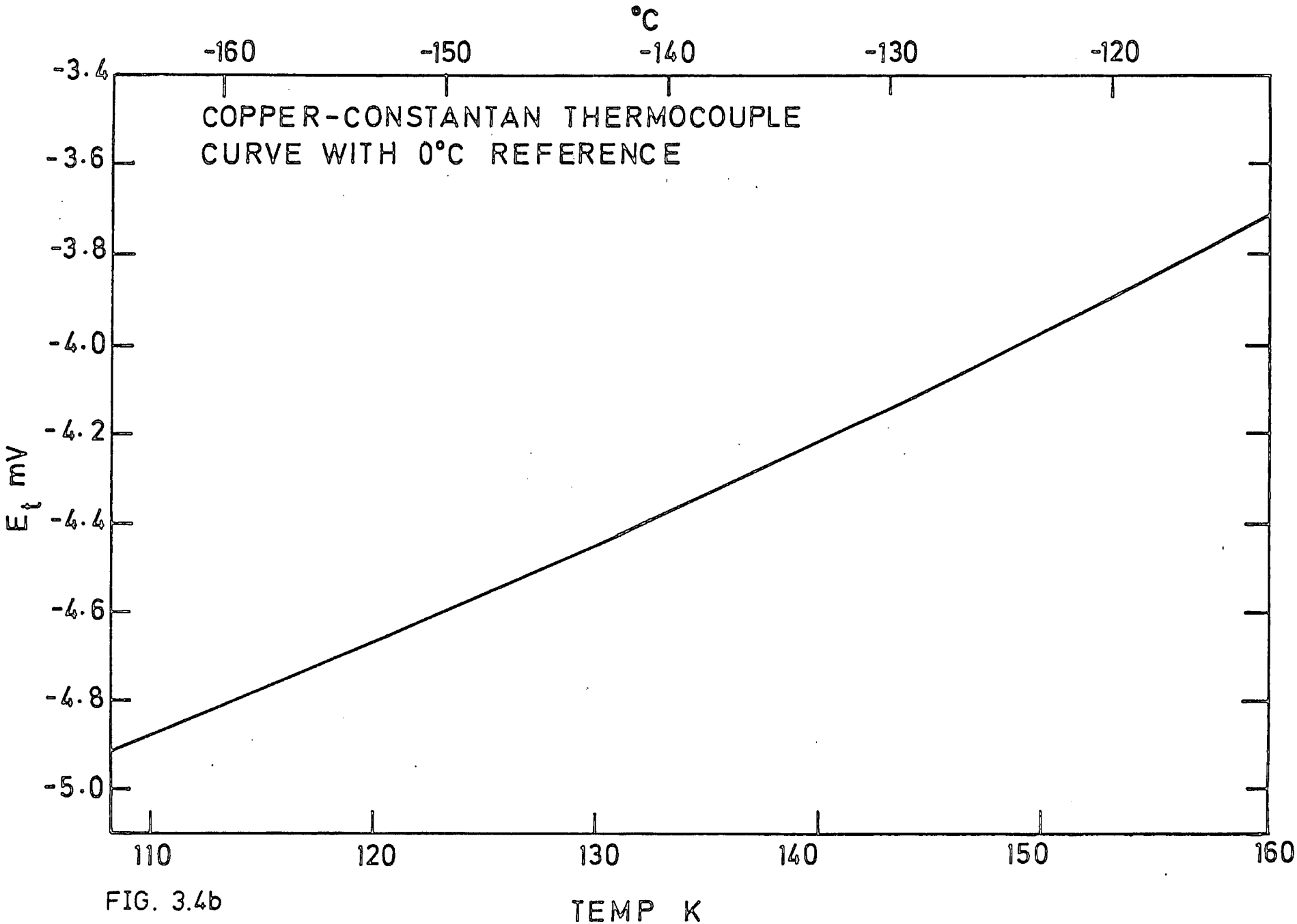

°C
-160
-150
-140
-130
-120
COPPER-CONSTANTAN THERMOCOUPLE
CURVE WITH 0°C REFERENCE
-3.4
-3.6
-3.8
-4.0
-4.2
-4.4
-4.6
-4.8
-5.0
E_t mV
110
120
130
140
150
160
TEMP K

FIG. 3.4b

Table 3.1

Standard temperatures used in the calibration of the copper-constantan thermocouple

Source	True Temperature t_t °C	Measured emf mV	Calculated Temperature t_c °C	$t_t - t_c$ °C
Liquid nitrogen	-195.8	-5.310	-186.0	-9.8
Acetone/ dry ice	- 78.5	-2.600	- 74.6	-3.9
Melting ice	0.0	-0.015	- 0.3	+0.3
Boiling water	+100.0	+4.040	+ 94.9	+5.1

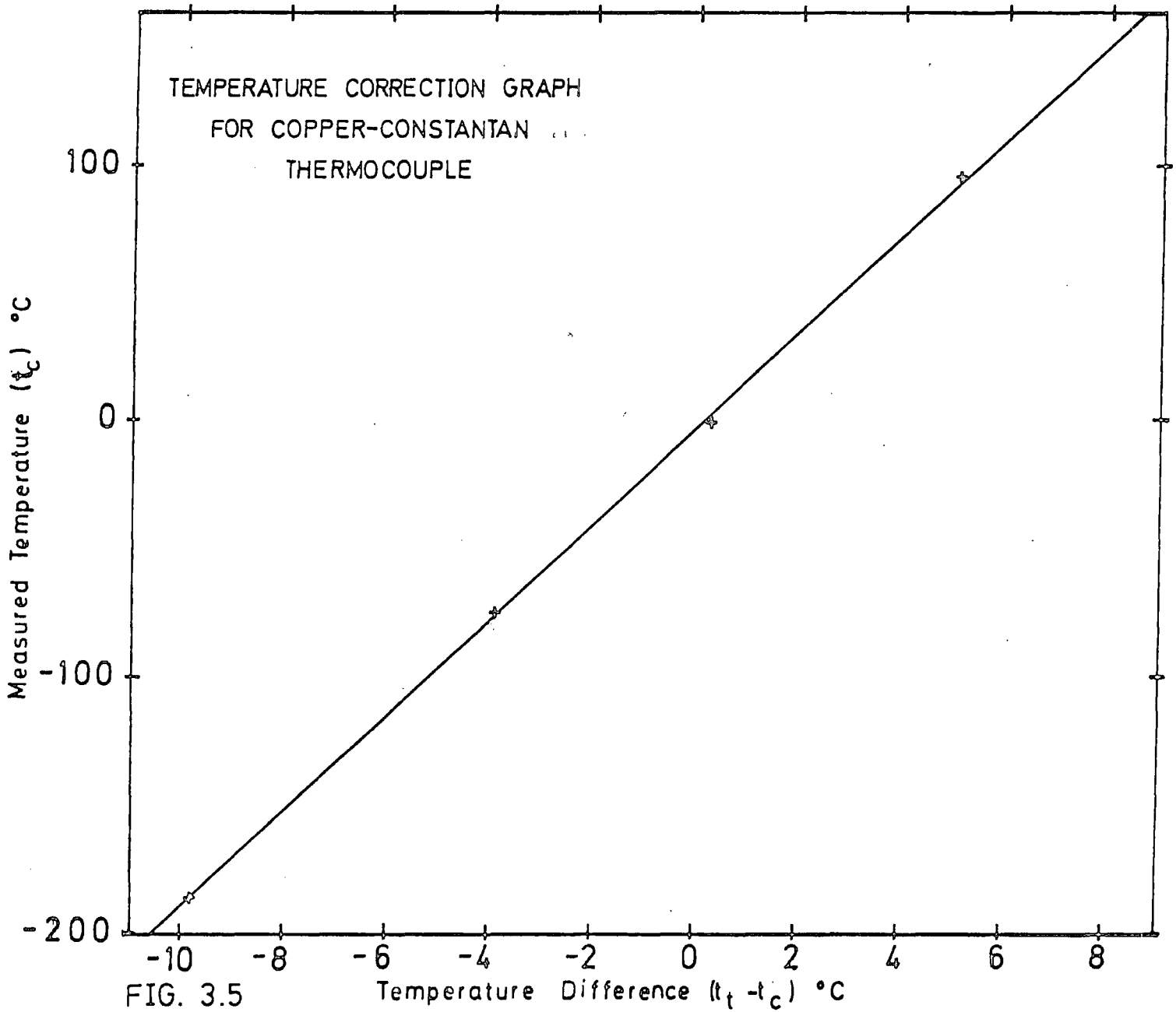

TEMPERATURE CORRECTION GRAPH
FOR COPPER-CONSTANTAN
THERMOCOUPLE
100
0
-100
-200
Measured Temperature (t_c) °C
-10
-8
-6
-4
-2
0
2
4
6
8
Temperature Difference ($t_t - t_c$) °C

FIG. 3.5

3.1.3 Temperature differences

In order to investigate temperature variations across the copper base plate in the cryostat, the thermocouple (TC1) was repositioned onto the uppermost of the cold finger mounting bolts on the base plate. A second thermocouple (TC2) was fitted onto the lower cold finger mounting bolt on the base plate and the temperature difference across the copper plate was measured. At a mean temperature of $-130^{o}C$, the variation across the copper base plate was $\sim 10^{o}C$. In order to give a better indication of the temperature of the array itself, TC1 was moved nearer to the array. It was sandwiched between the shaped copper block and the rectangular copper slab (see Figure 3.6). The measured temperature at this point behind the chip was $-122^{o}C$, while the lower mounting bolt was at a temperature of $-134^{o}C$.

3.1.4 Rates of change of temperature

With the cryostat evacuated to $\sim 4 \times 10^{-2}$ torr, the liquid nitrogen was poured into a dewar surrounding the cold finger. The temperature inside the cryostat was recorded at several minute intervals after the first addition of liquid nitrogen. The level of liquid nitrogen in the dewar was maintained such that the exposed copper part of the cold finger was always submerged in the coolant. A graph was plotted of the temperature of the thermocouple inside the cryostat against time for several cooling down procedures. Figure 3.7 shows a typical cooling curve. Errors in the reading of the potentiometer were within $\pm$ 0.005 mV, giving temperature errors of $\pm 0.3^{o}C$ near room temperature and $\pm 0.4^{o}C$ at temperatures near $-120^{o}C$. Timing errors were negligible ($\pm$ 5s) - much too small to be indicated on the graph. It was observed that after $\sim$100 minutes, the temperature inside the cryostat had reached a level

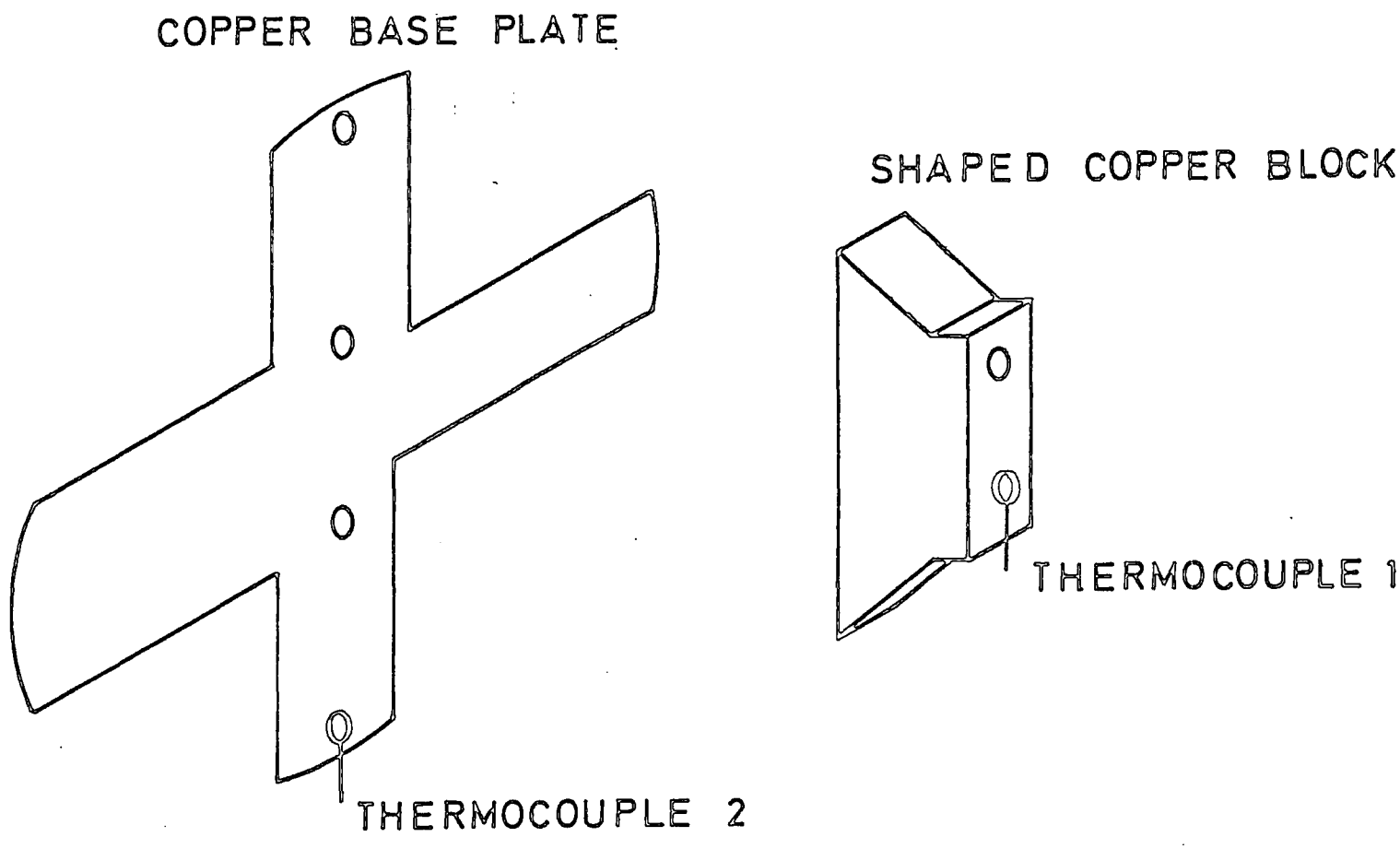


FIG. 3.6 LOCATION OF THERMOCOUPLES

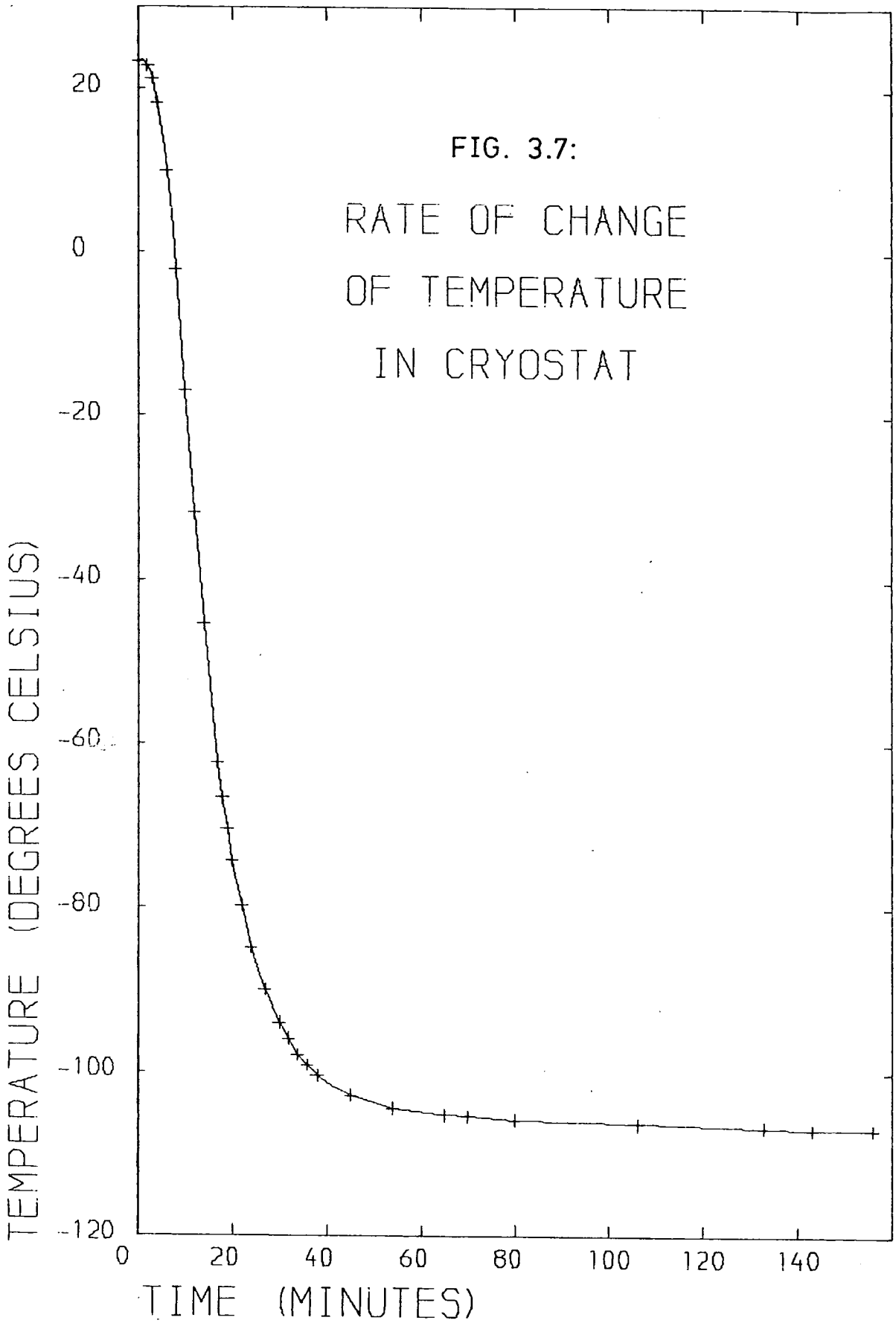


FIG. 3.7:
RATE OF CHANGE
OF TEMPERATURE
IN CRYOSTAT
TEMPERATURE (DEGREES CELSIUS)
20
0
-20
-40
-60
-80
-100
-120
0
20
40
60
80
100
120
140
TIME (MINUTES)

where the variation was not more than $\pm$ 1°C.

Figure 3.8 shows the effect on temperature when all of the coolant is allowed to evaporate. Errors are of similar value to those of the previous graph. Zero time was taken as the time of the last addition of liquid nitrogen. The temperature was monitored while the amount of coolant dropped from the maximum height in the dewar to below the level of the base of the cold finger. A typical time of approximately 2.5 hours could elapse before the temperature would begin to rise inside the cryostat, provided that the measured temperature was initially below -100°C. However, even if the temperature was only -50°C upon addition of the final LN_2, after a further 50 minutes of cooling down, the cryostat stabilised for about 80 minutes. The difference between stabilised temperatures for these two cases was only ~2°C, not significant enough to seriously affect the characteristics of the diode array. Furthermore, the overall level of liquid nitrogen in the dewar did not affect the temperature, so long as the exposed base of the cold finger remained immersed in the coolant.

3.1.5 Effect of heater - Temperature reproducability

With the cryostat cooled down to a reasonably stable temperature, a voltage was applied across the 120.4Ω resistor heater. The graph of Figure 3.9 shows the effect of the heater on the temperature inside the cryostat. Different heater voltages create different temperature plateaus on the graph. Within less than a minute of switching the heater on or off, the measured temperature changed.

The ability to reproduce a given temperature was investigated. The cryostat was first cooled down for at least one hour before a potential was applied to the heater, and the temperature was monitored. After stabilisation of temperature (at least 1.5 hours), the heater was

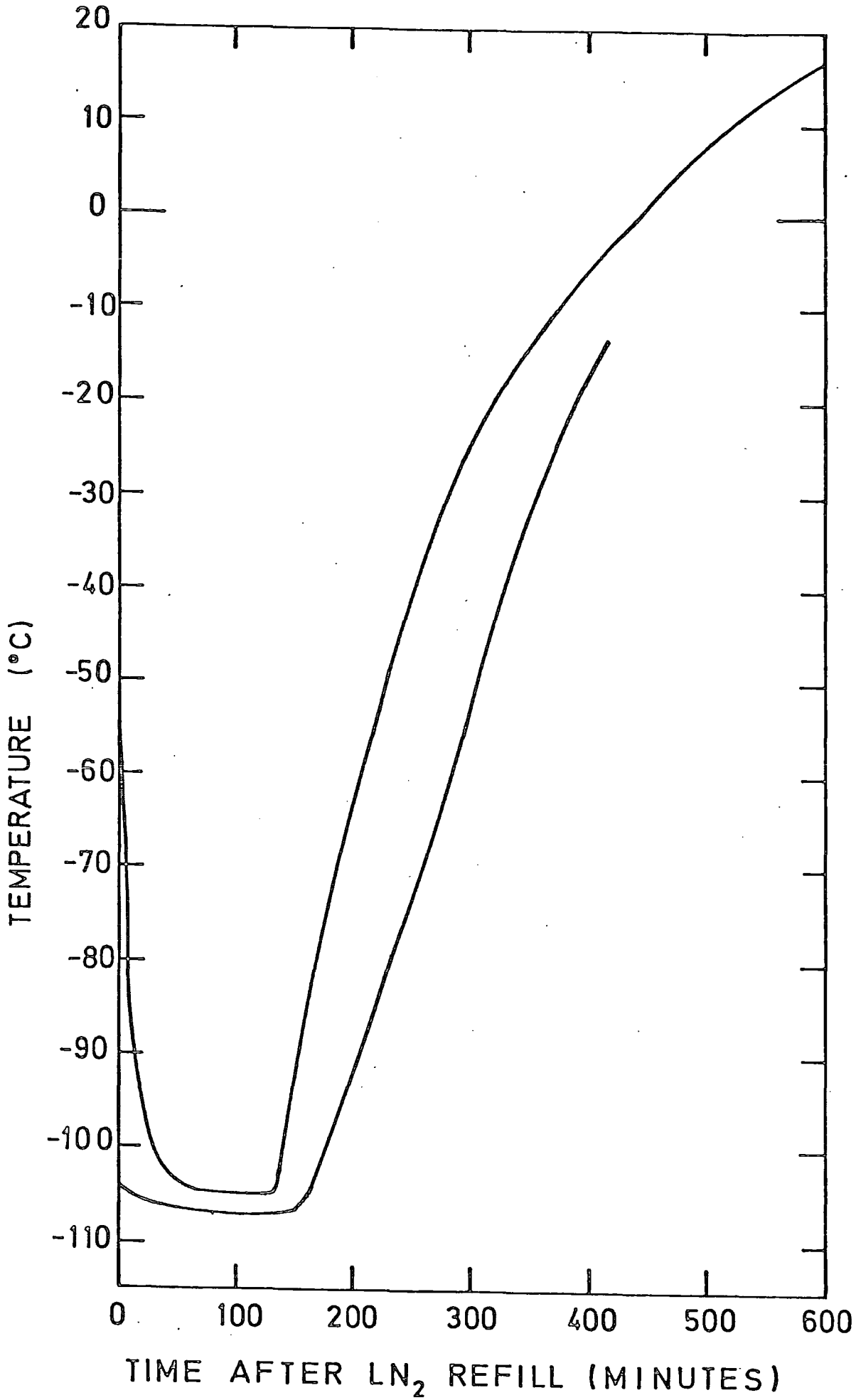


FIG. 3.8: EFFECT OF LOSS OF LN_2 ON CRYOSTAT TEMPERATURE

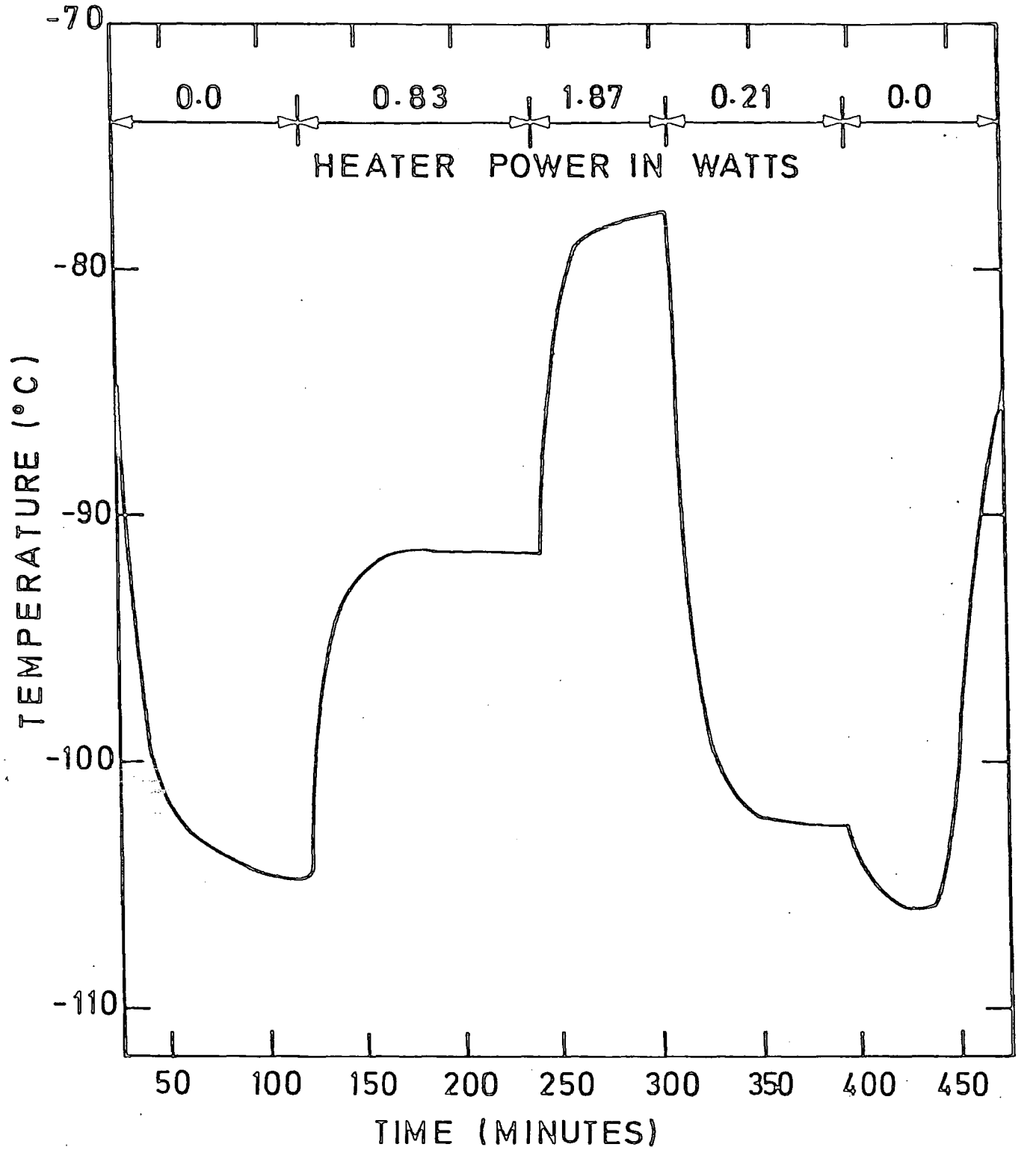


FIG. 3.9: EFFECT OF HEATER ON TEMPERATURE OF DIODE ARRAY

switched off for approximately 30 minutes to allow the cryostat to re-cool. The heating and cooling procedure was repeated several times for predetermined values of power dissipation of the heater. The results are illustrated graphically in Figures A1 to A7 in Appendix A. It can be seen that after the heater has been on for ~100 minutes, the temperature reached is constant to within $\pm$ 1.5°C for a given heater power. Taking the temperature measured after 100 minutes from each graph, a further graph was plotted of heater power dissipation against temperature (see Figure 3.10). Any desired temperature could therefore be attained inside the cryostat by applying the voltage determined by this graph.

3.1.6 New cryostat construction

A new cryostat and dewar were designed and constructed to conform to space limitations at the coudé focus of the 30" telescope at the RGO. The overall height of the cryostat had to be less than the original so as not to interfere with the light beam from the telescope (see Figure 3.11). This shortening of the cold finger resulted in two other problems:

1) a lower temperature of the chip than before,

2) a shorter dewar having to be refilled with liquid nitrogen more often.

The latter problem was solved simply by constructing a dewar of greater internal diameter (95 mm I.D.) than the original dewar (70 mm I.D.). The former problem was remedied by making the cryostat head longer, and mounting the array further from the cold finger, thus increasing the length of the cold path. Although an increase in the length of the cryostat head increases the overall volume to be evacuated, no significant changes in either evacuation time or pressure were encountered. Figure 3.12

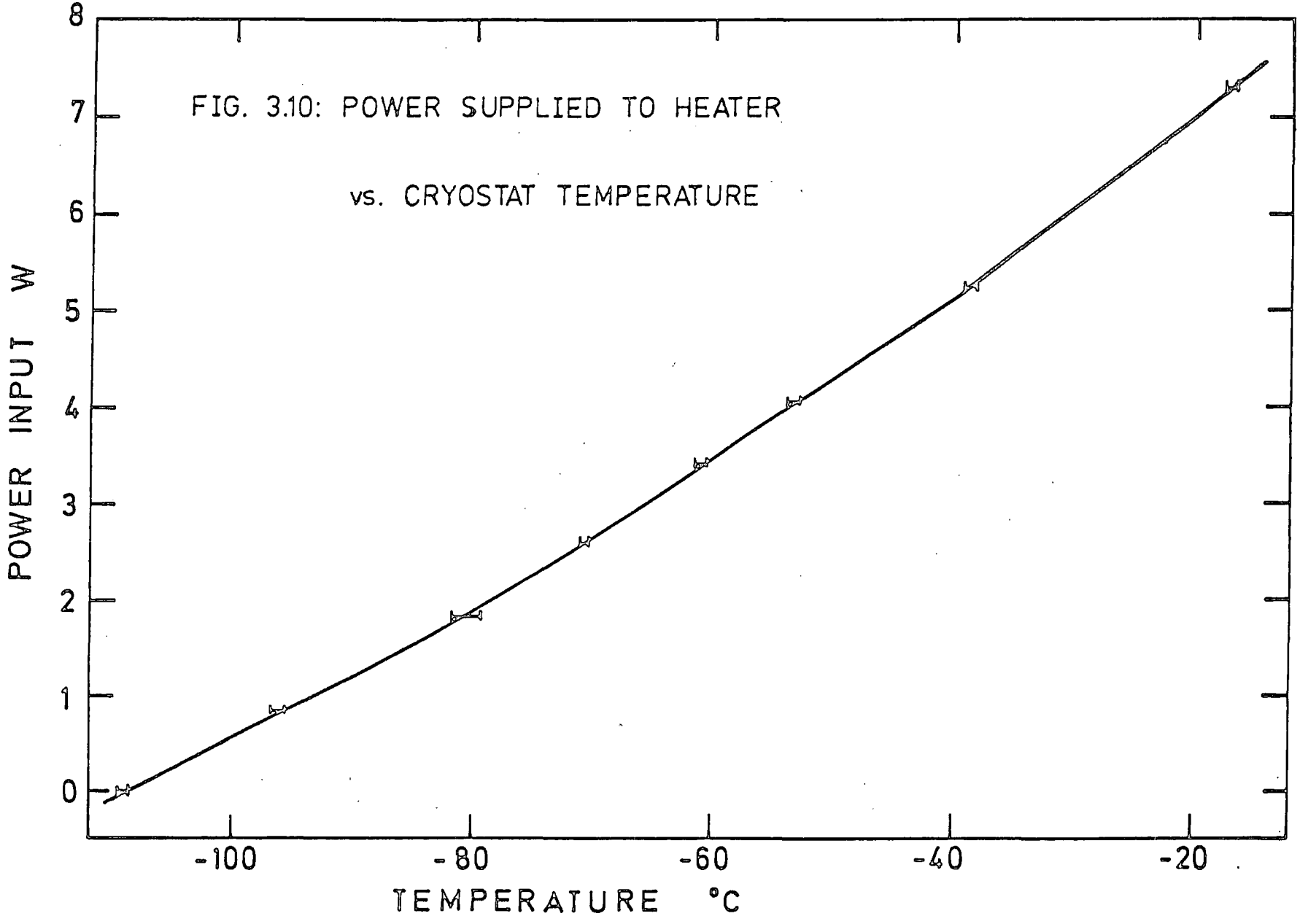


FIG. 3.10: POWER SUPPLIED TO HEATER
vs. CRYOSTAT TEMPERATURE
POWER INPUT W
8
7
6
5
4
3
2
1
0
-100
-80
-60
-40
-20
TEMPERATURE °C

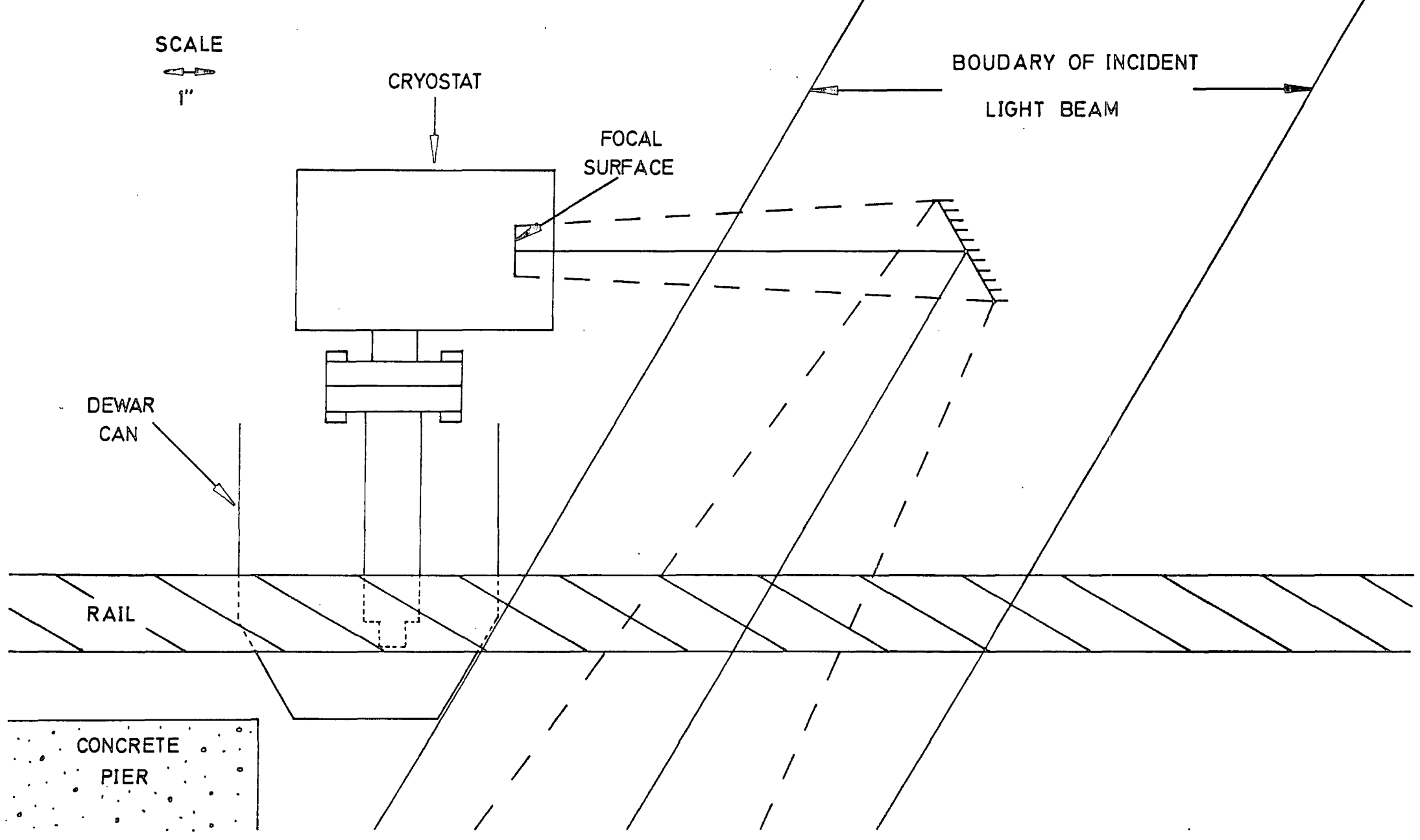


FIG. 3.11: COUDÉ FOCUS OF TELESCOPE

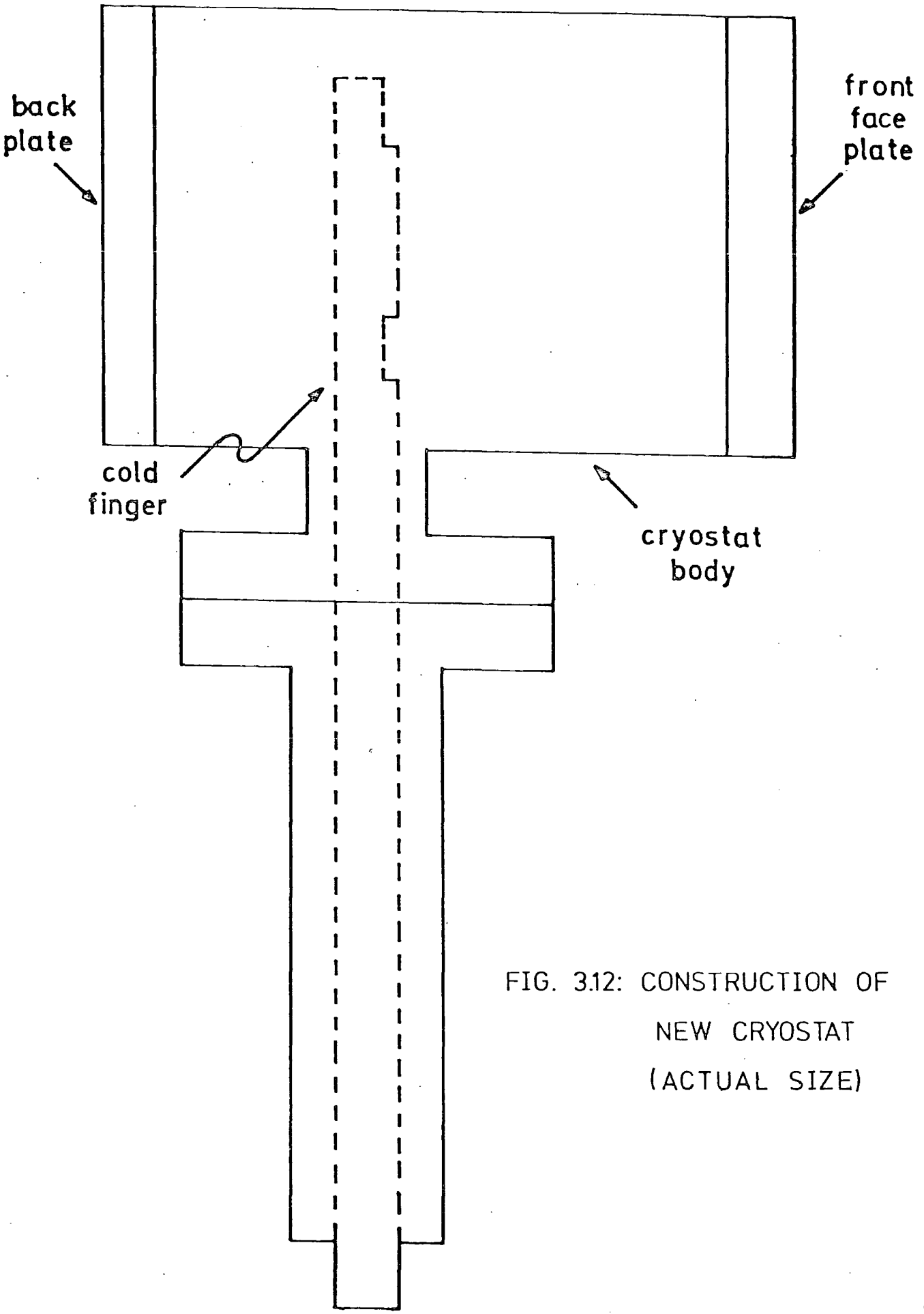


FIG. 3.12: CONSTRUCTION OF NEW CRYOSTAT (ACTUAL SIZE)

shows the construction of the new cryostat. Figure 3.13 illustrates the thermal connections to the cold finger. One possible problem with a longer cold finger to diode array path was discovered while the cryostat was in use at the RGO. The emerging stellar spectrum had to be accurately positioned on to the 200μm wide strip of diodes on the detector chip. Failure to do so was immediately evident upon examination of the results, causing splitting of the recorded spectrum. Movement of the array was done with a micrometer pushing the cryostat horizontally (see section 3.2). Once correctly aligned, it should not have been necessary to realign the array if the same region of the spectrum was being observed. However, frequent re-alignments were necessary. This could be explained by the geometrical properties of the cold finger section of the cryostat altering at different temperatures. Modifications to the existing cryostat are under consideration to give greater geometric stability to the diode array.

3.1.7 Thermocouple instrumentation amplifier

Measuring temperatures with a thermocouple can prove to be cumbersome, necessitating the use of a potentiometer, plus a power supply and a galvanometer. A better method would be to replace these with a digital voltmeter. Problems arise here because the emf generated by the thermocouple is only in the region of a few millivolts. At $\sim -100^{o}C$, a $1^{o}C$ temperature change will only result in a 0.02-0.03 mV change in the emf generated by the thermocouple. This voltage is much too small to give accurate temperature indications on a DVM (accurate to only $\pm \sim 1$ mV).

A Burr Brown 3660 instrumentation amplifier (IA) was chosen for its high gain (1 to 1000 V/V) and low drift ($<2.5\mu V/^{o}C$ at a gain of 1000 V/V) characteristics. The circuit diagram is shown in Figure 3.14. The offset voltage is adjusted for 0V output measured on the DVM when the two inputs are short circuited to earth.

The IA had also to be calibrated. This was done by taking the

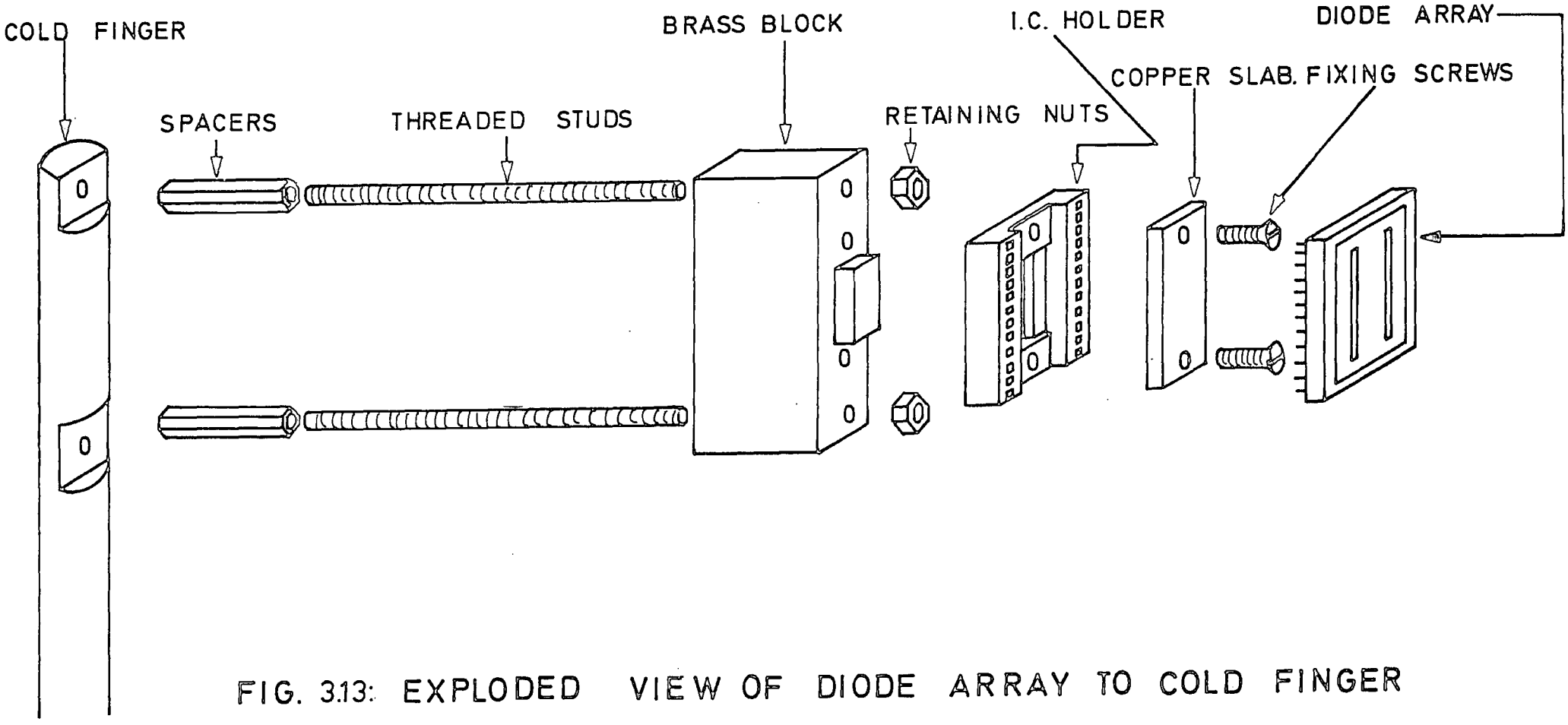


FIG. 3.13: EXPLODED VIEW OF DIODE ARRAY TO COLD FINGER MOUNTING

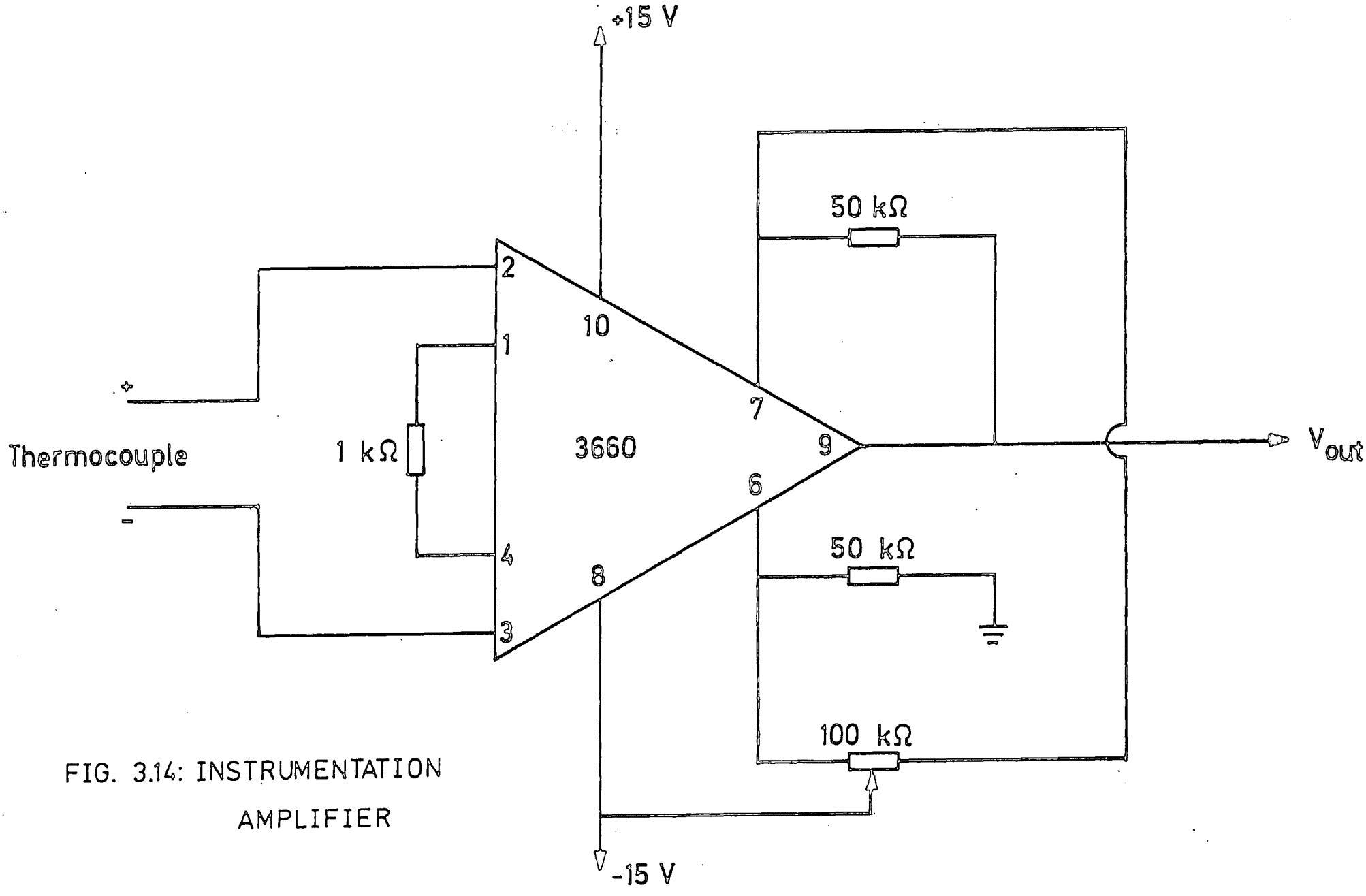


FIG. 3.14: INSTRUMENTATION AMPLIFIER

corrected potentiometer reading (from section 3.1.2) as the true emf and comparing the IA output to this. At reasonably constant temperatures, emf values were recorded using both the potentiometer and the IA. These results are shown graphically in Figure 3.15. The slope of the best straight line gives the gain of the IA. This slope was calculated using the least squares method, and gave a value of 93.7. Any reading on the IA DVM could now be converted to the potentiometer emf value, and thus give the thermocouple temperature from standard tables. A graph of the output from the IA against the calculated corresponding temperature is shown in Figure 3.16. This calibration curve was drawn by simply multiplying the ordinate values of the graphs in Figures 3.4a and 3.4b by the gain of the IA. Ordinate values above -0.2V for the IA output were calculated by direct conversion from standard thermocouple tables, using the IA gain of 93.7. From this graph, IA output values could give temperature indications to an accuracy of $\pm$ 2^{o}C.

3.2 Diode Array Alignment

3.2.1 Initial considerations

In order to ensure the correct location of the diode array in the focal plane of the spectrometer, and that the light falls squarely onto the array of diodes, a method of positioning the array in three orthogonal directions had to be devised and constructed. Before any designing of this mechanism was undertaken, the following considerations were taken into account:

1. The approximate position of the focal area in space was calculated from the engineering drawing of the coudé focus (see section 4.1). The accuracy of this drawing could not possibly be anywhere near the accuracy of the 10μm order of magnitude measurements required by the

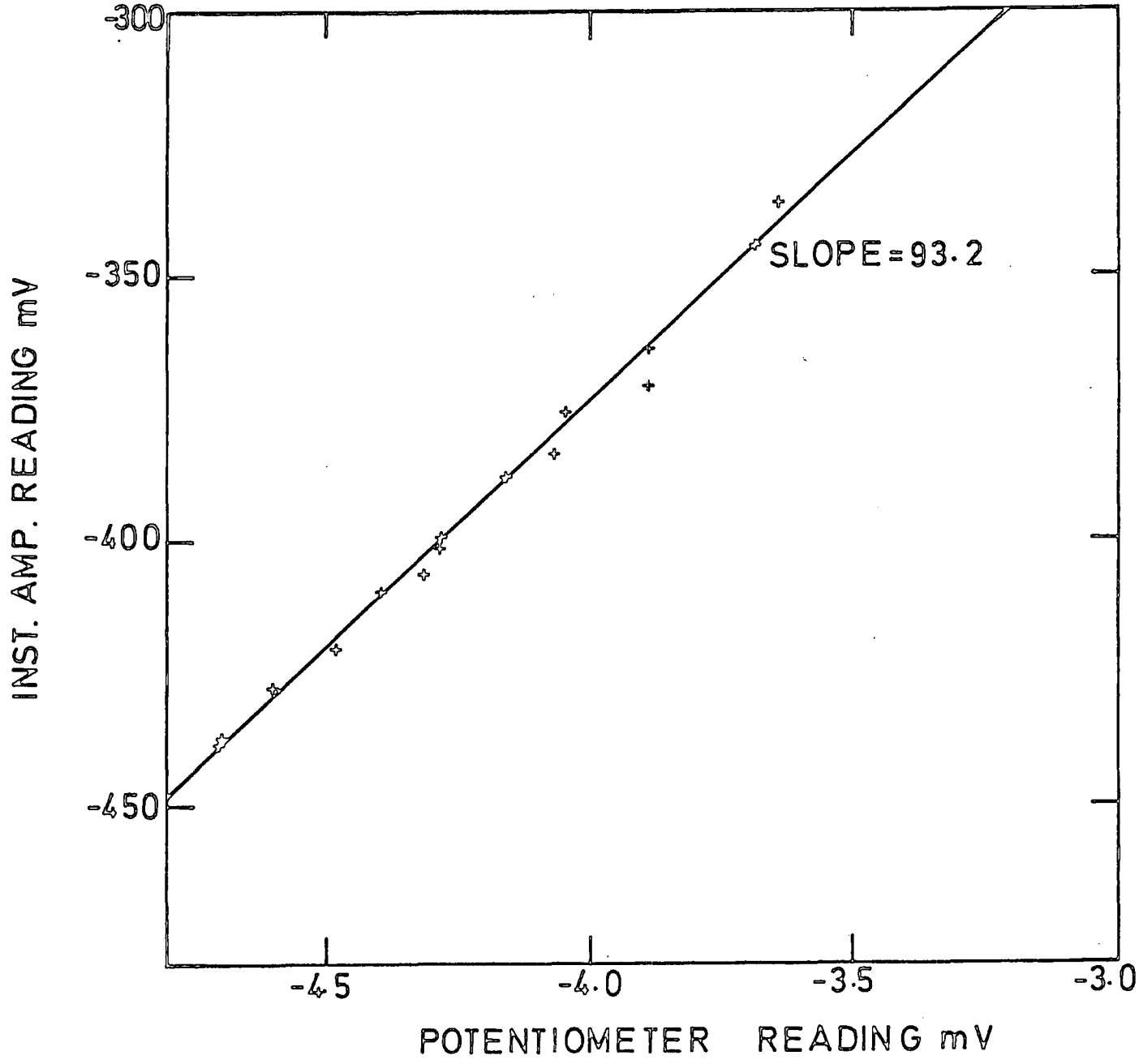


FIG. 3.15:
GAIN DETERMINATION OF INSTRUMENTATION AMPLIFIER

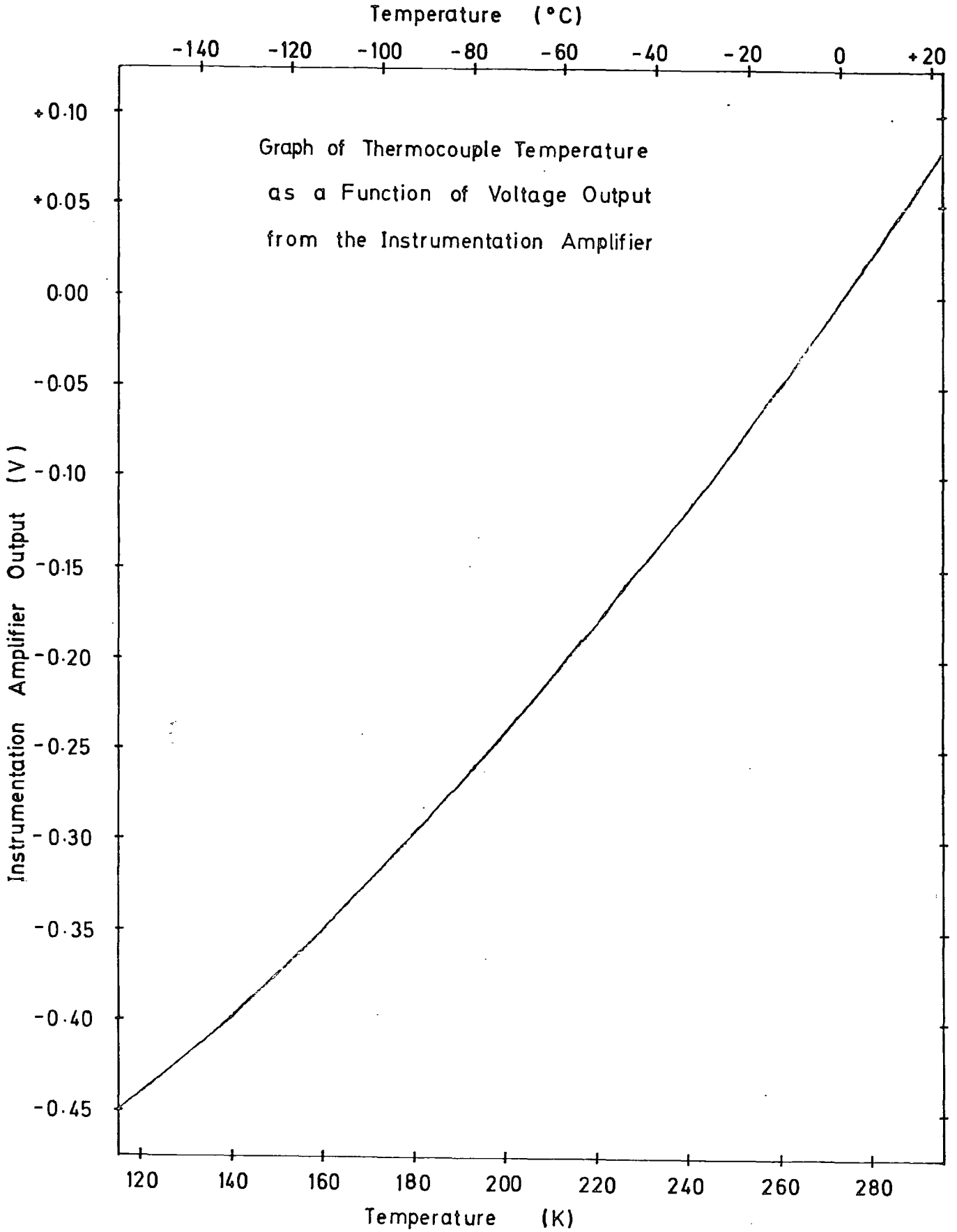

Temperature (°C)
-140
-120
-100
-80
-60
-40
-20
0
+20
+0·10
+0·05
0·00
-0·05
-0·10
-0·15
-0·20
-0·25
-0·30
-0·35
-0·40
-0·45
Instrumentation Amplifier Output (V)
Graph of Thermocouple Temperature
as a Function of Voltage Output
from the Instrumentation Amplifier
120
140
160
180
200
220
240
260
280
Temperature (K)

FIG. 3.16

40μm by 200μm diodes. For an extreme margin of error, it was therefore decided to allow the array to be positioned anywhere within ±2cm of the position of the focal area calculated from the engineering drawing.

2. A trolley was available from the RGO for use at the 30 inch coudé focus. The base of the trolley had bearings on three points of the bottom face. These bearings rested on the two parallel rails straddling the coudé focus (see section 4.1.3). The trolley gave a portable frame of reference on which to construct the diode array alignment mechanism in the laboratory before moving the experiment to the telescope.

3. Movement of the diode array chip inside the cryostat was not a practical consideration. It was decided to subject the whole cryostat to movement in the three planes of freedom.

4. The depth of field of the coudé focus was approximately 1cm. The trolley (with the cryostat) could be maneouvered manually to position the plane of the diode array into the field of focus. A small clamp was attached to one of the support rails, and a screw and dial were mounted on this clamp. Turning the screw against the trolley gave a fine control for movement of the array in the horizontal plane, parallel to the emerging spectrum (defined here as the z-direction). This method was considered sufficient to control movement in the z-direction.

5. The principle material used in the construction of the alignment mechanism was to be brass. This was decided upon because of the resistance of brass to corrosion, its stability (minimal creep), and its hardness (for bearing surfaces). Although it would make the alignment mechanism quite heavy and cumbersome, it is quite easily machined. Thus, brass has the desirable qualities of both stainless steel and aluminium, with few of the disadvantages.

3.2.2 X-direction movement

Movement of the cryostat in the x-direction is defined here as being movement in the horizontal plane, perpendicular to the emerging beam of light from the spectrometer. To achieve this motion, a tongue and groove sliding plate system is used (see Figure 3.17). Two sliding plates are required (labeled A in the diagram), one to fit over each end of the cryostat body. Four rods, R, are situated between each of the four corners of the sliding A plates. These rods are necessary to maintain the required rigidity between the sliding plates holding the cryostat. The rotational orientation of the cryostat with respect to the sliding plates is maintained by the two nylon screws, N, holding the cryostat in place. One pair of opposing edges of each of the A plates is machined to a 45 degree tongue. These tongues fit into mating grooves in two horizontally fixed plates, B, only one of which is shown in Figure 3.17. These two fixed plates are held in position by two other plates, C, again of which only one is shown here. A micrometer screw, M, is mounted to pass through the centre of one of the C plates (see Figure 3.18). When the cryostat and the A plates (not shown here) are in place in the B plates, the micrometer screw is used to move the cryostat (and A plates) horizontally. Two compression springs are located on the opposite side of the cryostat from the micrometer screw. These springs ensure that the cryostat is held firmly against the micrometer screw, removing any backlash in the screw thread.

3.2.3 Y- direction movement

Movement of the cryostat in the vertical plane, perpendicular to the light from the spectrometer is defined here as movement in the y-direction. In similar fashion to the x-direction movement, a tongue and groove

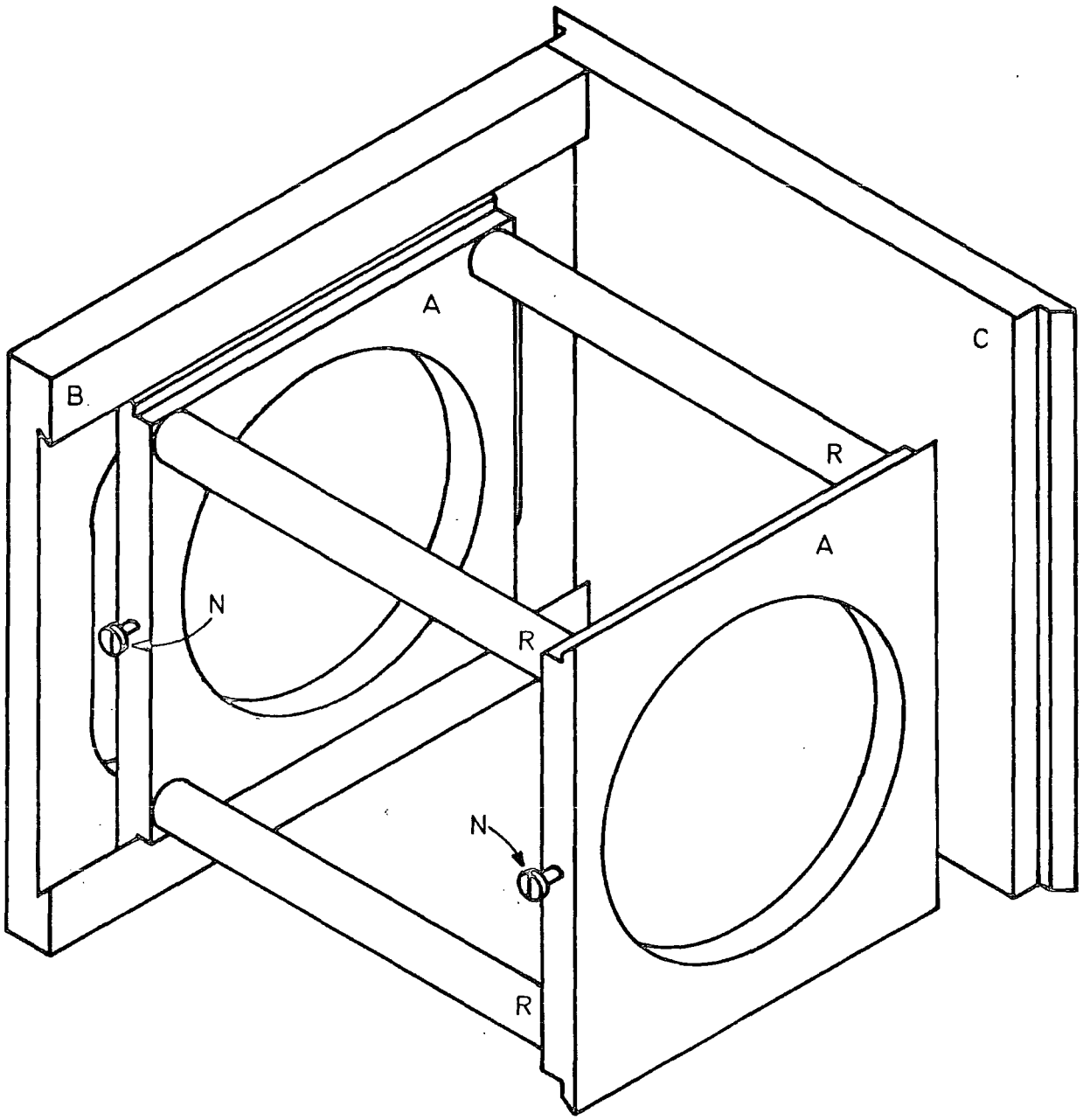


FIG. 3.17 : X-DIRECTION MOVEMENT MECHANISM

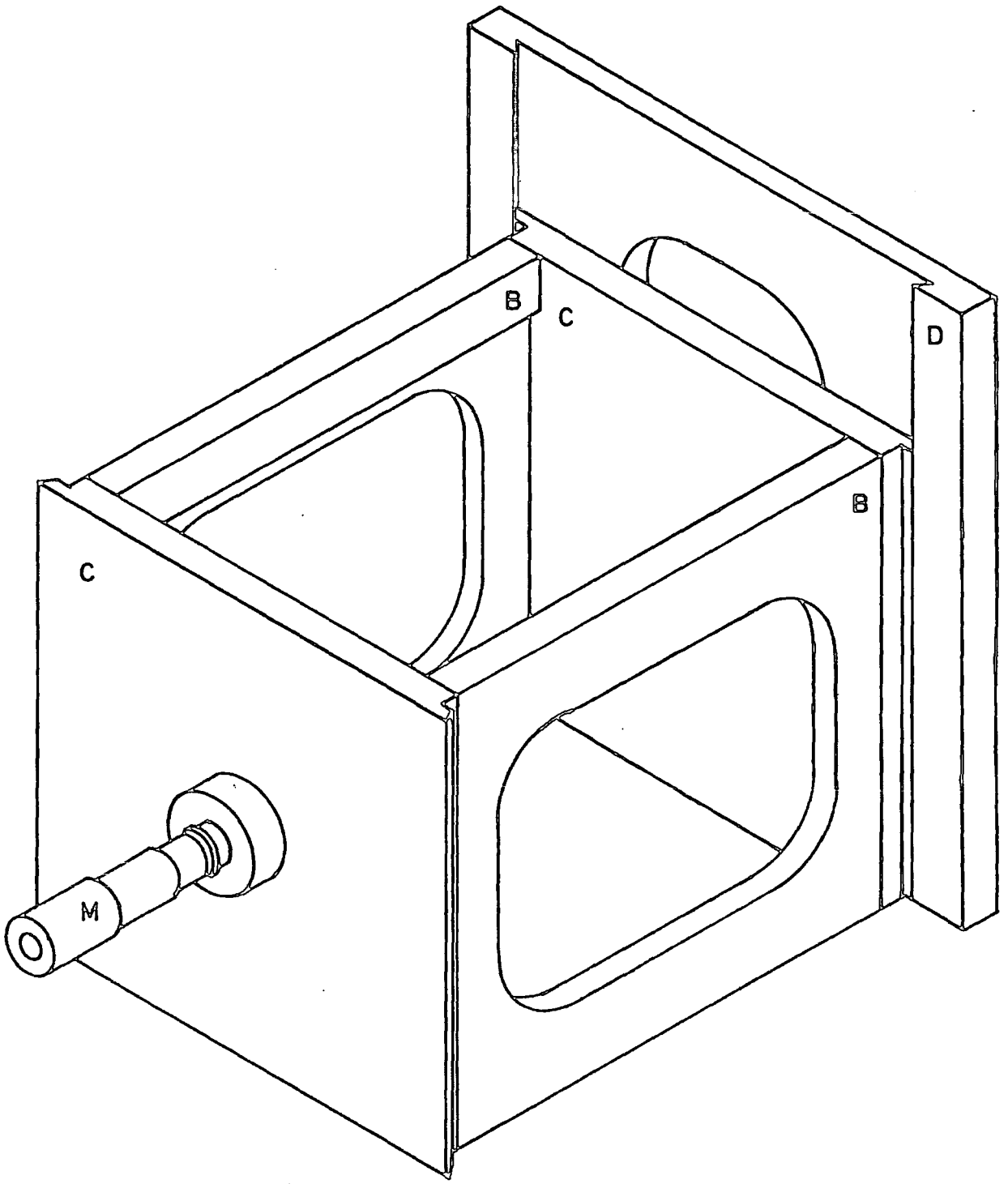


FIG. 3.18: X-DIRECTION AND PART OF Y-DIRECTION MOVEMENTS

sliding plate system was utilized here as well. The tongued plates are plates C in Figures 3.17 and 3.18. The "box" comprising plates B and C moves vertically along grooves in plates D. Only one plate D is depicted in Figure 3.18. Rigidity is achieved by firmly fixing the bottom ends of plates D to the spectrometer trolley on L-shaped mounting blocks, L (see Figure 3.19). The cryostat can undergo vertical movement by means of a steel screw and two brass plates (S, E, and F, respectively). Figure 3.20 shows the complete movement without the horizontal slides or cryostat. A steel screw is used in preference to brass to avoid local "welding" of the threaded assembly due to the close contact of identical materials. Plate E is fixed to the top of the B-C plate "box". Plate F is fixed to the top ends of plates D, and keeps these plates parallel. The hole in the centre of F is tapped to take screw S (see section 3.2.4 for a description of the threaded assembly). The lower end of screw S is held onto plate E by means of a split brass collar, G. Clearance between G and the bottom of S is just sufficient to form a bearing surface. Turning screw S by means of a knurled knob on the top end provides the desired vertical movement of the cryostat. The photographs of Figures 3.20a,b and c show three views of the cryostat mounted in the finished assembly.

3.2.4 Screw assembly

The design of the screw assembly for lifting the cryostat in the y-direction was carefully considered. The threads should be strong enough to lift the cryostat (plus horizontal movement components) without shearing, and fine enough to permit accurate positioning of the cryostat containing the array.

Since the diode-to-diode spacing is 50 μm, an equivalent amount of controlled movement must be achieved by the screw. A rotation of the screw of between 10^{o} and 20^{o} was considered sufficient for the desired amount of control. Therefore, an approximate value for the pitch of

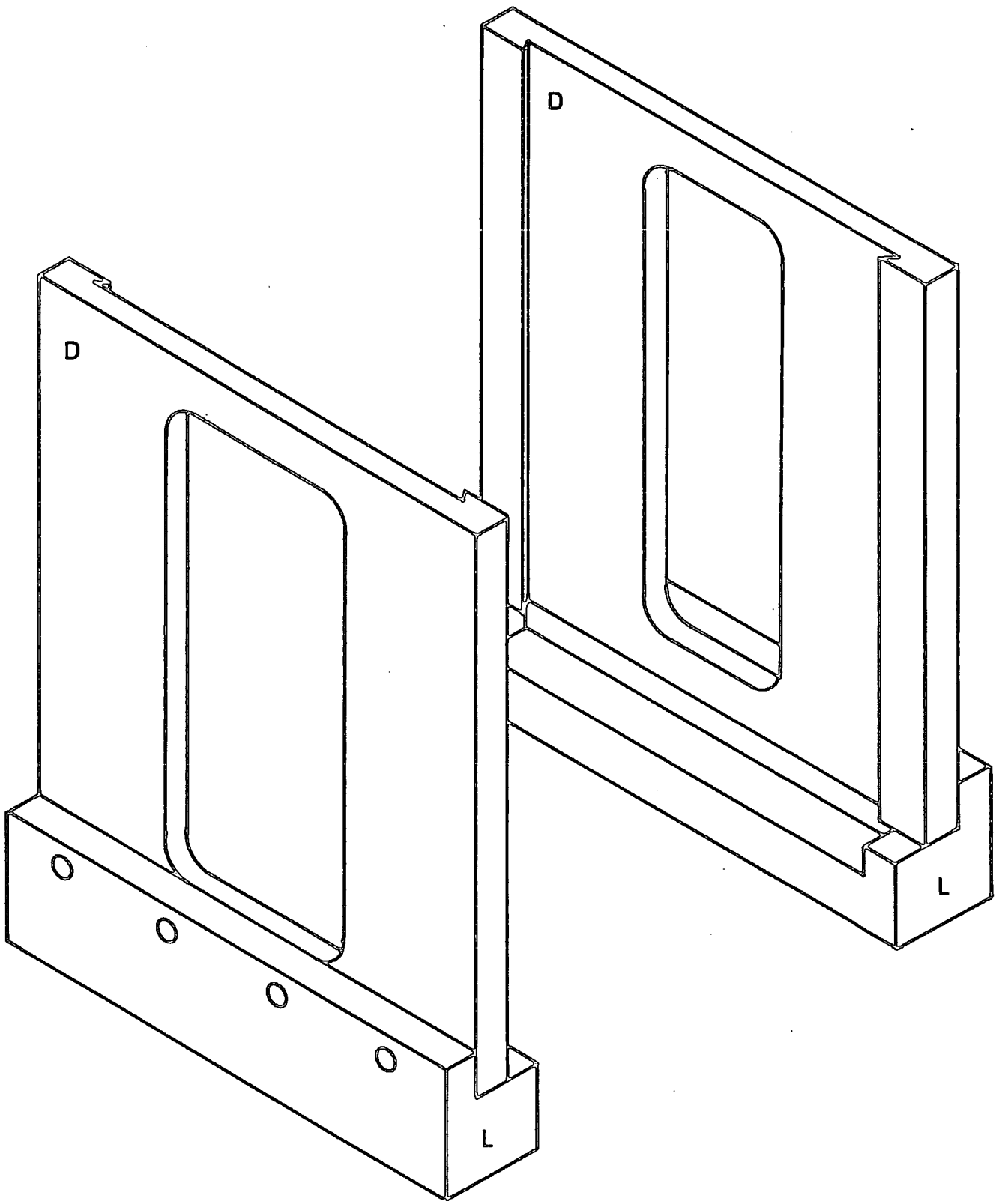


FIG. 3.19: MOUNTS FOR FIXED PLATES OF Y-DIRECTION MOVEMENT

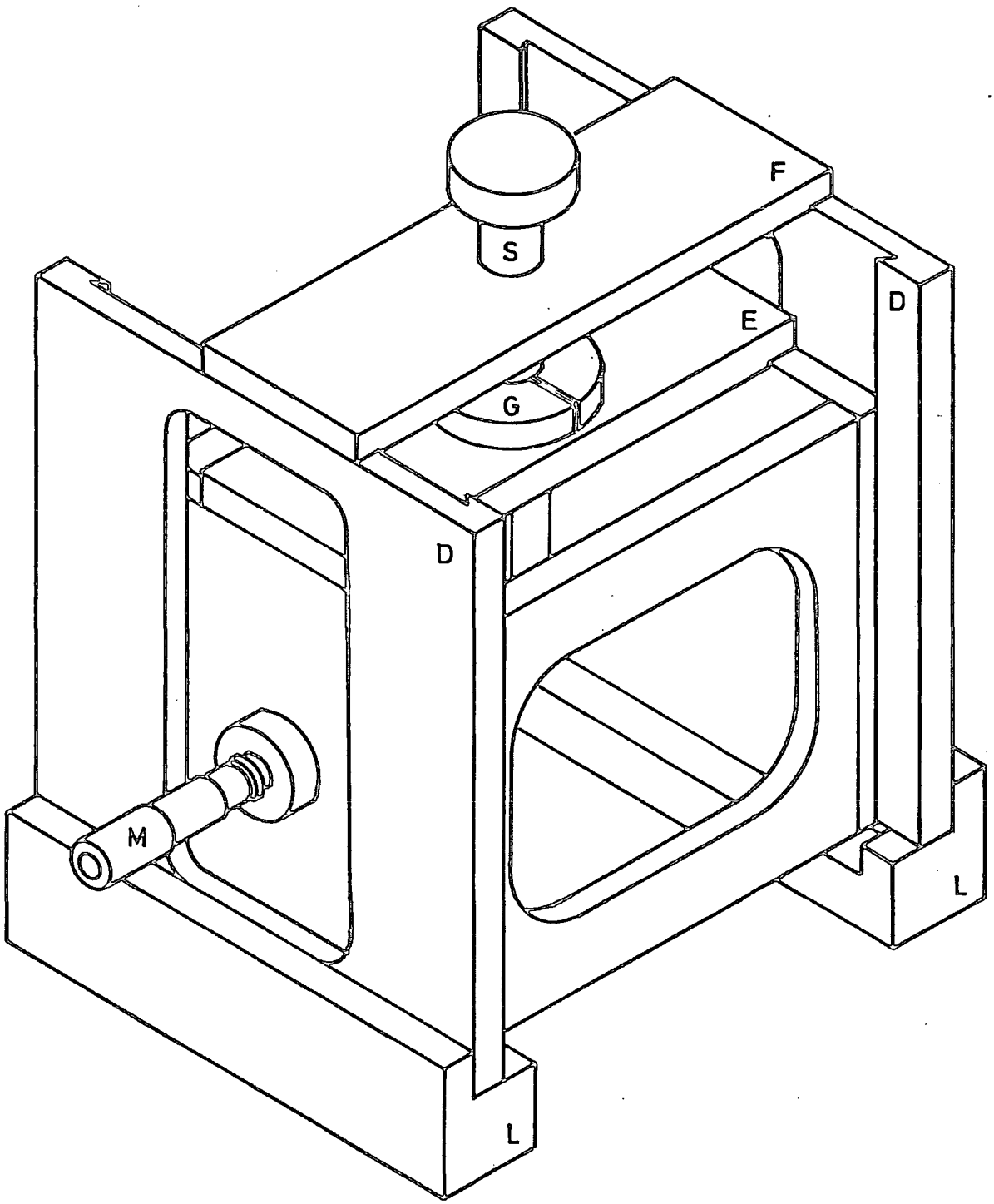


FIG.3.20: COMPLETE XY-MOVEMENT WITHOUT HORIZONTAL SLIDES

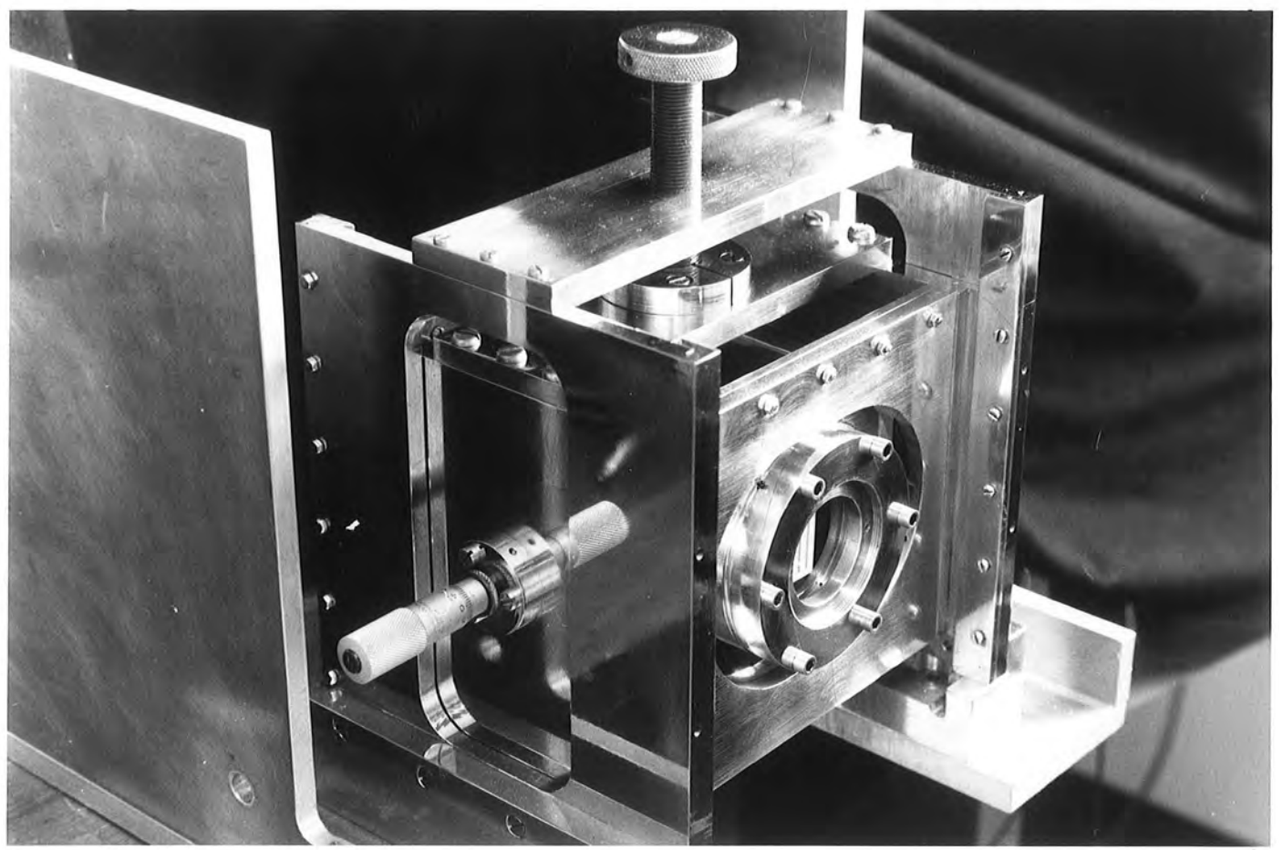


FIG. 3·20a: XY-MOVEMENT WITH CRYOSTAT MOUNTED ON TROLLEY

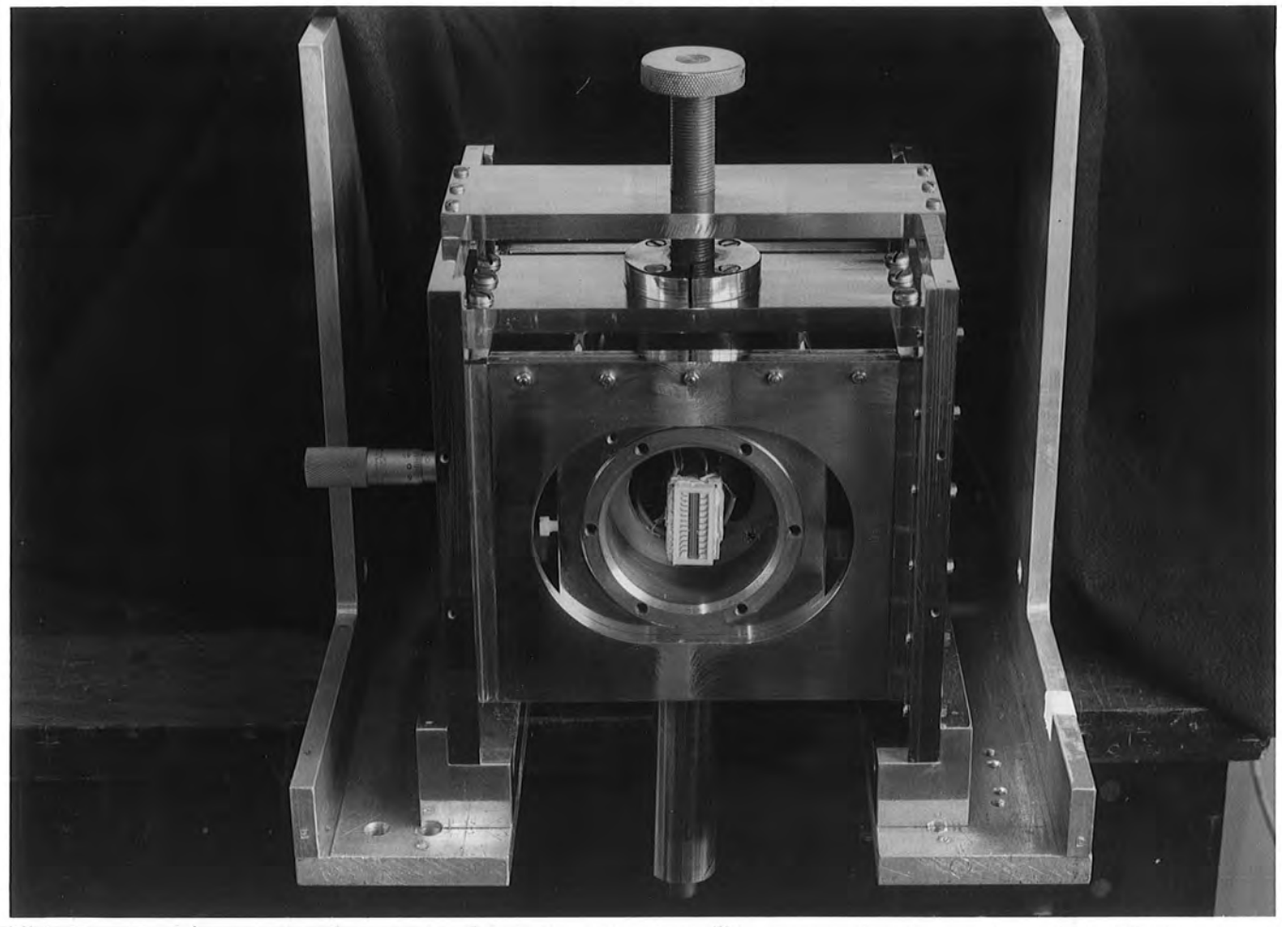


FIG.3·20b: FRONT PLATE REMOVED SHOWING ARRAY CHIP

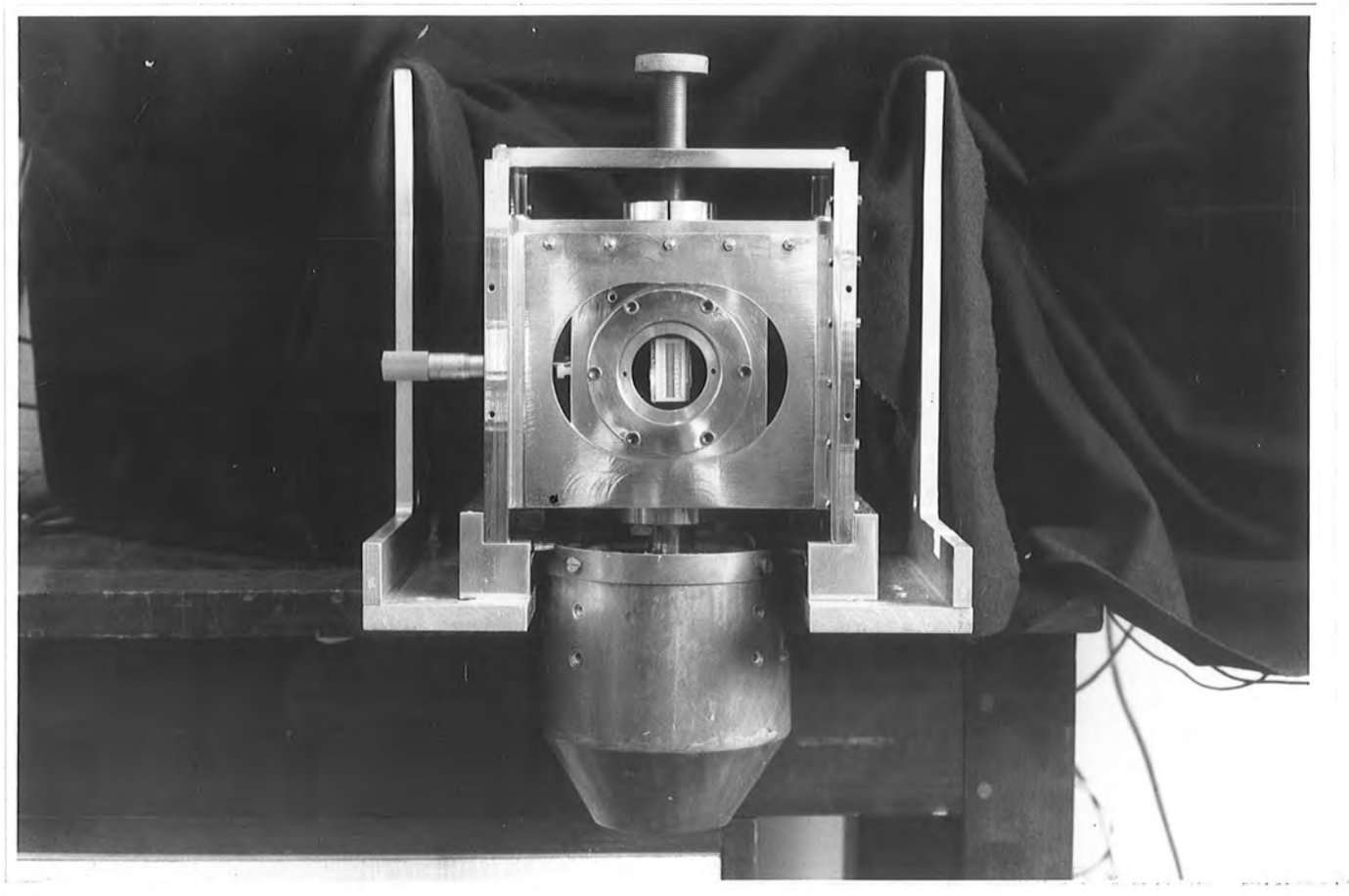


FIG. 3·20c: ENTIRE ASSEMBLY WITH LIQUID NITROGEN DEWAR

the screw was sought, being in the region of

$$\frac{15}{360} \times 50 \times 10^{-6} = 833.33 \text{ turns per meter}$$

$$\approx 20 \text{ turns per inch (tpi).}$$

The next tasks were to calculate both the required length of engagement of the threaded assembly, and the load necessary to break the screw without stripping the threads.

In the event of failure of the threaded assembly, it was preferred that the screw itself should break rather than the threads strip. Therefore, the length of engagement of the mating threads must be enough to support the full load required to break the screw without stripping the threads. To prevent the external thread stripping, the minimum length of engagement is given by (Oberg and Jones, 1962):

$$L_e = \frac{2 \cdot A_t}{\pi \cdot Kn_{max} \cdot \left[\frac{1}{2} + 0.57735n \cdot (Es_{min} - Kn_{max})\right]} \qquad (3.1)$$

where L_e = length of engagement in inches

n = number of turns per inch

Kn_{max} = maximum minor diameter of the internal thread

Es_{min} = minimum pitch diameter of the external thread

A_t = tensile stress area of the screw, from

$$A_t = 0.7854 \cdot (D - \frac{0.9743}{n})^2 \qquad (3.2)$$

where D = basic major diameter of the thread.

In equation (3.1), it is assumed that the area of the screw in shear must be twice the tensile stress area. Hence, the value of 2 in the numerator. In fact, this is slightly greater than necessary, so there is a small safety factor involved.

The major and minor diameters above are the largest and smallest

diameters of the straight screw threads, respectively. The pitch (or effective) diameter on a straight thread is the diameter of an imaginary cylinder, the surface of which would pass through the thread profiles at such points as to make the width of the grooves equal to one-half of the basic pitch. On a perfect thread, this occurs when the thread width equals the groove width.

Assuming that there are no shear or torsional forces present, the load required to break the threaded portion of the screw without stripping the threads is given by

$$P = S.A_t \qquad (3.3)$$

where P = load required to break the screw, in pounds

S = ultimate tensile strength of the bolt material, in pounds $\cdot$ (inch)$^{-2}$

A_t = tensile stress area, in (inches)^2, from equation (3.2).

Following investigations regarding the availability of various taps and dies with the value of the pitch sought (approximately 20 tpi), it was decided to use a $^3/_8$ - inch British Standard Pipe (BSP) thread. The characteristics of this thread are as follows:

$^3/_8$ - inch BSP	
pitch	19 tpi
major diameter	0.656 inches
effective diameter	0.6223 inches
minor diameter	0.5886 inches

For a screw elevation of 50μm, a thread of pitch 19 tpi requires a rotation of

$$\frac{50 \times 10^{-6} \times 19}{2.54 \times 10^{-2}} = 0.0374 \text{ turns}$$

or $0.0374 \times 360 = 13.46$ degrees.

Using equation (3.2), the tensile stress area is calculated to be

$$A_t = 0.7854 \left(0.656 - \frac{0.9743}{19}\right)^2$$
$$= 0.2872 \text{ in}^2.$$

Substituting this value of A_t into equation (3.3) gives the maximum load that the screw will sustain as being

$$P = (10^4 - 10^5)\ 0.2872$$
$$= (1.3 - 13) \text{ tonnes}$$
$$\therefore P \sim 10 \text{ tonnes}.$$

Finally, using the calculated value of A_t in equation (3.1) gives the minimum length of engagement to support this load as

$$L_e = \frac{2 \times 0.2872}{3.1416 \times 0.5886 \times \left[0.5 + 0.57735 \times 19 \times (0.6223 - 0.5886)\right]}$$
$$= 0.3573 \text{ inches}$$
$$= 0.907 \text{ cm}.$$

In summary, for a $^3/_8$-inch BSP threaded assembly, a length of engagement of 0.907 cm will sustain a load of up to ~10 tonnes without stripping the threads. Furthermore, a twist of approximately 13°.5 will change the height of the screw by one unit of diode spacing. The actual length of engagement of the threaded assembly in the plate F and screw S constructed is 1.0 cm.

3.3 Experimental Operation and Control

3.3.1 General layout

The centre of control for the collection of data from the stellar observations is a PDP11-03 mini-computer. The programming language is a form of BASIC called CATY 2 (see section 4.3.1). All peripheral interfacing is achieved via CAMAC (Hedge et al, 1978). A 512 word buffer memory is a Durham produced CAMAC module, as are the integration and readout control modules. The remaining electronics are constructed in NIM modules. A block diagram of the acquisition and control logic is shown in Figure 3.21.

3.3.2 Signal processing

The diode array chip requires five input signals for operation, namely LSR, CP1, CP2, Ø1 and Ø2. LSR is the Load Shift Register pulse. This pulse determines the beginning of each frame, and is used to start a scan. CP1 and CP2 are the Charge Pulses, used to recharge the diode signal levels to their fully charged state. CP1 recharges the odd-numbered diodes, and CP2 recharges the even-numbered diodes. Ø1 and Ø2 are two clock pulses in a non-overlapping two-phase configuration (see Figure 3.22).

A basic two-phase clock is shown in (a). In the ideal case, the sum (Ø1 + Ø2) will be a continuous output equal to the voltage of the clock pulses. In practice, the voltage switch from high to low (and vice versa) takes a finite amount of time. Therefore, at the time of the switching of the clocks, the sum (Ø1 + Ø2) can be anywhere between zero volts and twice the clock voltage (b). This may cause problems further along the logic circuit by allowing more than one supposedly mutually

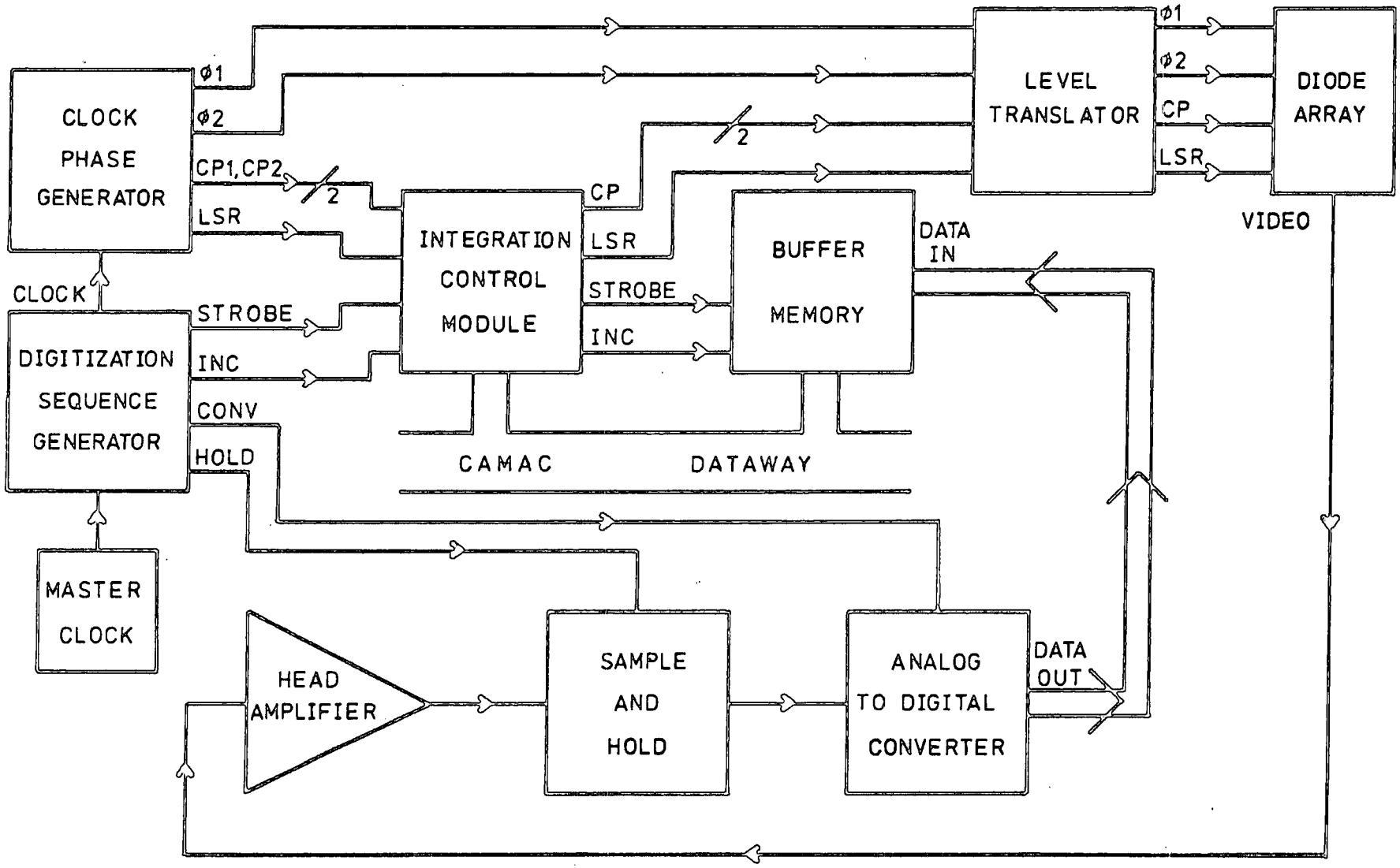


FIG. 3.21: DATA ACQUISITION AND CONTROL SYSTEM

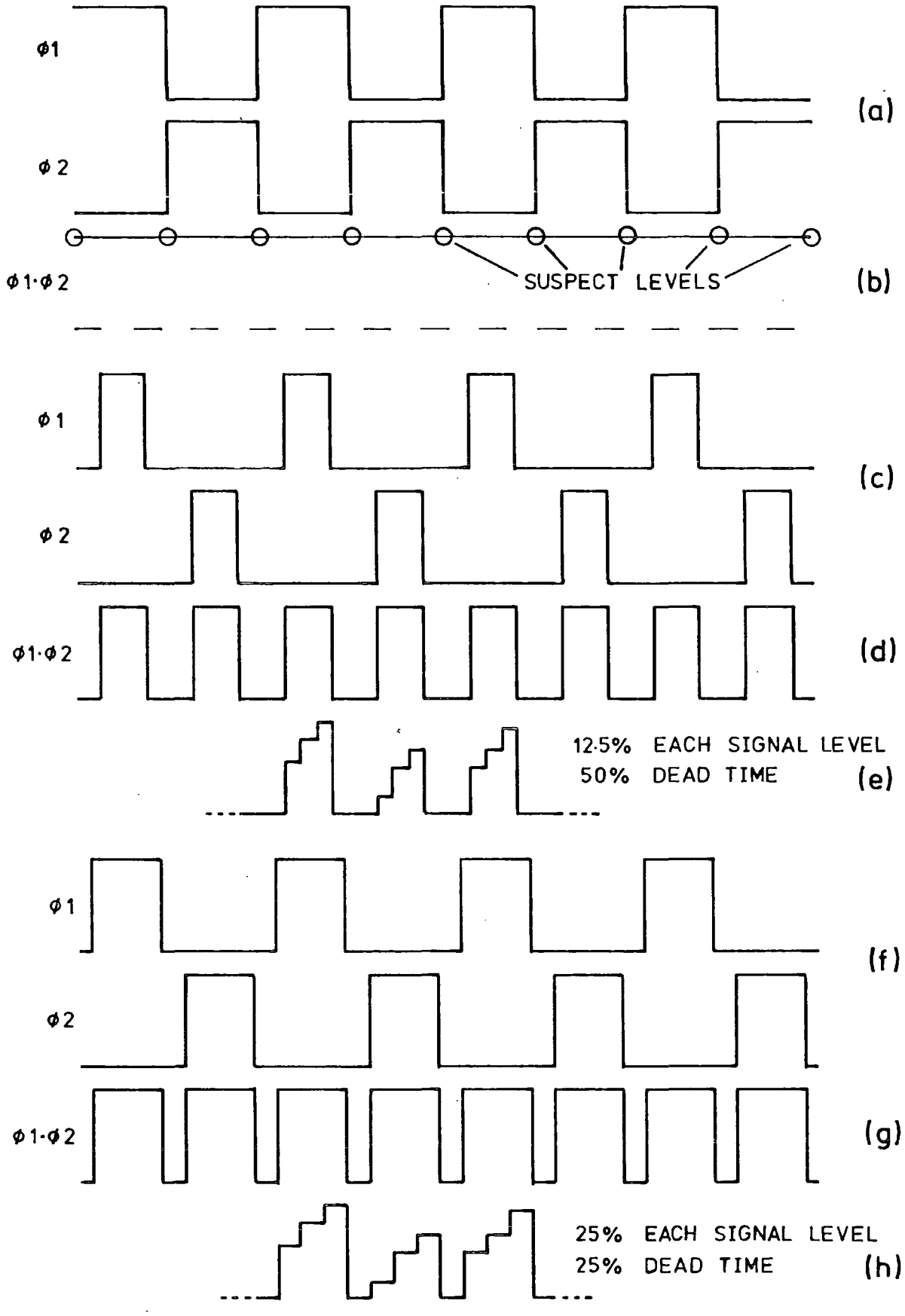


FIG. 3.22: DEVELOPMENT OF OPTIMUM CLOCK TIMING

exclusive events to occur simultaneously.

The first attempt at solving this problem was to use a 50% two-phase clock, which ensured that the pulse edges did not coincide (c). This meant that, in effect, 50% of the time was being wasted (d). Since the video output for one pair of diodes is digitized at four points (base line, diode sum, one diode recharged, both diodes recharged), only 12.5% of each clock phase is available for each signal level (e).

Finally, it was decided to use a mark to space ratio of 3:5 for each clock (f). This results in a 3:1 mark to space ratio in the sum $(\emptyset 1 + \emptyset 2)$ (g). The 3:1 ratio allows the digitization to be carried out at four equidistant points in each clock phase. In this way, each of the four levels can be digitized during 25% of the clock phase (h).

Before digitization of the video signal is achieved, it must be converted from a current signal to a voltage signal. This conversion is carried out by a head amplifier, mounted close to the cryostat, consisting of a Burr Brown 3550 operational amplifier (see Figure 3.23). The gain of this amplifier is altered by different values of the feedback resistor, R_F, in the circuit. The output signal is that voltage that would exist across R_F if this were the only circuit component. Using a resistance of 47kΩ for the feedback resistor gives an amplifier output ranging from 0V to +8V, which is suitable for the processing electronics.

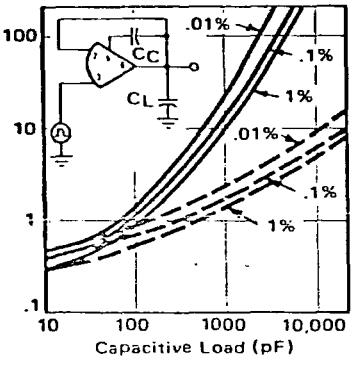
From the head amplifier, the signal is fed into a sample and hold unit employing a Hybrid Systems S/H 730. The signal is sampled at three points in a clock phase, each sampling occurring after any ringing has subsided. Not sampling during the "video off" quarter of the cycle has the effect of suppressing the unwanted pedestals of the ADC base line. The time duration of the sample is then stretched to the same length as the original input signal, thus removing any large spikes from the video signal.

The signal is then input to the analog to digital converter (ADC).

TYPICAL PERFORMANCE CURVES

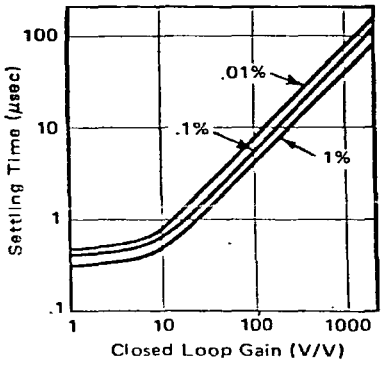
$T_A = 25^oC$ $V_S = \pm 15$ Vdc unless otherwise indicated.

SETTLING TIME vs. CAPACITIVE LOAD

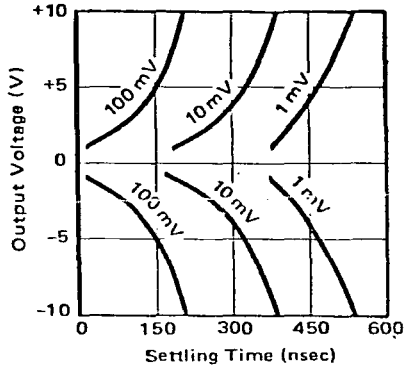


$C_C = 0$ — — — $C_C = .02 \times C_L$

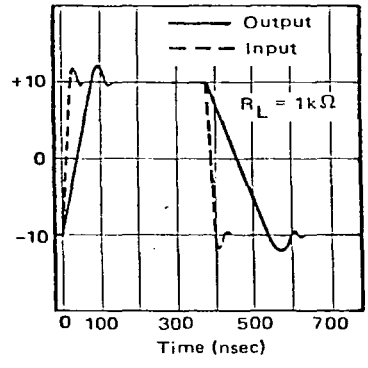
SETTLING TIME vs. CLOSED LOOP GAIN



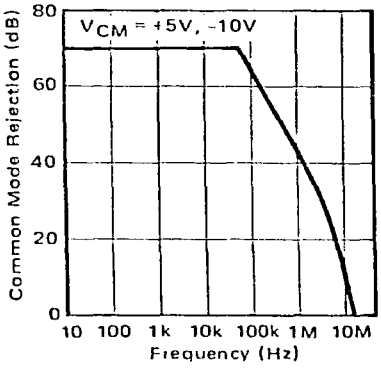
SETTLING TIME vs. OUTPUT VOLTAGE CHANGE



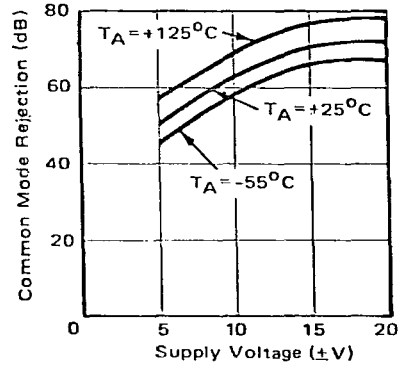
LARGE SIGNAL VOLTAGE FOLLOWER PULSE RESPONSE



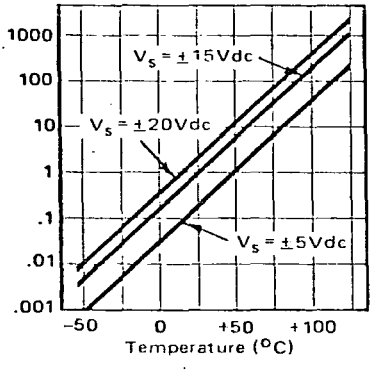
COMMON MODE REJECTION vs. FREQUENCY



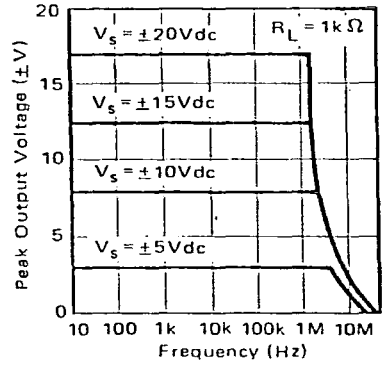
COMMON MODE REJECTION vs. SUPPLY VOLTAGE



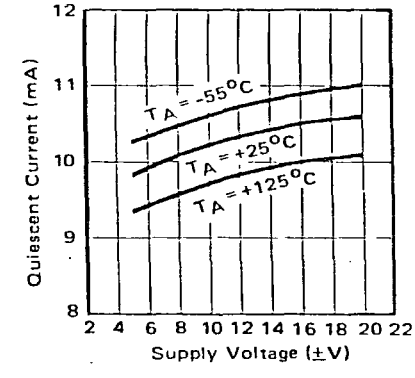
NORMALIZED INPUT BIAS CURRENT vs. TEMPERATURE



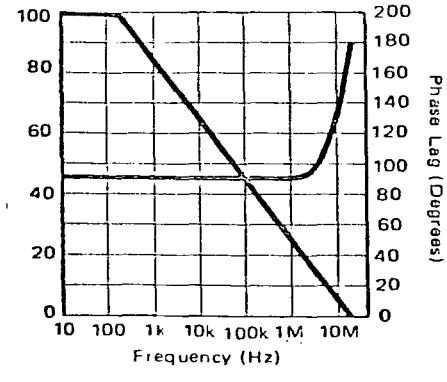
OUTPUT VOLTAGE vs. FREQUENCY



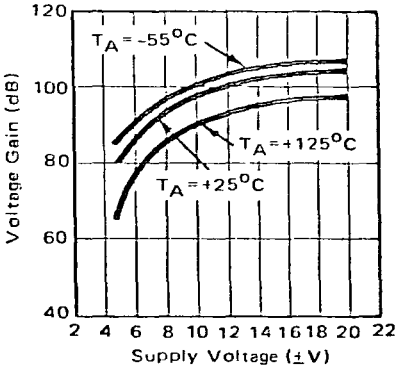
QUIESCENT CURRENT vs. SUPPLY VOLTAGE



OPEN LOOP RESPONSE



OPEN LOOP GAIN vs. SUPPLY VOLTAGE



MAXIMUM POWER DISSIPATION

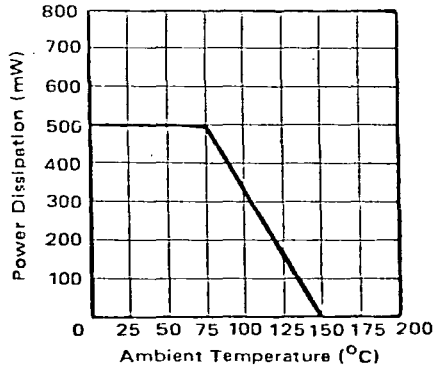


FIG. 3·23: PERFORMANCE CURVES OF BURR BROWN 3550 OPERATIONAL AMPLIFIER

The ADC used is a Burr Brown ADC80AG-12, which digitizes the signal by successive approximation to twelve bit accuracy (ie. digitization steps of 1/4096 times the available voltage range of the ADC). The ADC operates in the -2.5V to +2.5V range, thereby giving digitization steps of 1.22 mV.

From the ADC, the digital signal is read into the sequential access memory. Since each diode pair signal is digitized in four places, 4 x 128 twelve bit words of data are generated. These data are stored here in the buffer memory until the readout cycle is complete. Under program control, the data can then be transferred to the resident memory of the computer for display or output.

3.3.3 Control logic

The digitization sequence generator (DSG) generates four signals, ie. STROBE, INC, CONV, and HOLD. The HOLD signal is used to enable the sample and hold module, thereby removing spikes from the video signal. CONV enables the ADC, thus converting the signal from analog to digital form. INC is used to increment the current memory word address for the input and output of information in the buffer memory. The STROBE pulse determines when in the clock cycle the information is to be read to or from the buffer memory. The timing of these four signals from the DSG is set by two hexadecimal thumbwheel switches on the front panel of the module. The first switch determines the start time (in the clock cycle) of any one of the four signals, and the second switch determines the time duration of that signal. There are two push switches for each signal on the front panel as well, one to load the set timing and the other to inspect. Two hexadecimal LED displays indicate the recorded values. The four signals can therefore be set to occur anywhere in the clock cycle.

The STROBE and INC (as well as the CP and LSR) pulses are input to

the integration control module (ICM), which controls the integration and readout sequences. By choosing the appropriate control word for the ICM via the computer program, one of the following sequences will be initiated:

(a) integrate - commences a scan and allows the diodes to discharge due to light and/or thermal leakage;

(b) recharge - charges the diodes back to their original fully charged state by sending a charge pulse 25 times;

(c) non-destructive readout - outputs the signals from pairs of diodes to the buffer memory without re-charging the diodes;

(d) double sample readout - outputs the signals on individual diodes by first recharging one diode and then the other diode in the pairs; the output is sent to the buffer memory.

Two manual requests are also available from the ICM:- non-destructive readout and terminate run. Either of these two can be selected at any time during a run by means of two push-button switches on the front panel of the ICM. When it is required for the data to be stored in the buffer memory, a 9 bit counter assigns each word a memory location. The counter is initially set to zero and is incremented (by the INC pulse) after each word has been strobed in (STROBE) by the ICM.

The clock phase generator (CPG) generates all four of the signals required by the array, itself requiring only a clock input from the DSG. The master clock, which is the fundamental time base for the whole experiment, is input to the DSG. The level translator (LT) translates the TTL input signals it receives into the MOS levels necessary for the operation of the diode array chip.

3.3.4 Peripheral devices

The general layout of the experimental package is shown in Figure 3.24, including all of the computer peripheral devices. Since the PDP 11-03

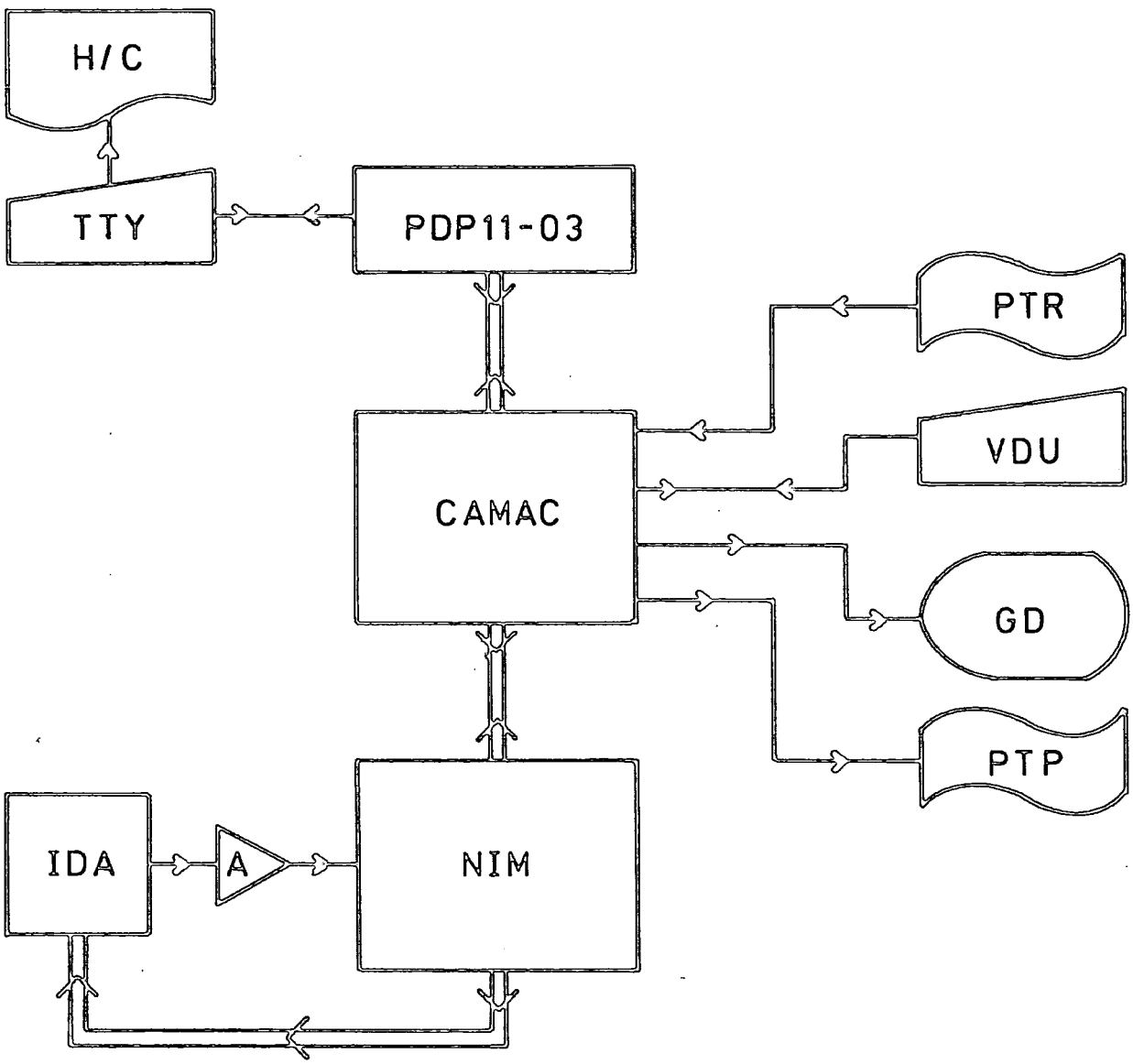


FIG. 3.24: COMPUTER AND PERIPHERALS

has no operators panel, initial communication is achieved via an ASR 33 teletype (TTY). Computer control can be transferred to the Lynwood DAD-1 visual display unit (VDU). The TTY can still be used for hard copy printouts of data, as well as retaining the ability to break communication between the computer and all other peripherals, and restore control to the operator. Program input to the computer is via a Lynwood ATR 2 paper tape reader (PTR), and paper tape output is on a Facit 4070 punch (PTP). Visual output (ie. non-destructive readout, double sample readout, etc.) is displayed on a Tektronix 603 graphics display (GD). All four of these peripheral devices (VDU, PTR, PTP, GD) are interfaced via CAMAC to the PDP 11-03. The TTY is on a direct link with the computer and requires no separate interface. The electronics associated directly with the operation of the diode array are contained in the NIM bin, while the actual control of the experiment during a run is maintained via CAMAC.

The computer to CAMAC interface is a Hytec CAMAC Programmed Dataway Controller type 1100. It occupies both the control station and the adjacent normal station of the crate, and acts in two ways. Firstly, it is the interface which directly adapts the Dataway to the PDP 11-03 computer. Secondly, it converts the crate into a System Crate, behaving as the equivalent of an Executive Controller. In this capacity, it controls the transfer of commands and data between the interface and the computer. The advantage of this special controller is evident if only one crate is used, where cost can be kept down and simplicity maintained. If a second crate is used, another Hytec 1100 Controller is required (ie. one special controller for each crate).

REFERENCES

Hedge, A.R., Breare, J.M., Campbell, A.W., Hopkinson, G.R., Humrich, A.: 1978, ESO/SRC Conference on "Applications of CAMAC to Astronomy", 210.

Kaye, G.W.C., Laby, T.N.: 1960, "Tables of Physical and Chemical Constants", 12th ed., Longmans, London.

Oberg, E., Jones, F.D.: 1962, "Machinery's Handbook", 16th ed., Industrial Press, New York.

CHAPTER FOUR

OBSERVATIONAL OPERATION

4.1 Telescope and Spectrometer

The telescope used for the observations of the stellar spectra to be described was the 30-inch reflector situated in dome A of the equational group of the Royal Greenwich Observatory. This telescope was originally one-half of a twin telescope mounted in Greenwich in the late 1800's, the other instrument being a 26-inch refractor. When the observatory was moved from Greenwich to Herstmonceux after World War II, the 30-inch was removed from the mounting and replaced by a counterweight. A new fork mounting was constructed for the reflector. At first, this telescope was used in the Cassegrain form, but later a coudé spectrograph was installed to afford the opportunity of high dispersion spectrophotometry.

4.1.1 Dome A

The telescope and spectrometer are located in dome A of the equational group. Figure 4.1 (from Harding, et al. 1968) shows the complete coudé arrangement. The spectrometer collimator is supported on the ground floor, while the top level (observing room) contains the telescope proper and the slit assembly. The mezzanine level houses the diffraction grating and telescope drive, together with the experimental package and any electronics necessary for the control of the experiment.

The spectrometer extends from inside the top level of the building down through to the ground floor. The collimator mirror is mounted separate from the building on a concrete block at the ground floor level. This arrangement gives the collimator a focal length of 224 inches (5.69m). The main spectrometer framework extends from the mezzanine level upwards

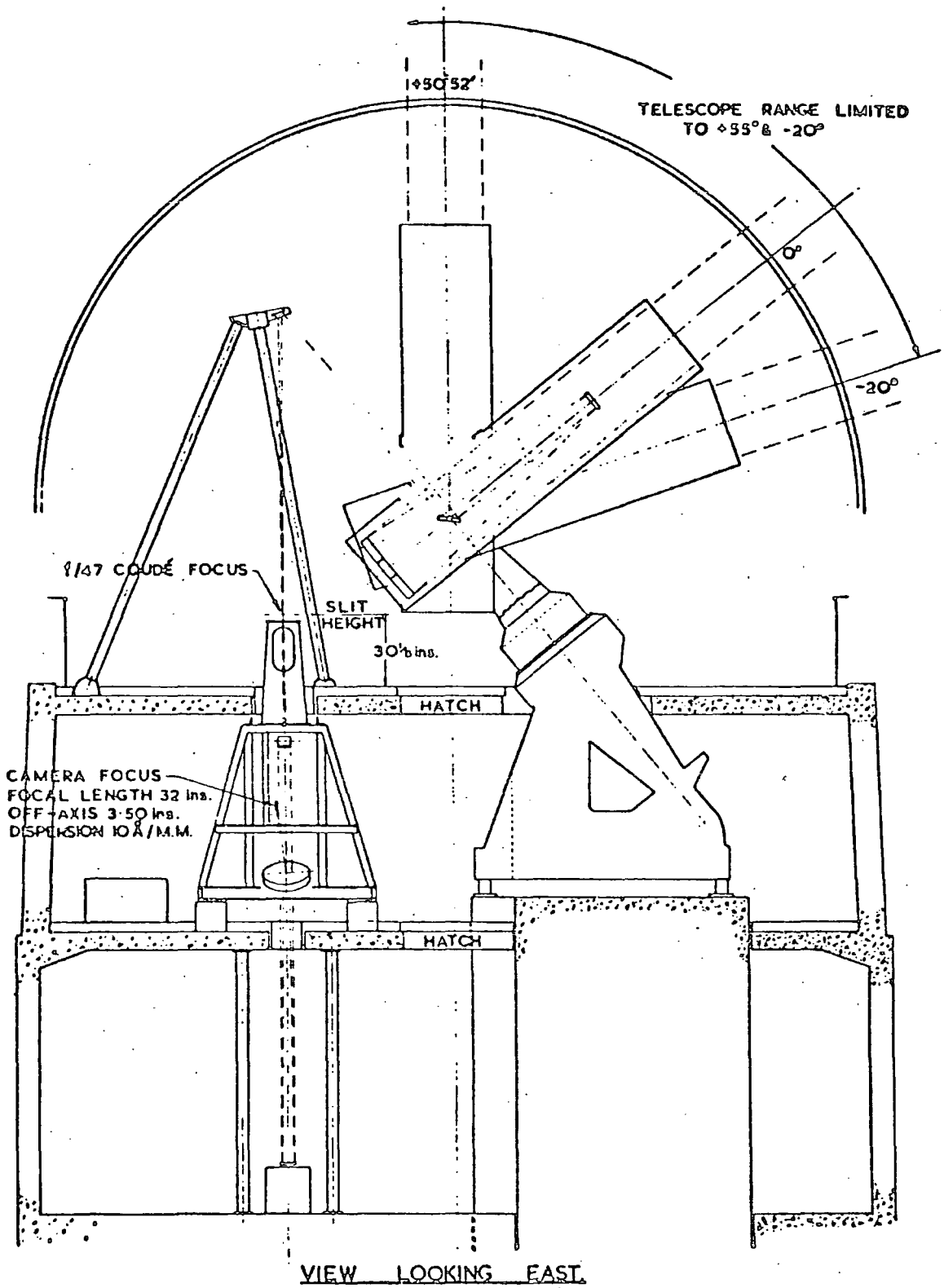


FIG. 4.1: CUTAWAY VIEW OF DOME 'A'

through the floor of the observing room, providing a mounting for the slit assembly. This steel framework also provides the support for the grating and detector components inside the spectrometer room.

The tripodal support for the top mirror in the dome (on the northern extension of the polar axis of the telescope) consists of steel tubing filled with sand. The sand is present to reduce any vibrations in the tubular framework as the three legs are only supported by the floor of the observing room.

4.1.2 Light path

The light from a star enters the telescope and is reflected off the primary mirror onto a hyperbolic mirror. This secondary mirror is designed to produce an f/47 beam of light. The beam is directed onto a flat mirror mounted inside the telescope tube at the intersection of the polar and declination axes. By means of a specially designed mechanism inside the telescope tube, any change in declination of the telescope is accompanied by half of that change in the declination of the flat. This ensures that the beam is always directed upwards along the polar axis through an aperture in the side of the telescope. Another flat is located further up along the polar axis mounted on the tubular steel frame. This mirror reflects the beam vertically downwards onto the slit of the spectrometer. The width of the slit can be varied in steps of 12.5μm. This coudé arrangement restricts movement of the telescope such that only objects with declinations between -20 and +55 degrees can be observed. However, the advantage of this system is that the axis of the spectrometer is made to be vertical.

Above the slit assembly is a retractable flat mirror. This mirror can be used to deflect the beam of light through 90 degrees at a convenient height so that the image of the star can be observed through

an eyepiece. The field of view at this eyepiece is approximately 7 minutes of arc.

After passing through the slit, the light travels down to the collimator mirror at the ground floor level. From here, it is reflected back up onto the diffraction grating.

The diffraction grating is ruled with 831 lines $\cdot$ mm^{-1}, and blazed at 8465Å in the first order. It can be rotated about an axis which is parallel to the rulings and passes through the face of the grating. The rotation is achieved by a drive mechanism controlled from outside the spectrometer dark room. A digital display on the drive control indicates the angle between the normal to the grating and a constant fixed plane. The display is accurate to ± 1 in the least significant figure, which in this case is ± 0.1 degrees. The dispersion of the grating varies according to the wavelength under observation, but is approximately 10Å $\cdot$ mm^{-1} in the first order and 5Å $\cdot$ mm^{-1} in the second order.

4.1.3 Spectrometer room

The spectrometer enclosure (or dark room) is located on the mezzanine level and houses the coudé focus. Figures 4.2 and 4.3 illustrate the path of the beam of light in this room. The diffracted light from the grating is reflected off of a 39-inch focal length mirror mounted on the lower section of the main spectrometer framework. This mirror focuses the spectrum onto a flat which turns the beam through an angle of approximately 60 degrees to enter the detector horizontally. A concrete pier is situated at one end of the room onto which two parallel guide rails are firmly secured. A three-wheeled trolley sits on these rails and the cryostat and the XY-movement are mounted on this trolley. Precise positioning of the cryostat such that the spectrum is focused onto the diode array is described in section 3.2.

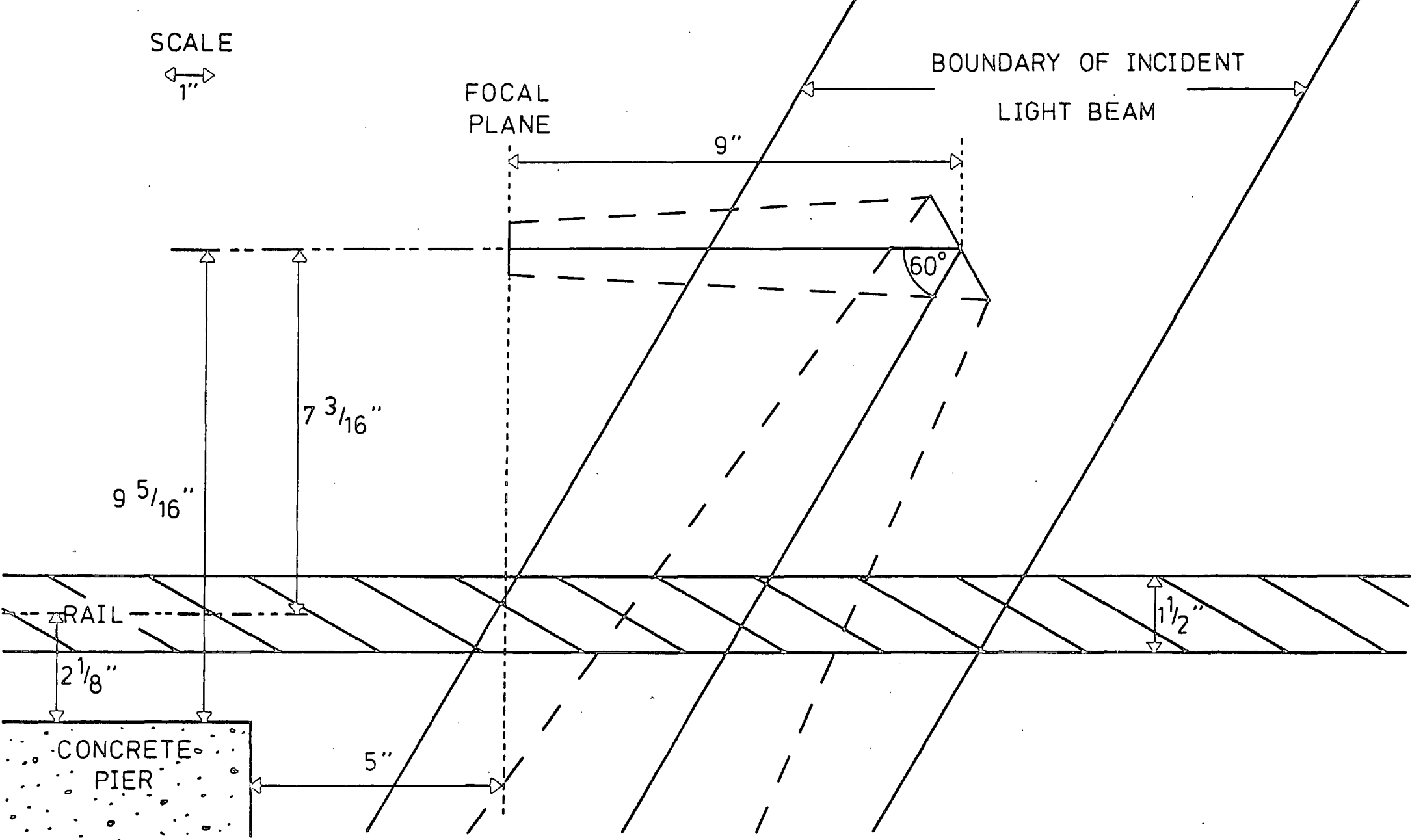


FIG. 4.2: DIMENSIONS AT THE COUDÉ FOCUS

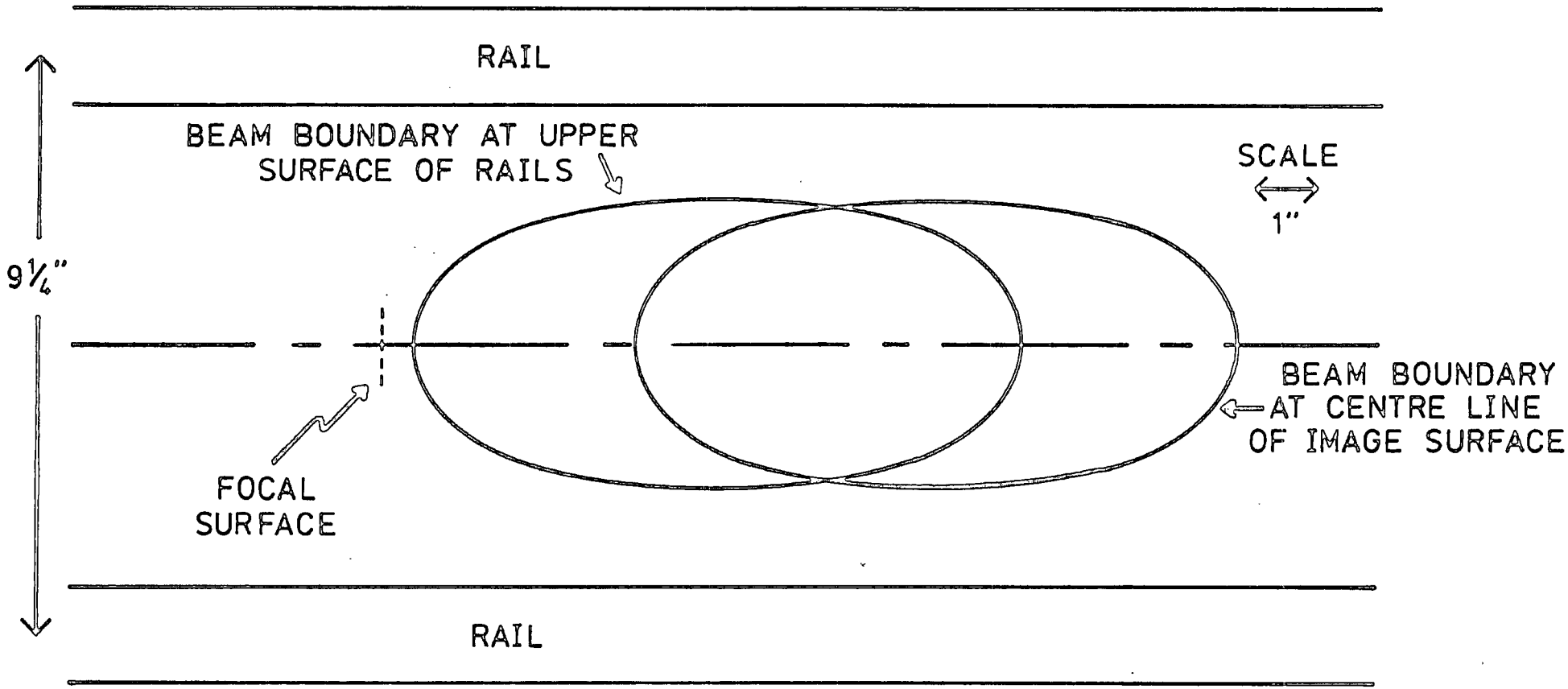


FIG. 4.3: OUTLINES OF DOWNCOMING BEAM SHOWING REGION WHICH SHOULD NOT BE OBSTRUCTED (from C.F.W. HARMER, personal communication)

4.1.4 Star location and tracking

The initial position of the telescope is set manually and can locate a star to an accuracy of 0.2 minutes right ascension and 1 arc minute declination. A finder is attached to the side of the telescope. The field of view in the eyepiece of this finder is approximately 1 degree of arc.

The telescope drive operates at sidereal rate. Manual and autoguide override controls are available. The Grubb Parsons autoguider operates on receiving reflected light from the jaws of the slit. At the slit, the scale of the image is 174μm to 1 arc second. The field of view at the autoguider eyepiece is only about 20 arc seconds, so correct centering is important. The faintest object able to be held on the slit by the autoguider is a ninth magnitude star. Manual override controls can be used instead of the autoguider during variable seeing conditions, or when large amounts of background light are present.

4.1.5 Calibrations

A tungsten lamp is mounted between the top mirror and the slit. It can be swung around into the optical path of the spectrometer for flat field observations. The results obtained from the spectrum of the tungsten lamp indicate the differing responses of individual diodes to white light. Dividing any recorded stellar spectrum by the appropriate tungsten spectrum (for that particular wavelength range) will therefore remove any diode-to-diode responsivity variations.

For calibration purposes, a neon arc lamp is also available. This lamp is mounted off-axis above the slit. The neon light can be directed down the spectrometer by means of a retractable flat mirror. This is the same flat that is used to direct the light of the image of the star

through to the 7 arc minute eyepiece mounted above the slit.

In order to locate a specific region of the spectrum to expose the detector to, some method of calibrating the angle of the diffraction grating in terms of the waveband available for observation had to be devised. Dr. Dianne Harmer (the astronomer in charge of the 30-inch telescope) had previously exposed a number of photographic plates to spectra produced by the neon lamp on the spectrometer. Five exposures were taken on each plate. The setting of the grating angle was altered by 0.5 degrees between each of the five exposures. Seven plates were available for the calibrations.

The wavelength was determined for the central position of each spectral "strip" on the plates by comparison to a reference set of neon spectral lines borrowed from Dr. Harmer. Table 4.1 shows the values of the grating angle and the corresponding central wavelength for each strip on four of the seven plates. These twenty points gave sufficient information to plot a graph of the diffraction grating angle against the centre of the available waveband (see Figure 4.4). The best straight line is drawn through these points using the least squares fit method.

It was highly unlikely that the position of the diode array would exactly coincide with that of the previously analysed photographic plates. However, it was now known approximately which region of the spectrum was expected to be observed for any particular grating angle setting.

The angular dispersion of a grating spectrometer is given by (Jenkins and White, 1957):

$$\Delta\theta/\Delta\lambda = m/d\cos\theta \qquad (4.1)$$

Table 4.1

Wavelengths associated with each grating angle

on the analysed photographic plates

Plate number	Grating angle (degrees)	Centre of observed spectrum (Å)
9	46.0	5521 (first order)
	45.5	5702 "
	45.0	5866 "
	44.5	6042 "
	44.0	6215 "
12	38.8	4063 (second order)
	38.3	4167 "
	37.8	4257 "
	37.3	4351 "
	36.8	4450 "
7	36.5	4511 "
	36.0	4607 "
	35.5	4695 "
	35.0	4786 "
	34.5	4871 "
8	34.5	4869 "
	34.0	4964 "
	33.5	5052 "
	33.0	5148 "
	32.5	5248 "

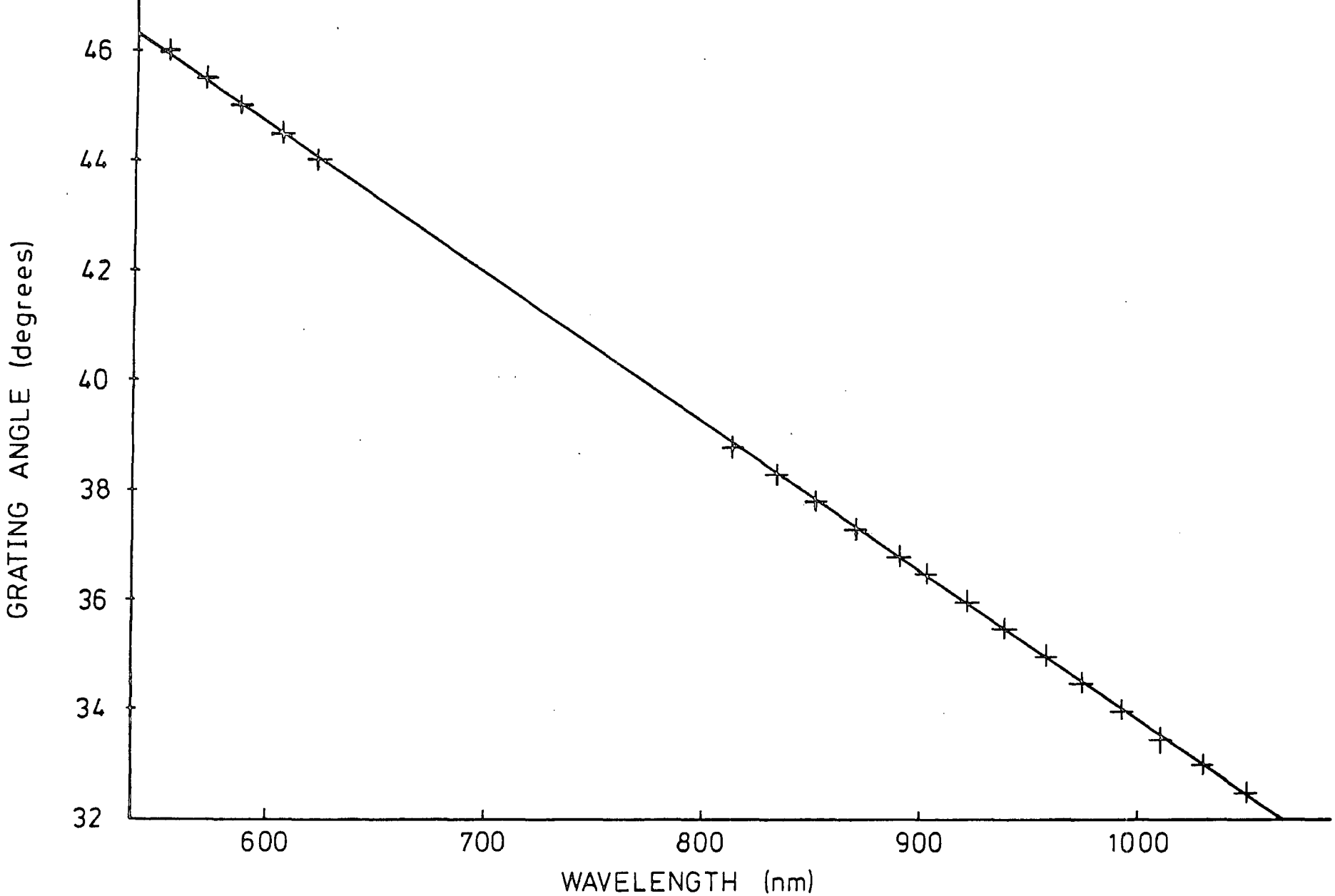


FIG.4.4: GRAPH OF GRATING ANGLE AGAINST WAVELENGTH OF PLATE CENTRE

where θ = angle of the diffracted light from the grating normal

m = order of the observed spectrum

d = line spacing on the grating

From equation (4.1), it can be seen that the smallest dispersion is along the grating normal (where $\theta = 0$), and the dispersion increases slowly on either side of the normal. For small values of θ,

$$\cos \theta \approx 1$$

$$\text{ie. } \Delta\theta/\Delta\lambda \approx m/d$$

Therefore, for different spectral lines in the same order,

$$\Delta\theta \propto \Delta\lambda$$

In other words, unlike a prism spectrometer, the grating spectrometer provides a linear wavelength scale. A small displacement in the position of the array from the position of the photographic plates would therefore not significantly affect the slope of the line in Figure 4.4. Because of this, it was only necessary to locate the centre of the spectra observed with the diode array for a minimum number of different grating angle settings. In this case, four points were located using the diode array data. A line was drawn through these four points parallel to the line obtained from the photographic plate data. This graph (shown in Figure 4.5) now indicates the approximate centre of the diode array (expressed in units of wavelength) for any particular value of the diffraction grating angle.

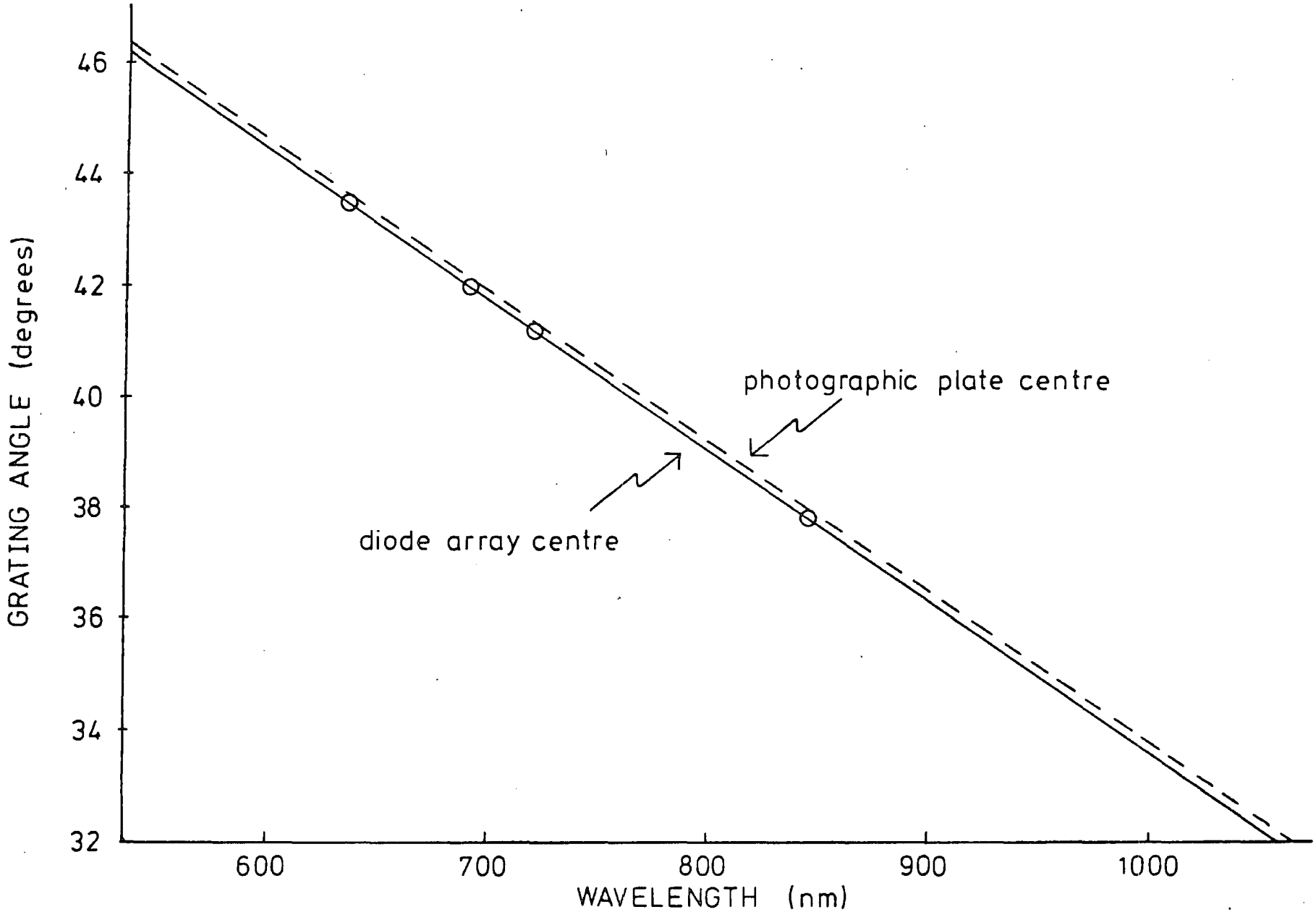


FIG. 4·5: GRAPH OF GRATING ANGLE AGAINST WAVELENGTH OF DIODE ARRAY CENTRE

4.2 Observation Program

The initial planned objective of this project was to observe δ Scuti type variables. A comprehensive survey on δ Scuti stars is given in Baglin, et al, (1973). These stars are short period variables with possible non-radial pulsations. The periods are not necessarily constant. Beat periods have been observed which are believed to be due to the superposition of two or more sinusoids. Difficulties arose in this line of research since not a great deal of work had been published on this type of star. As we were using a new type of detector, it was desirable to compare some of our preliminary results with published work. Also, the response linearity of the Plessey photodiode array had not yet been investigated. Light changes of less than 0.05 magnitudes are experienced by δ Scuti stars, and it was not known if these variations would be observable. However, one observation was made of δ Sct itself, but any features present were smeared out during the long (over 90 minutes) exposure time (see Section 5.6).

The second choice of research area was to investigate the β Cepheid (or β Canis Majoris) type of variable. This is the hottest type of variable known. Much observational research had been carried out on this type of star (eg. see Lesh and Aizenmann, 1978, for a summary of the collected data on β Cepheid stars). Light variations can be as much as 0.1 magnitudes, with approximately 180° phase lag. However, upon determining the positions of well-observed β Cep stars in the sky, it was found that only one or two were visible to us at the time. But quite a long period of time was spent observing one of the β Cep stars with the diode array, namely γ Pegasi (see Section 5.5).

After consultation with Dr. Dianne Harmer of the RGO, it was decided to spend the majority of the telescope time observing luminous supergiants. Dr. Harmer had an interest in supergiants, and it was felt that we could benefit from her knowledge of the subject, and she could possibly gain

further information from our near-infrared observations.

4.2.1 Description of observations

All of the observations described in this work were carried out at the Royal Greenwich Observatory at Herstmonceux. The telescope used was the 30-inch reflector plus spectrometer located in dome A of the Equatorial Group. The Plessey diode array was placed at the coudé focus of the telescope in order that the detector system should remain fixed while the telescope was free to track a particular star across the sky. The observations took place during the period from August 8th to August 26th, 1978. Over this period, data were collected for 107 separate observations of 19 different stars. A further 212 runs were also logged. These latter runs consisted of setting up procedures (eg. optimization of the diode array input voltage levels), noise measurements, detector linearity measurements, neon calibration runs, flat field tungsten calibration runs, and several scrapped runs. A run was scrapped for one or more of the following reasons:

(a) equipment failure - In particular, problems were experienced with the buffer memory in the early stages of the project. Bad contacts in the logic circuitry caused a loss of memory during some runs.

(b) weather conditions - Poor seeing conditions could cause the autoguider to guide off of the star. If this mis-tracking was not noticed and corrected for in time (by manual guidance of the telescope), the run could be contaminated by the moon or other stars. This problem was especially serious when observing stars close to the horizon, where one is looking through a much thicker layer of the earth's atmosphere.

(c) operator error - In particular, the switch activating the run terminate sequence was, on occasion, mistakenly selected (instead of the NDRO switch). If this mistake occurred near the beginning of a

run before a significant signal had been collected, the run was scrapped. This error was primarily due to the close proximity of the two identical switches on the front panel of the integration control module. An improvement on this design will eliminate the problem in the future.

Table 4.2 shows the characteristics of the observed stars, and the number of runs taken at each wavelength interval for each star. The first three columns are self-explanatory. Column four indicates the spectral type of the star. Columns five and six show the absolute and apparent magnitudes (where known) from the authors given in the references. On the continuation of table 4.2, column seven repeats column one. Columns eight to ten indicate the number of runs taken at each of the wavelength regions observed. The particular features under observation were located near to the centre of the 110Å range seen by the array in one run. This positioning was to try to avoid any end effects in the subsequent analysis of the data which might affect the spectral feature. Column eleven lists any particular comments regarding the stars, and column twelve shows the more common name for the brighter stars.

4.2.2 Instrumental set-up

Once the decision was taken as to which stars to look at, and what particular region of the spectrum to investigate, it was necessary to determine which part of the spectrum was available from the angle of the spectrometer grating. See section 4.1.5 for a description of the calibration for the grating. In brief Dr. Dianne Harmer gave us photographic plates exposed to the spectrum of an iron-argon arc. The spectrometer grating angle was known for each strip of spectrum recorded on the plate. By comparison with iron-argon wavelength charts and MIT

Table 4.2

Characteristics of observed stars

Name	HD	HR	Sp	M_V	m_V j
γ Peg	886	39	B2IV	-3.0 a	+2.8
β And	6860	337	M0III	+0.2 b	+2.0
α Per	20902	1017	F5Ib	-4.5 c	+1.8
ν Per	23230	1135	F5II	-2.2 d	+3.8
ε Aur	31964	1605	A8Ia	-7.0 e	+3.0
α Boo	124897	5340	K2IIIp	-0.3 b	+0.1
89 Her	163506	6685	F2Ia	-7.0 f	+5.5
α Lyr	172167	7001	A0V	+0.5 d	0.0
δ Sct	172748	7020	F3III-IV	-	+4.7
γ Cyg	194093	7796	F8Ib	-5.6 g	+2.2
α Cyg	197345	7924	A2Ia	-7.5 h	+1.3
ξ Cyg	200905	8079	K5Ib	-4.4 i	+3.7
β Aqr	204867	8232	G0Ib	-4.5 i	+2.9
ε Peg	206778	8308	K1Ib	-4.4 i	+2.4
α Aqr	209750	8414	G2Ib	-4.5 i	+2.9
ρ Peg	216735	8717	A1V	+0.7 d	+4.9
56 Peg	218356	8796	K0Ib	-2.3 i	+4.8
-	218396	8799	A5	-	+5.9
γ And	223047	9003	G5Ib	-4.5 i	+5.0

Table 4.2 (continued)

Name	n(H_α)	n(7774)	n(8446)	Comments	Common name
γ Peg	12	1	4	β Cep	Algenib
β And	-	-	4	-	Mirach
α Per	-	3	-	-	Mirfak
ν Per	-	1	-	-	-
ϵ Aur	-	3	-	P Cyg	-
α Boo	3	-	-	-	Arcturus
89Her	2	1	1	H_α emission	-
α Lyr	21	14	4	-	Vega
δ Sct	-	-	1	δ Sct	-
γ Cyg	-	3	-	-	Sadir
α Cyg	2	4	1	P Cyg	Deneb
ζ Cyg	4	-	-	-	-
β Aqr	-	3	-	-	-
ϵ Peg	5	-	3	-	Enif
α Aqr	-	1	-	-	Sadalmelik
ρ Peg	-	2	-	Broad lines	-
56Peg	-	1	-	-	-
(HR8799)	-	1	-	-	-
γ And	2	-	-	-	-

Table 4.2 (continued)

References

a: Lesh & Aizenman (1973)

b: Bonsack & Culver (1966)

c: Sorvari (1974)

d: Allen (1976)

e: Hack & Selvelli (1979)

f: Osmer (1972)

g: Baker (1974)

h: Abt (1957)

i: Kraft et al (1964)

j: Abt & Biggs (1972)

wavelength tables, the wavelength of the center of each strip was calculated for that particular grating angle, both in the first and second orders. For the first order the reciprocal dispersion was calculated as being $\sim 9\text{Å mm}^{-1}$, and $\sim 4.5\text{Å mm}^{-1}$ in the second order. Because the exact position of the diode array relative to that of the photographic plate was not known, several calibration lines had to be located on the diode array output. Since the mean dispersion of the spectrometer remained relatively unchanged over the focal area, a line with the same gradient as above could be drawn through these new calibration lines. This yields the approximate central wavelength of the diode array for a given angle of the spectrometer grating. Since the array could observe $\sim 110\text{Å}$ at one time (in the first order), the accuracy of the grating angle ($\pm 0.^{\circ}1$) was sufficient to ensure that the desired spectral feature would fall on the array. Nevertheless, before and after each wavelength run a neon spectrum was recorded. This allowed a check to be made on the range of wavelengths observed across the diode array.

4.3 Data Acquisition and Reduction

4.3.1 Data acquisition

As described in section 3.3 on the operation and control of the experiment, the control of the equipment is carried out by means of a PDP11-03 minicomputer. The communication language is CATY2.

The CATY language was developed at the Daresbury Nuclear Physics Laboratory jointly with Francis Golding of Applied Computer Services Limited, Manchester. CATY is a subset of BASIC with many advantages, such as ease of modification and editing of programs. In BASIC, each statement is compiled and executed sequentially. This differs from CATY. On the RUN command, the entire CATY program is first compiled into object

form and then executed. The program must be compiled again each time the RUN command is issued, as the object program is not stored. Since the entire program is compiled, the time between the execution of consecutive statements is greatly reduced, thus favouring real-time applications. Also the direct manipulation of CAMAC hardware is a prominent feature of the CATY language.

The original version of CATY (called CATY1) has few logical and arithmetic facilities. It is intended mainly for the testing of equipment with a small computer. CATY2 is an upgraded version of CATY1, with, among other improvements, an extended range of arithmetic functions. The CATY2 compiler for the Digital Equipment Corporation PDP11-03 computer was obtained from Francis Golding Associates of Manchester. Modifications to the system are being carried out to include a floppy disk unit to run under RT11. FGA are to supply RT11 handlers to support our existing peripherals. A version of CATY which runs under RT11 has already been obtained, but presently, without the extra RT11 device handlers, all input and output is via disk.

The data acquisition program (Hedge, et al, 1978) collects information about the conditions of the diodes in the array. The program stores the diode array information as four blocks of computer memory. An intermediate CAMAC memory exists in which each of the four blocks in turn is sequentially built up and stored. When a readout frame is complete, the 512 words of information are transferred from this buffer memory to the resident memory of the computer. At the end of an observing run, when all four blocks have been sent to the computer memory, the operator can decide whether to display, print, or dump the data.

4.3.2 On-line data storage

The first block of memory contains the fixed pattern noise (FPN1) on all of the diodes prior to any signal being induced in the array. Figure 4.7 shows an example of FPN1 before commencing an observation. The signal levels in FPN1 are collected after recharging the diodes for twenty frames. This repetitive recharge ensures the removal of any residual signals present on the diodes. With these dark level signals in the memory, the starting point of an integrating run is available.

The second memory block contains the most recent non-destructive readout levels. Figure 4.8 shows a typical NDRO collected from the array. The array is scanned as if for a double sample readout (see below), but the recharge pulses are suppressed. At any time during a light integrating run, a manually requested interrupt (via the front panel of the integration control module) will trigger a software routine to commence a NDRO cycle. As well as the absolute signal levels of the diodes being presented to the operator, a reduced NDRO is calculated and displayed. Figure 4.9 shows an example of this reduced output, consisting of 128 signal levels. These levels are composed of the sum of the signals on a pair of adjacent diodes, minus the sum of the two fixed pattern noise levels (from FPN1) for the same pair of diodes. In the case of the example illustrated, the central feature is the Balmer H-alpha absorption line in 89 Herculis. The displayed output of the reduced NDRO is scaled such that the vertical limits of the display approximate to the signal levels between which the ADC is able to digitize. In this way, NDRO observations can be used to monitor the build-up of the spectrum until a satisfactory amount of signal has been received. The run can then be manually terminated, knowing that no signal saturation in the electronics has taken place, and a good signal to noise ratio has been achieved. Since the construction of the on-board shift register is such that pairs

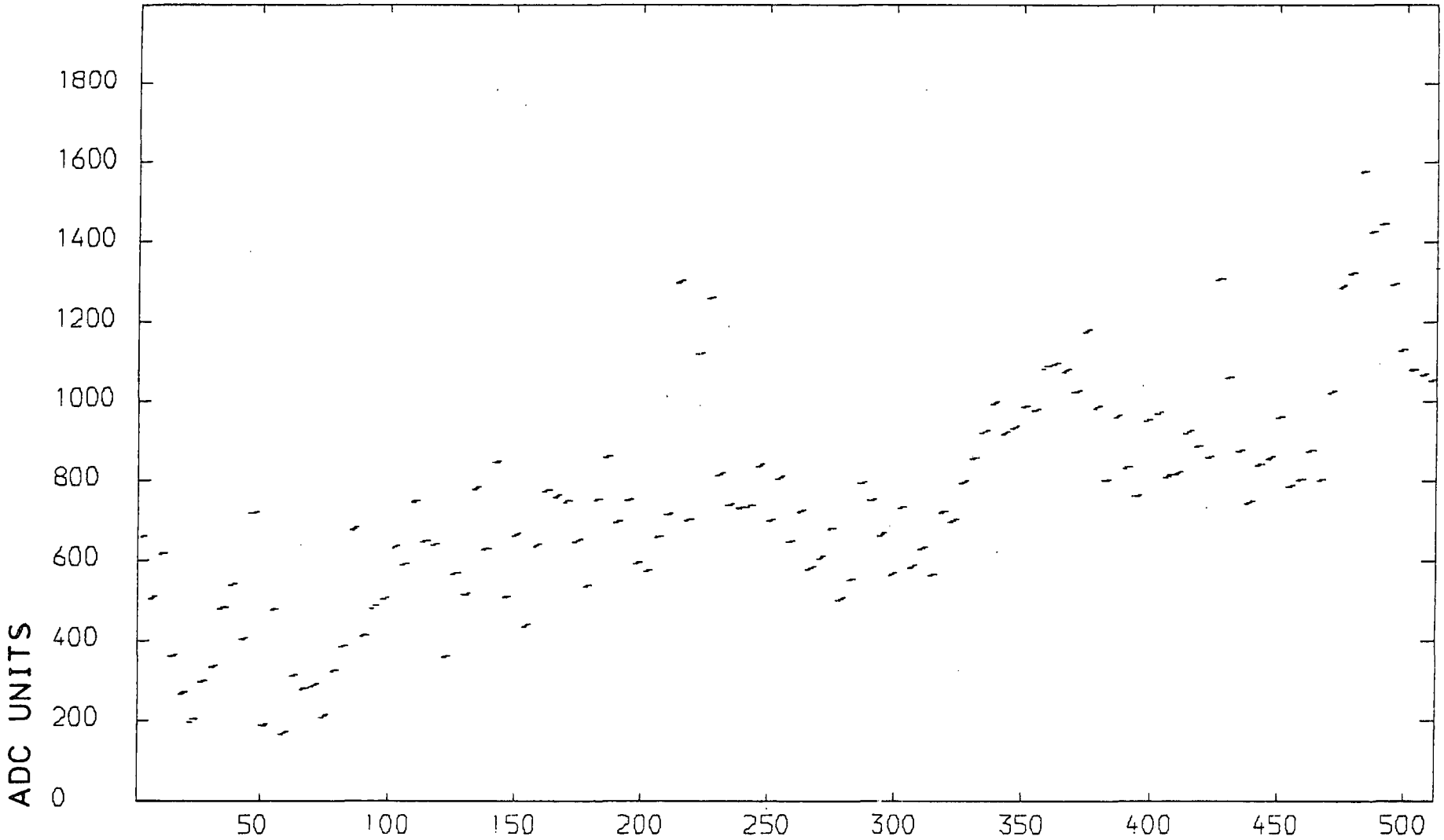


FIG. 4.7: EXAMPLE OF FIRST FIXED PATTERN NOISE MEASUREMENT

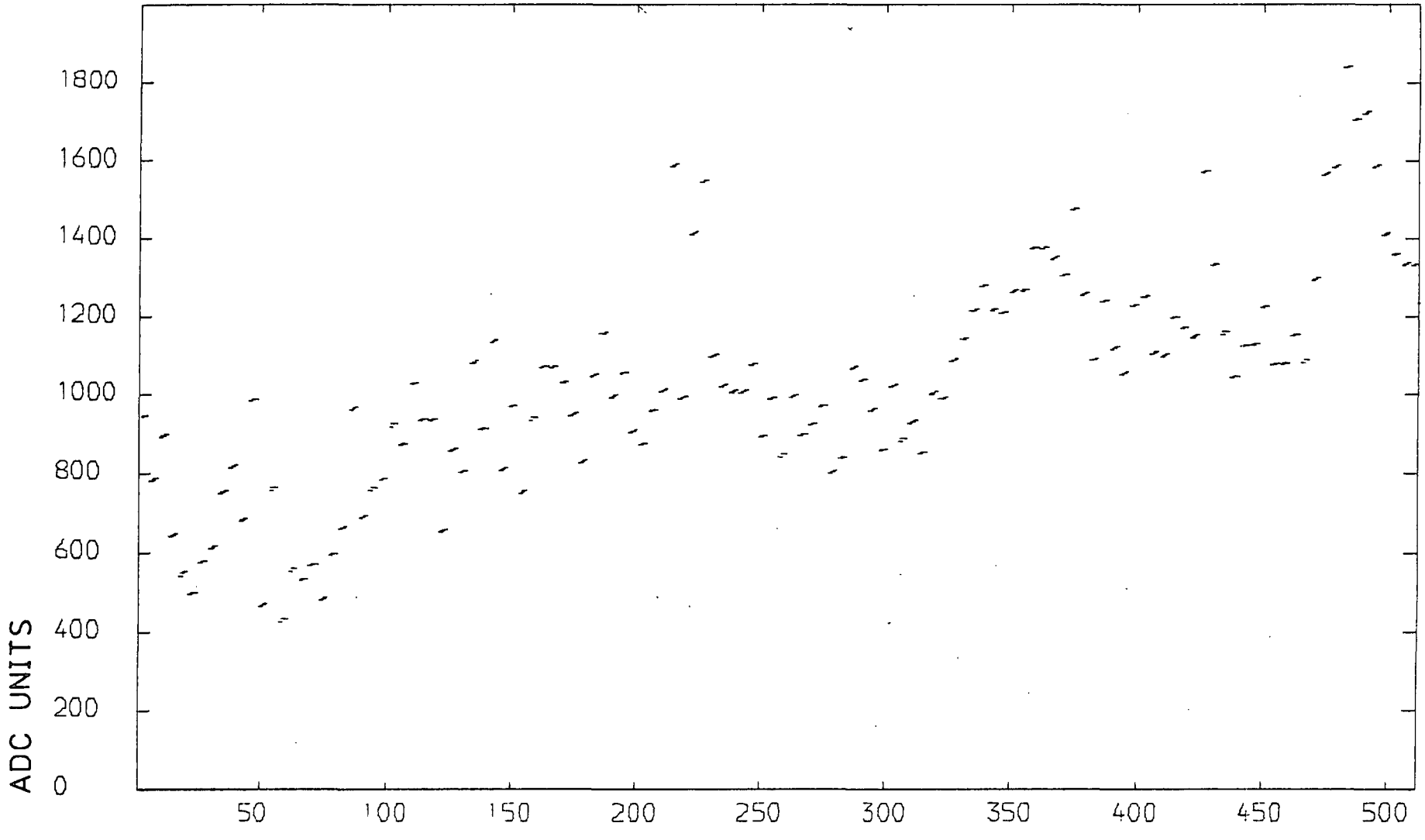


FIG. 4.8: EXAMPLE OF FINAL NON-DESTRUCTIVE READ-OUT

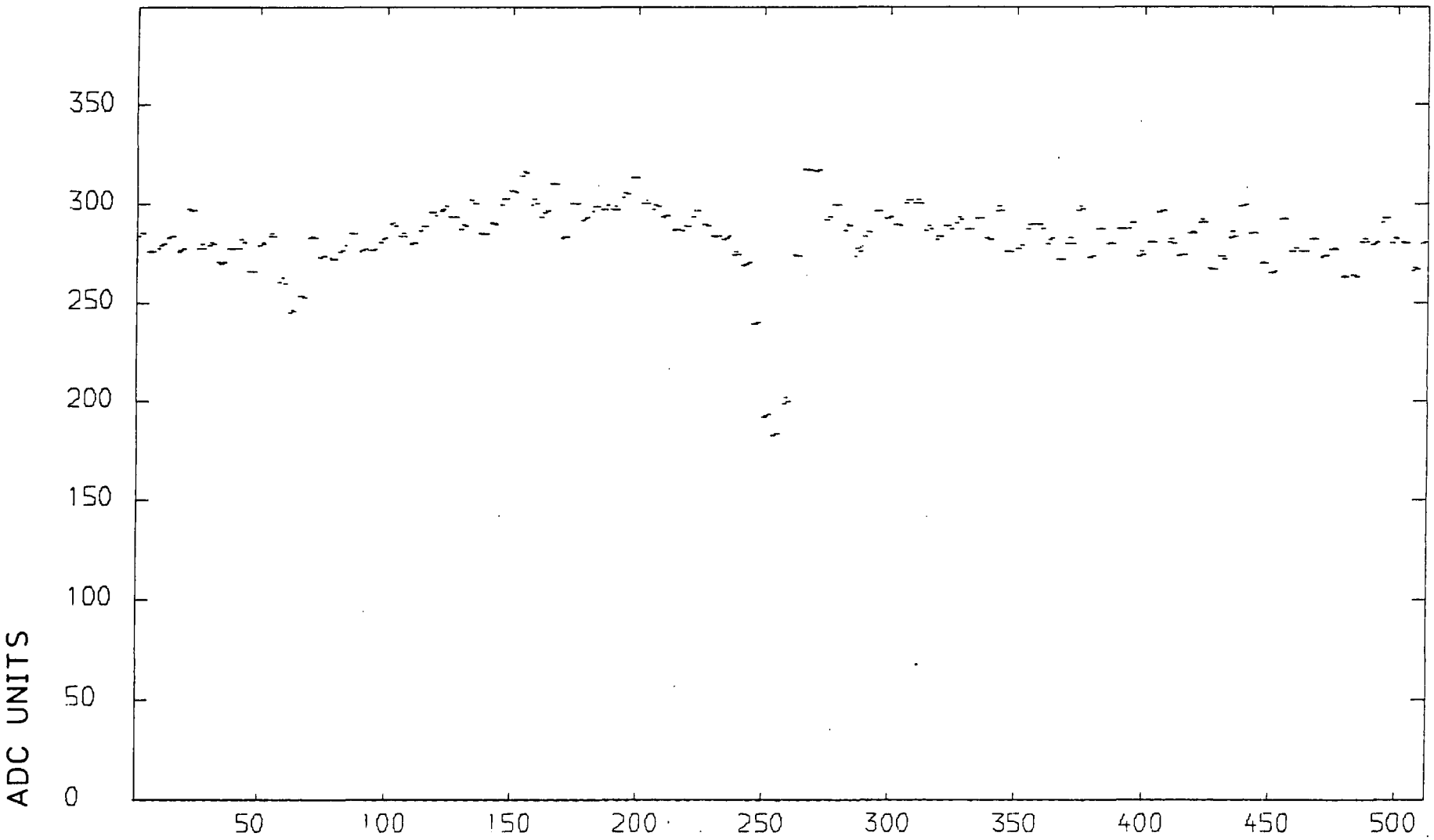


FIG. 4.9: EXAMPLE OF REDUCED NON-DESTRUCTIVE READ-OUT

of diodes are addressed at a time (to conserve area on the chip), the resolution of the observed spectrum is only half as good as that with the final double sample readout. Each time that a NDRO is requested and executed, the new NDRO levels are written into the second memory block. Since unlimited memory space is not available, and only the current NDRO is considered to be of interest, overwriting previous NDRO's is the most appropriate method of storing this information.

The third block of computer memory contains the double sample output from the diode array. Figure 4.10 shows the double sample readout after the NDRO from Figure 4.8. After the completion of a run, the data acquisition program is used to control the method of collecting the total array information. The double sample readout is achieved by first recharging one of the diodes in the linked pair, and then the other. For each pair of diodes, the following four signal levels exist:

a) the base line level of the ADC;

b) the sum of the signals on both diodes;

c) the signal after one diode has been recharged;

d) the resultant signal after both diodes have been recharged (i.e. the diode zero level).

By subtracting the appropriate levels, one can obtain the signal level on each individual diode due to the light (and/or the dark current) it has been exposed to during the integration period. Figure 4.11 illustrates this reduced double sample readout. This information plus the total array information (from Figure 4.10) are displayed to the operator. The reduced spectrum is that information which is to be saved in a short dump (see section 4.3.3).

The final block of memory in the computer contains a second fixed pattern noise readout (FPN2). After the execution of a double sample readout, the diodes are left in a fully charged condition, allowing another chance of observing the fixed pattern noise. After twenty recharges (to remove

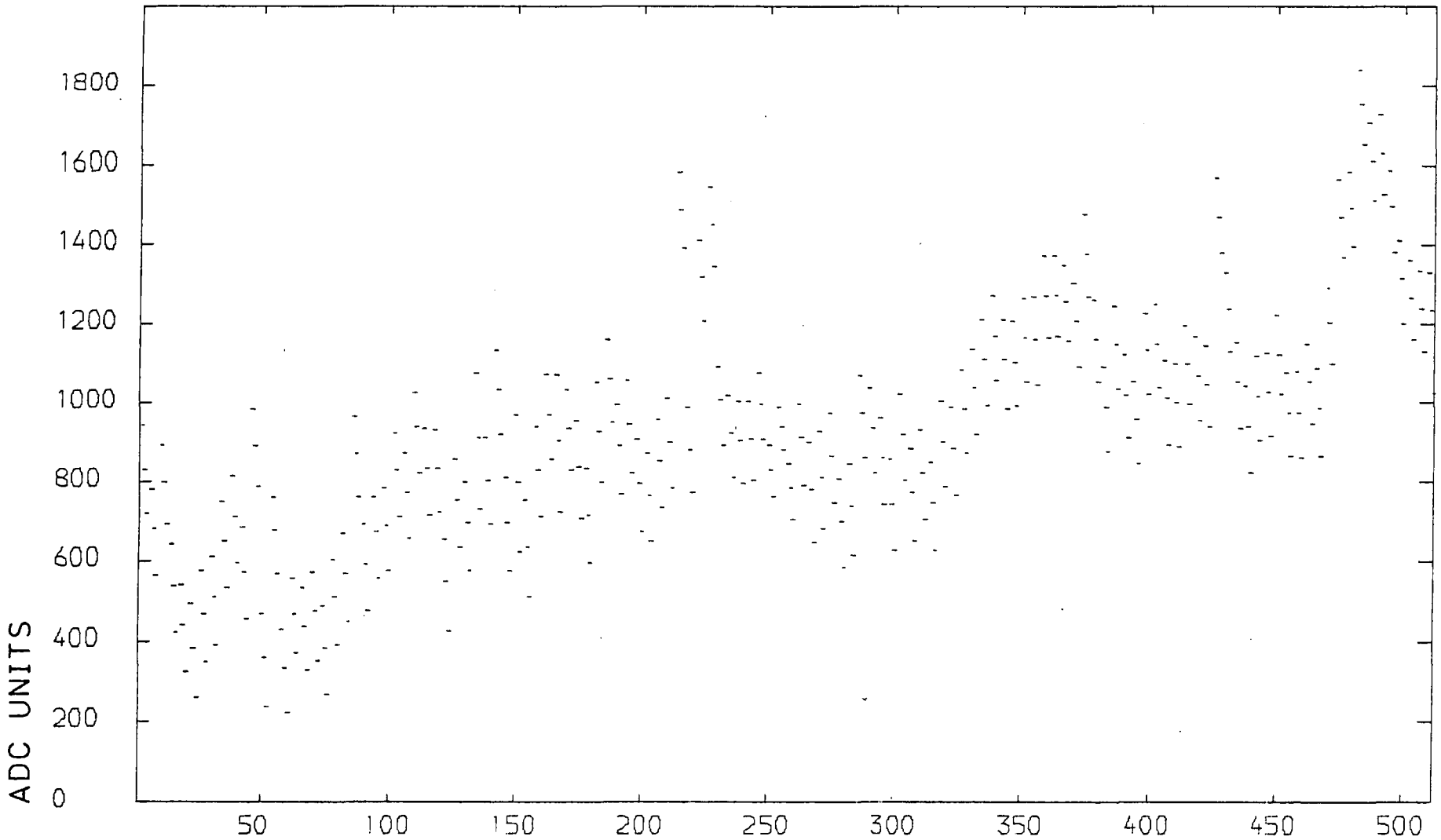


FIG. 4.10: EXAMPLE OF DOUBLE SAMPLE READ-OUT

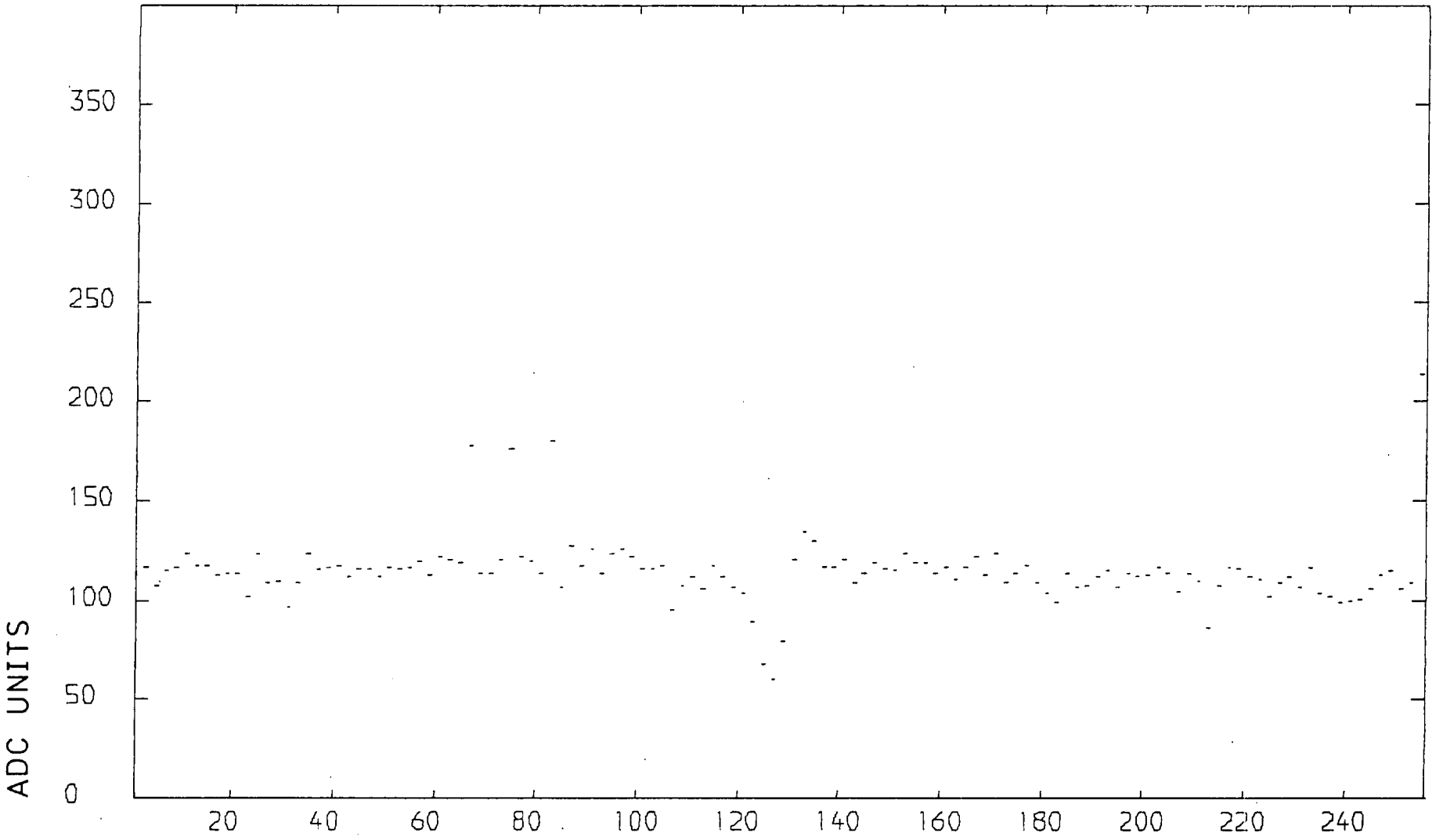


FIG. 4.11: EXAMPLE OF REDUCED DOUBLE SAMPLE READ-OUT

all of the accumulated signal), FPN2 is recorded (see Figure 4.12) so that it can be compared to FPN1. Comparison between the two fixed pattern noise spectra can indicate whether or not the diode array has experienced any overall drift in the spatial noise over the time taken to complete a particular integrating run. The differences between the two fixed pattern noise readouts described above is shown in Figure 4.13. One can see that, with the obvious exception of three diode pairs, a relatively constant change in the zero levels across the array has occured during this integration period. The length of this particular run was 29 919 frame times, or 49.87 minutes (allowing for the 100 ms duration of each frame). One ADC unit measured along the ordinate is equivalent to 1.22 mV (see section 3.3.1). This gives an average drift of the diode zero levels of $0.2\mu V.s^{-1}$ for this particular run, or a total drift of 61.0 mV. One possible cause of this drift could be attributed to small temperature changes in the diode array chip, being more noticeable after long integrations. The change in the ADC base line levels is very small. This is expected with the Burr Brown ADC used, with a maximum drift of 23 ppm $^{o}C^{-1}$. This is equivalent to a drift of 28 nV $^{o}C^{-1}$, or less than one ADC unit for a ten degree temperature change. The three "high" diode pairs are believed to be due to some form of pick-up during the double sample readout. This fault occured intermittently on previous and subsequent runs, and could not be entirely eliminated.

4.3.3 Off-line data storage

The only output system available for the off-line storage of data was paper tape. As each run was completed on the telescope, the collected data was output onto paper tape in what was termed either a long dump or a short dump. Both dumps contain descriptive parameters

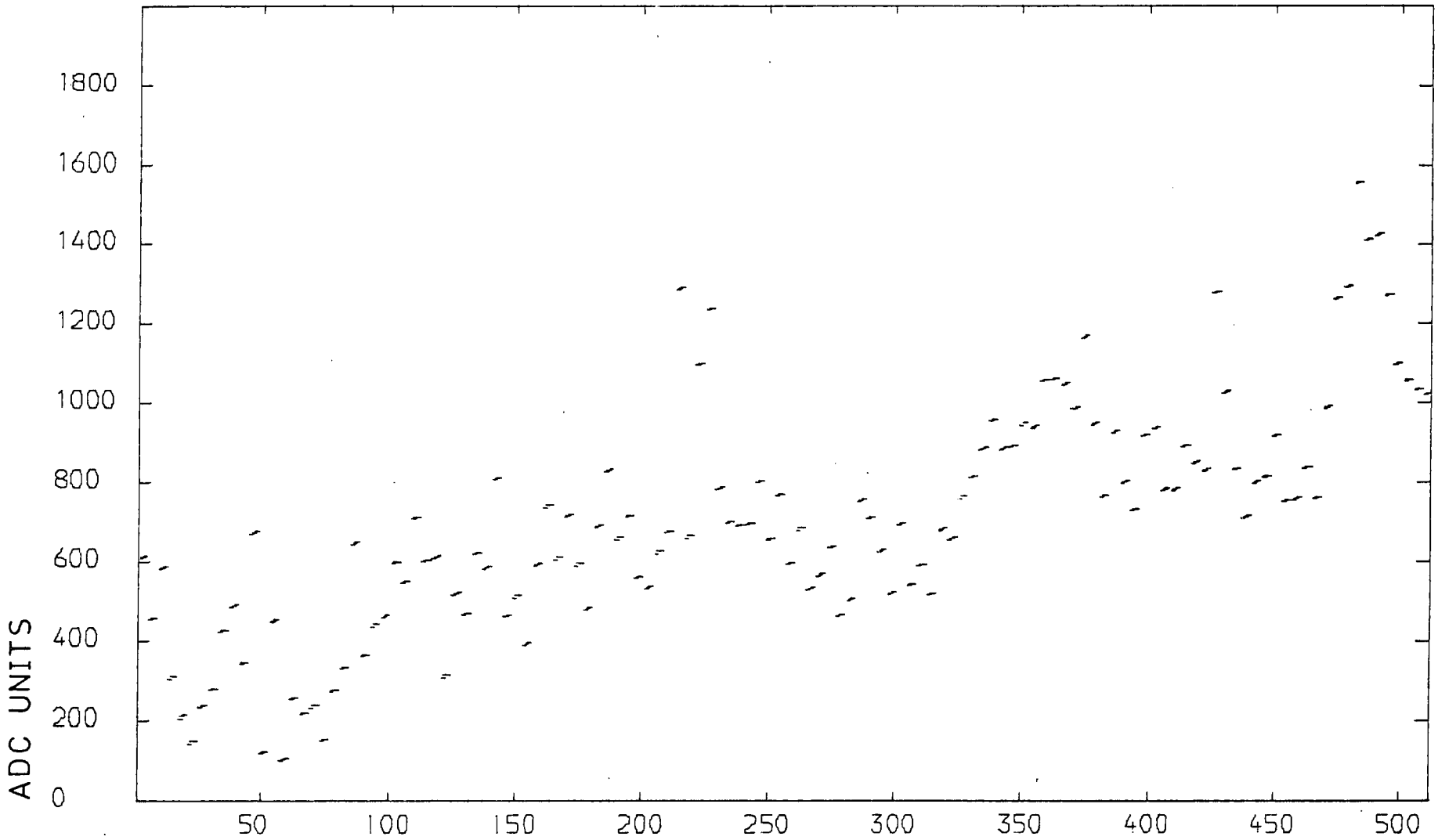


FIG. 4.12: EXAMPLE OF SECOND FIXED PATTERN NOISE MEASUREMENT

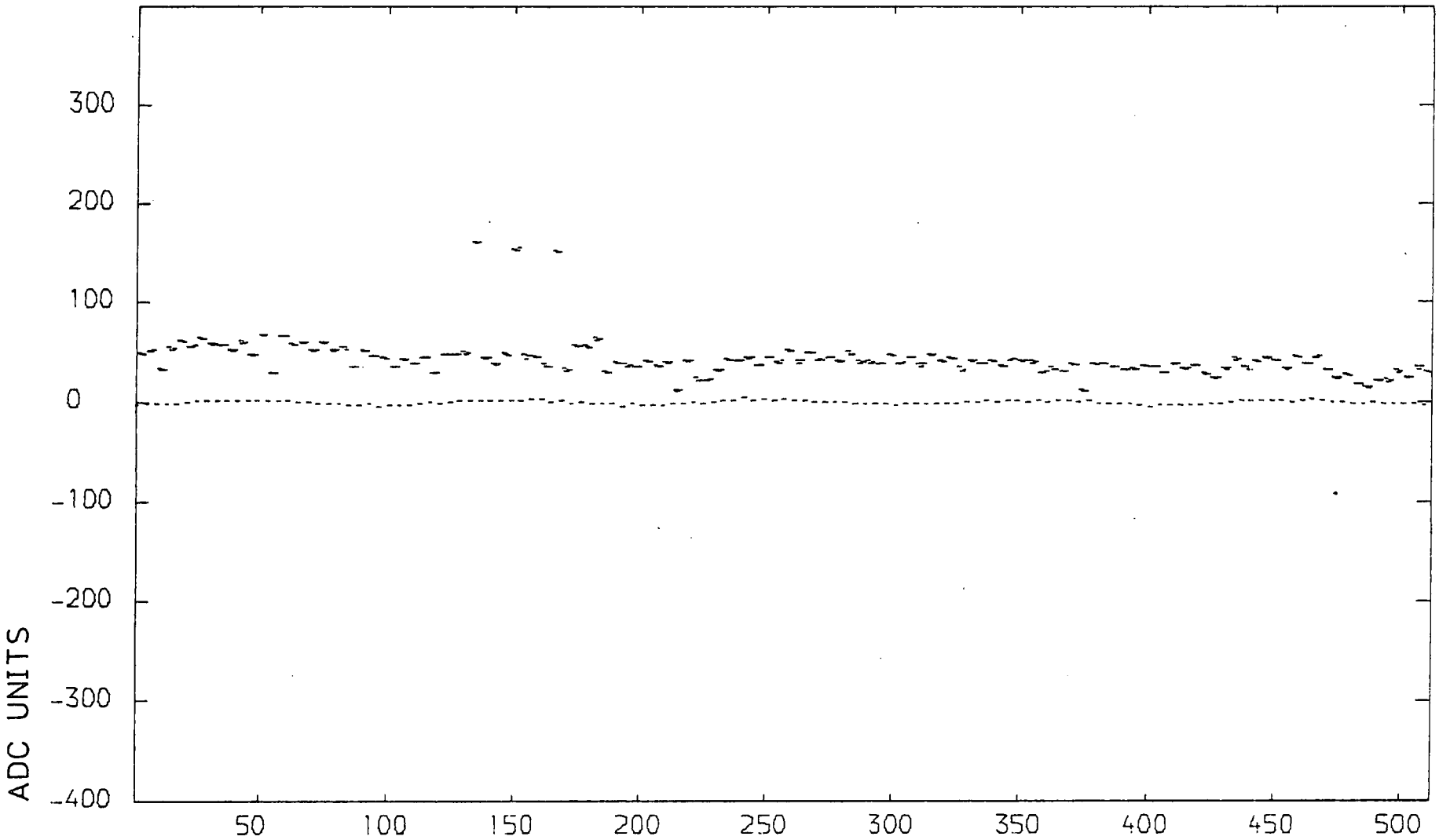


FIG. 4.13: EXAMPLE OF FIXED PATTERN NOISE DIFFERENCE

of that particular run, such as the run identification number, exposure time, time of last NDRO, etc. (Time in this sense is measured in the number of frames elapsed since the start of the run, each frame being 100 ms long). These parameters are located at the start of the tape and form the tape header. An example of the information available from a typical tape header is shown in Figure 4.14. This header is, in fact, from the run information displayed in section 4.3.2 above.

The short dump contains information regarding the amount of signal collected by each diode. This is the reduced double sample spectrum, and is the data from which the astronomical information is derived.

The long dump contains:

a) an initial fixed pattern noise spectrum (FPN1);

b) the most recent NDRO spectrum;

c) the double sample readout, exhibiting the sum of two adjacent diode signals, the signal with one of the pair of diodes recharged, the signal with both diodes recharged, plus an ADC base line level for each diode pair;

d) a final fixed pattern noise spectrum (FPN2).

Each of the four blocks of computer memory contains 512 words of information.

For the majority of runs, a short dump was considered sufficient, as the signal collected by each diode for that particular observation is the most important information to be gathered. However, if any problems or difficulties occurred during the exposure, or a star was thought to be particularly interesting at a given wavelength, then the information was retained in a long dump of that run. The information contained in a short dump can be retrieved from the long dump at any time after the execution of the observing run.

RUN NUMBER	114
INTEGRATION PERIOD (SELECTED)	100000
NO. OF INTEGRATIONS	1
FPN1 AT FRAME	20
INTEGRATION BEGAN AT FRAME	35
LAST NON DESTRUCTIVE READ AT	29815
LAST INTEGRATION FINISHED AT FRAME	29954
FPN2 AT FRAME	30004
INTEGRATION PERIOD (ACTUAL)	29919

FIG. 4.14: TYPICAL EXAMPLE OF TAPE HEADER

4.3.4 Data transfer to NUMAC

The reduction and analysis of the collected stellar spectra were carried out at Durham using the Northumbrian Universities Multiple Access Computer, or NUMAC. As no fast paper tape reader was available at Durham at the time, this left only two alternative methods for getting the data into NUMAC. The first method was by means of submitting a batch job to read in the paper tape on a mechanical reader and store the data on a personal disk file. The batch job had to be submitted on punched cards along with the paper tape to input. At times it would take several days for the file to appear on current disk space. Also, there was no guarantee that the paper tape would ever reappear. A quicker and more reliable method, though more tedious, was to first obtain a hard copy printout of the short dump data on our own ASR33 teletype via the PDP11-03 mini-computer system. From this hard copy the data were manually typed into a disk file via a NUMAC VDU. Each file of 256 numbers would take approximately fifteen minutes to input. In this way, the data were available for immediate use. Admittedly, for a large number of runs, the first method would be more efficient. But as each of the run outputs was dumped onto a piece of paper tape only about two metres long, a large number of runs would have to be copied onto a single paper tape so that the possibility of the misplacement of individual run data tapes could be eliminated. Subsequently, since usually only one or two spectra were to undergo analysis at any given time, the manual input of the data directly onto computer filespace was adopted.

4.3.5 Data reduction

The first step in the reduction of the data was to counteract the odd-even effect (ie. the observed splitting between adjacent diode signal

levels).

Due to the geometrical construction of the array chip, the on-board circuitry is located on either side of the line of detector elements. The electronics associated with odd-numbered diodes are situated on one side of the array, and those associated with even-numbered diodes are located on the other side (see Figure 4.16a). No part of the chip is masked, and as such, light falling onto areas other than the detector elements gives undesireable effects. When light overspills onto, say, the left hand side electronics, alternate diode signal levels appear to increase (see Figure 4.16b). This apparent increase is due to the gain of the on-board amplifiers being affected by the incident light. Overspill of light onto the right hand side electronics causes a similar effect on the complementary set of alternate diodes. The actual effect on the gain is a function of the sensitive area illuminated, which cannot easily be measured.

This splitting problem is further complicated by the fact that the spectrum of light and the array of diodes are not colinear (see Figure 4.16c). This results in varying degrees of amplifier gain effects along the line of array elements. A FORTRAN program was written to attempt to rectify this problem.

In the reduction program, the difference in signal levels between pairs of adjacent diode levels is divided by the sum of the signals of the same two diodes. Extreme values are removed, as these pertain to spectral features. Only from near the continuum can the odd-even effect be determined. At this stage in the analysis, there are up to 128 values with which to continue. These resultant values are used to calculate the best straight line that can be drawn through them all by the least squares method. The equation of the line yields the predicted amount of splitting between any two adjacent diodes along the spectrum. The

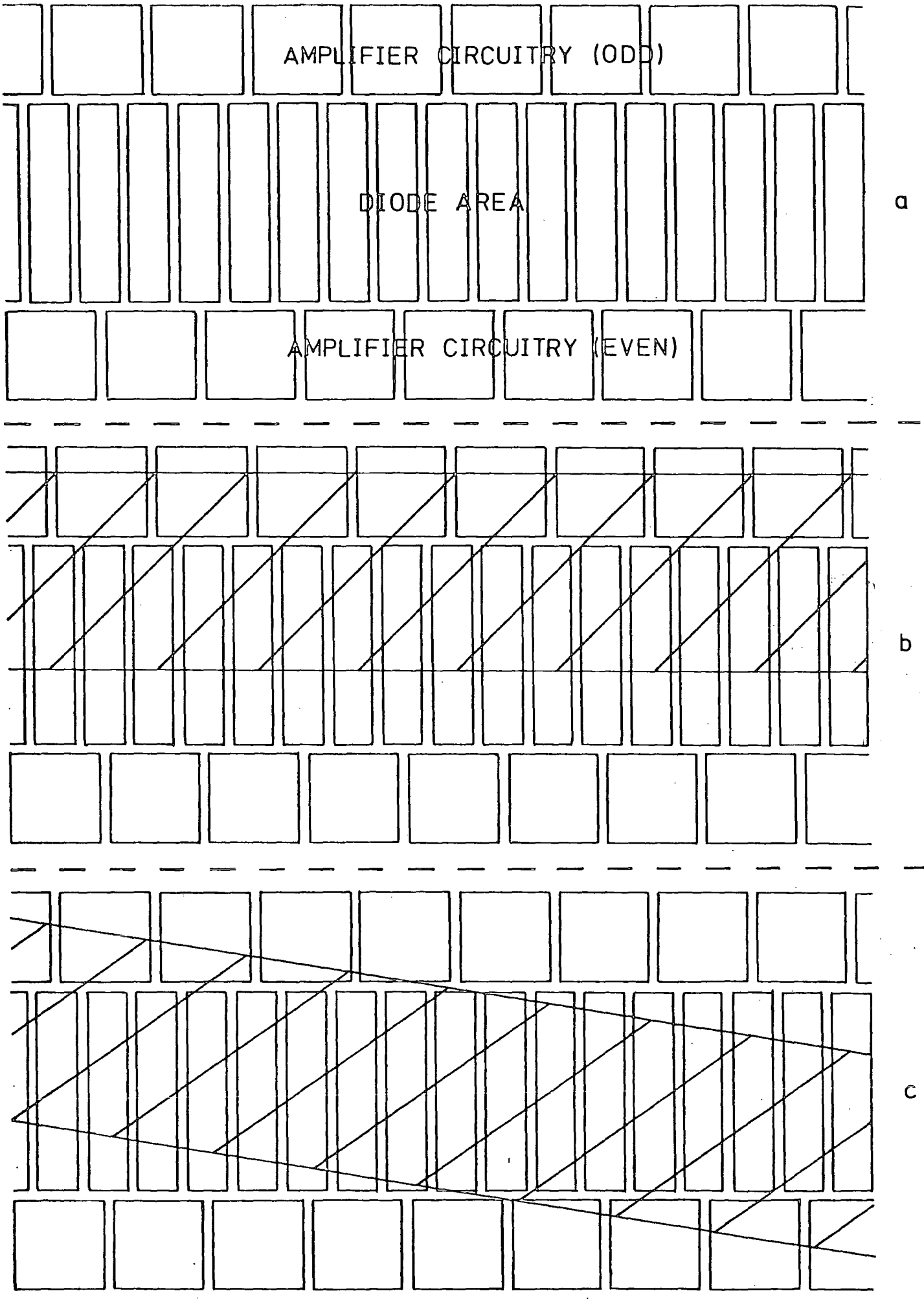


FIG. 4-16: SCHEMATIC OF DIODE ARRAY SHOWING HOW OVERSPILL OF LIGHT CAN AFFECT RESULTS

set of alternate diode levels which are high are reduced by the appropriate splitting factor in an attempt to diminish the unwanted signal as far as possible. The data are now ready for the removal of the diode to diode variations. Figure 4.17a shows the raw spectrum of the K-type supergiant epsilon Pegasi (ε Peg), with H-alpha as the strong absorption feature on the right hand side. Figure 4.17b illustrates the resultant spectrum when the odd-even effect has been removed as far as possible.

In order to remove the diode responsivity variations from the data, it is necessary to divide each stellar spectrum by a flat field spectrum. After a run was completed at a particular diffraction grating setting, a tungsten lamp was introduced into the light path in place of the star (see section 4.1.5). This source of white light gives the nearest approximation to a flat field. Small variations were encountered in the flatness of the illumination. For example, a fall off in intensity at the blue end of the tungsten spectrum was occasionally observed. The image of the star on the slit was believed to be not in the same position as the image of the centre of the tungsten bulb. Light variations across the image of the bulb could yield this effect. In the same manner as for the raw stellar spectra, the tungsten data are also corrected for splitting. In addition, the blue end fall off is also corrected for, to give a nearly straight line on which the diode variations are superposed. The stellar spectrum is then divided by the appropriate corrected tungsten spectrum (both spectra collected with the grating angle unchanged) to yield a spectrum in its best form for analysis. The final reduced form of the spectrum of ε Peg from above is shown in Figure 4.17c.

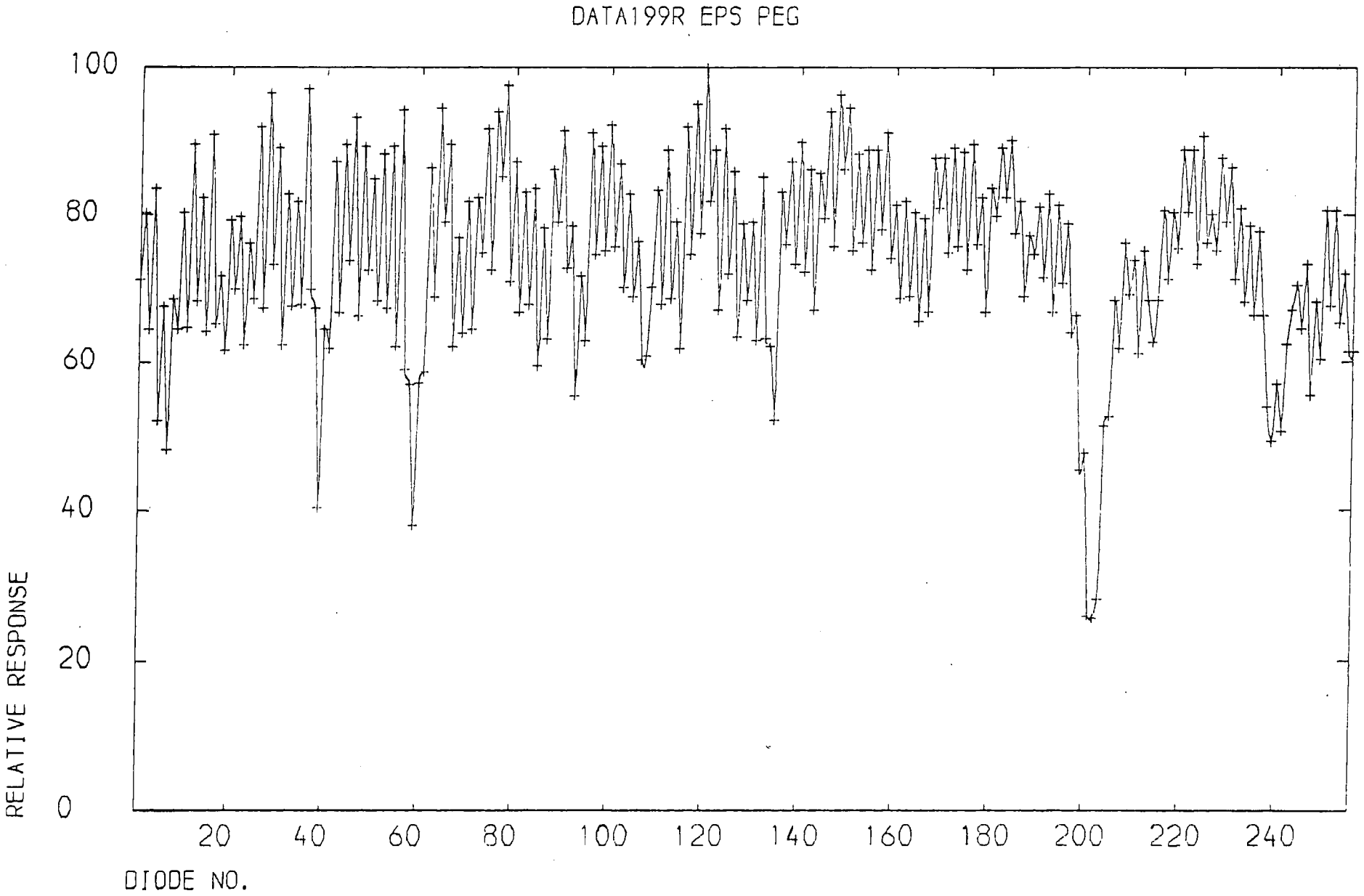


FIG. 4.17a : AN EXAMPLE OF A RAW SPECTRUM

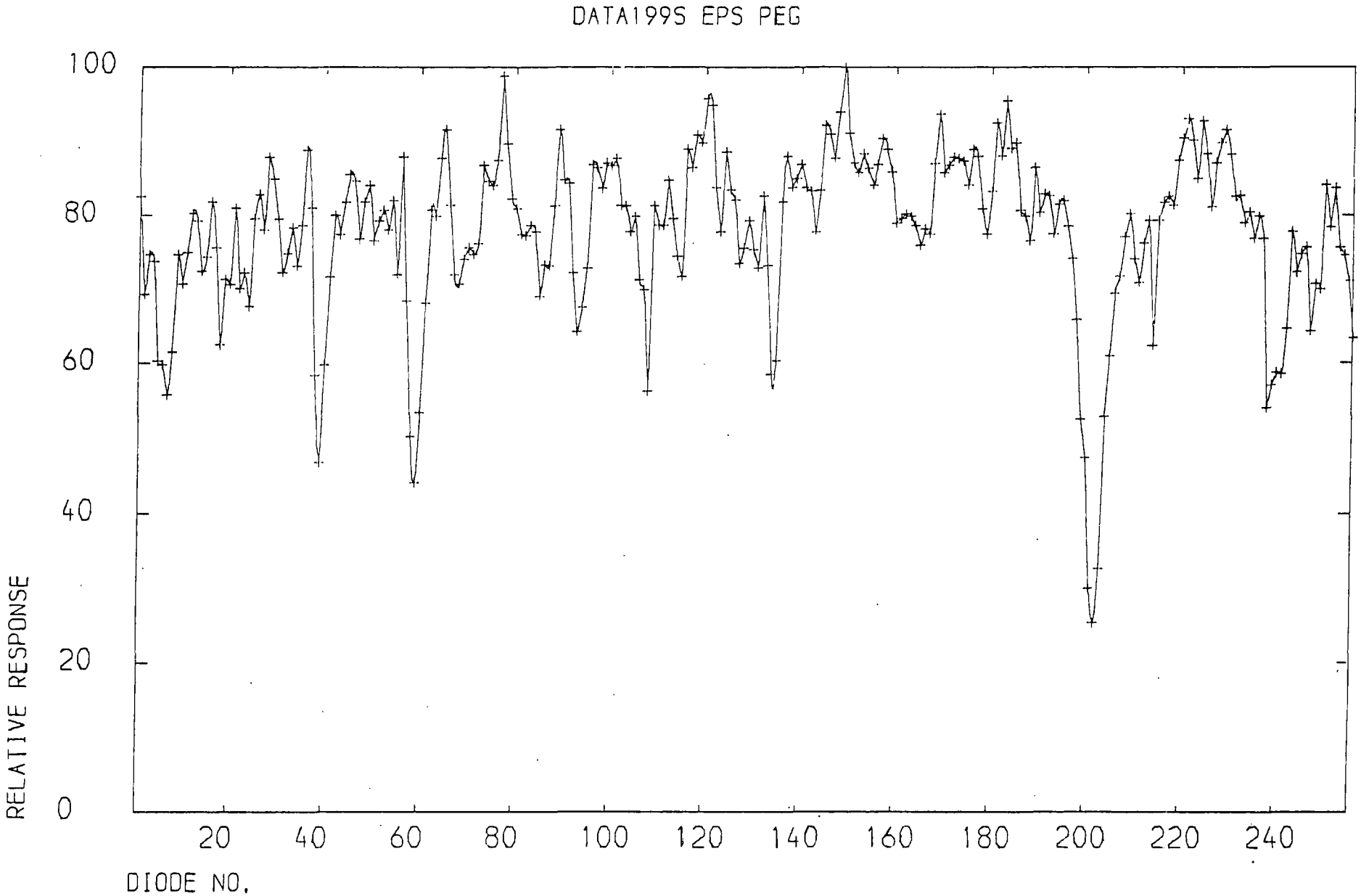


FIG. 4.17b : SPECTRUM CORRECTED FOR SPLITTING

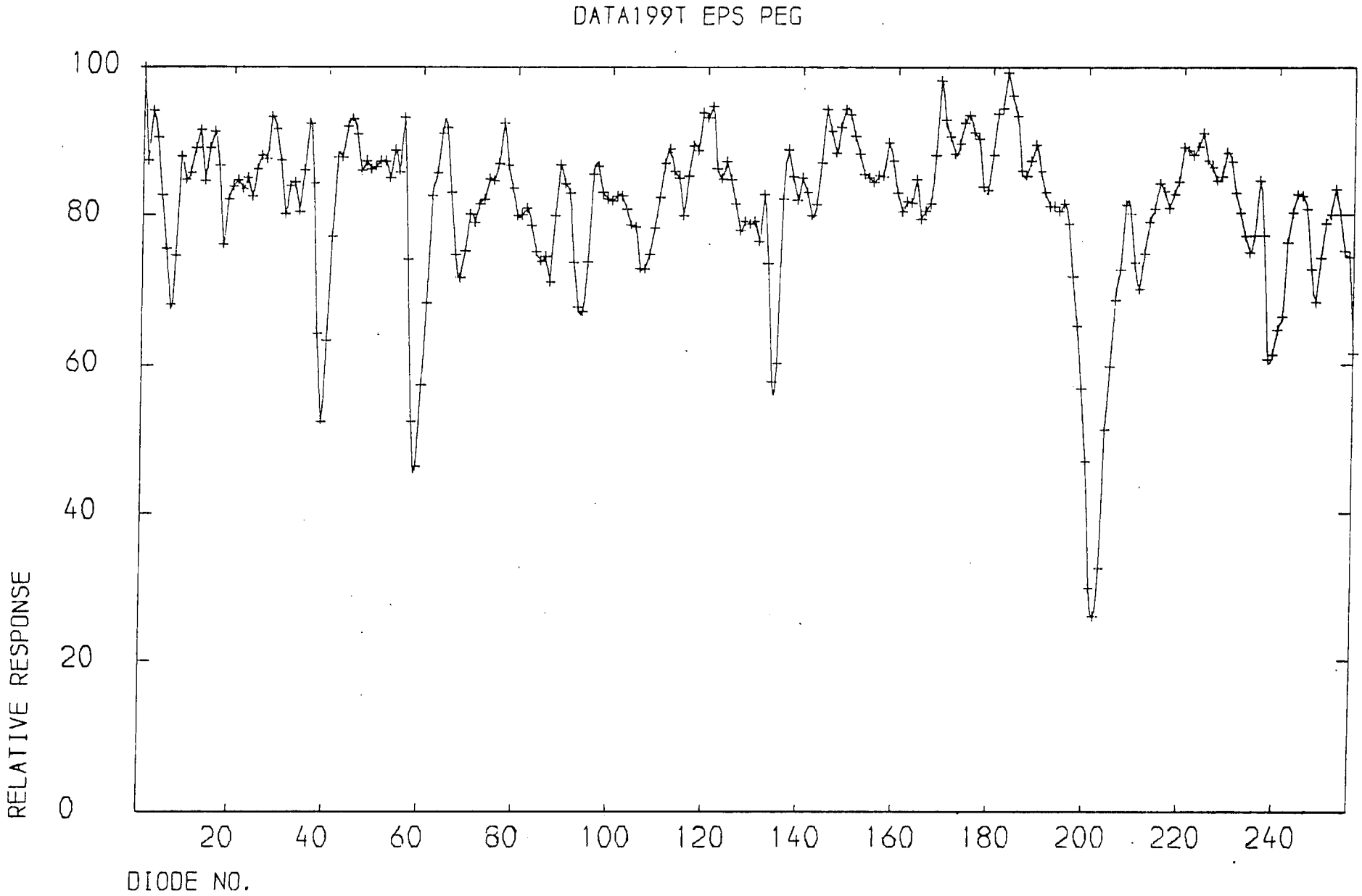


FIG. 4.17c : SPECTRUM IN FINAL REDUCED FORM

REFERENCES

Abt, H.A.: 1957, Astrophys. J., 126, 138.

Abt, H.A., and Biggs, E.S.: 1972, "Bibliography of Stellar Radial Velocities", Latham Process Corp., New York.

Allen, C.W.: 1976, "Astrophysical Quantities", Athlone Press, London.

Baglin, A., Breger, M., Chevalier, C., Hauck, B., le Contel, J.M., Sareyan, J.P. and Voltier, J.C.: 1973, Astr. Astrophys., 23, 221.

Baker, P.W.: 1974, Publs. astr. Soc. Pacif., 86, 33.

Bonsack, W.K. and Culver, R.B.: 1966, Astrophys. J., 145, 767.

Hack, M. and Selvelli, P.L.: 1979, Astr. Astrophys., 75, 316.

Harding, G.A., Palmer, D.R. and Pope, J.D.: 1968, R. Obs. Bull. No.145.

Harmer, C.F.W.: 1978, personal communication

Hedge, A.R., Breare, J.M., Campbell, A.W., Hopkinson, G.R. and Humrich, A.: 1978, "Applications of CAMAC to Astronomy", in Proc. ESO/SRC Conf., CERN, Geneva, 210.

Jenkins, F.A., and White, H.E.: 1957, "Fundamentals of Optics", 3rd ed., McGraw-Hill, New York.

Kraft, R.P., Preston, G.W., and Wolff, S.C.: 1964, Astrophys. J., 140, 235.

Lesh, J.R., and Aizenman, M.L.: 1973, Astr. Astrophys., 22, 229.

Osmer, P.S.: 1972, Astrophys. J. Suppl., 24, 247.

Sorvari, J.M.: 1974, Astr. J., 79, 1416.

CHAPTER FIVE

PRELIMINARY OBSERVATIONS, DATA HANDLING, AND DATA ANALYSIS

5.1 Spectrum Collecting Procedure

Once the angle of the spectrometer grating (with respect to a fixed plane) was calibrated on a wavelength scale (see section 4.1.5) it was possible to select any position of the spectrum from 4000 Å to 1.1 μm (in a 110 Å window) for observation. The summary of steps taken in procuring a stellar spectrum are as follows:

1) set the grating angle for the desired, approximate, wavelength region;

2) observe and identify neon lines with the array to discover the true spectral range being observed, and record the neon spectrum;

3) substitute the star light for the neon light and commence the run — by monitoring progressive non-destructive readouts, determine the attainment of a satisfactory signal-to-noise ratio, terminate the run, and record the spectrum;

4) substitute the tungsten source for the star light to obtain a flat field spectrum, and record this data.

5.1.1 System check-out

By far the most observed stellar feature during the entire experimental period was H-alpha (6563Å) in alpha Lyrae (α Lyr). The brightness of αLyr (+0.04 magnitude) allows a spectrum to be obtained in a very short time (approximately two minutes), and the strength of H-alpha (H_α) ensures positive identification. Figure 5.1 shows a typical reduced double sample recording (before any off-line data reduction) of H_α in αLyr after an exposure time of 121 seconds.

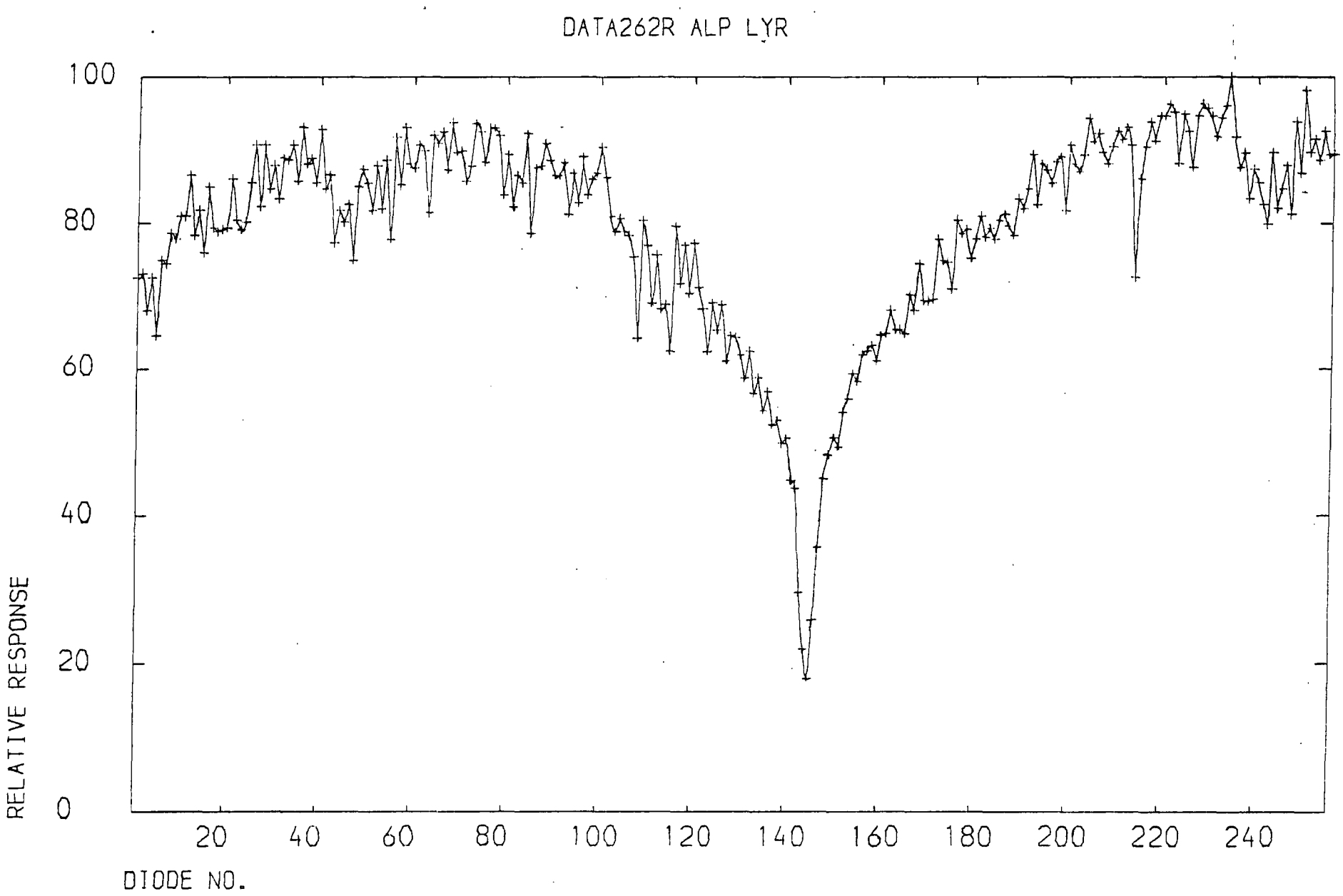


FIG. 5.1 : AN EXAMPLE OF H-ALPHA IN ALPHA LYRAE

When observing stars at H_α that are fainter than αLyr, a system check-out was carried out by first locating αLyr in the sky, and recording it's H_α spectrum. Fine positioning of the detector (see section 3.2) could then be carried out in minimal time to reduce the splitting effect in the final stellar spectrum. Once the optimum conditions were reached for αLyr, the telescope was pointed at the desired star, and the run commenced. In fact, αLyr was the first star looked at each night of observing, to ensure that the equipment was all functioning properly.

5.2 Interpolation of the Data

With the spectral data in the final reduced form (ie odd-even splitting removed and flat field response compensated for), several different analysis methods could be employed. Multiple, shifted spectra were necessary for sampling at the Nyquist rate or better. Therefore, interpolation of the data was carried out to different degrees both to calculate spectral dispersions and to locate the positions of spectral features to within fractions of a diode. Radial velocity and equivalent width measurements could then be determined from the interpolated data.

The method of interpolation follows that of Lanczos (1964). The 256 diode data points are read into a FORTRAN program which carries out the interpolation. One can represent the spectrum by the function f(x), for x = 0 to 255, and the 256 data points by the equidistant samples contained in the set $\{ f_m \}$. For the best results (ie. minimization of the Gibbs effect), it is necessary for the function and the first derivative of the function to have the same values at both the beginning and the end of the chosen period. This can be achieved by selecting the interval of operation as [-k, +k], where f(x) operates only in the positive half. Defining f(x) in the negative half as an even function circumvents the discontinuity problem, since f(-k) = f(k), but the discontinuity in the first derivative

may still exist. This problem is resolved by subtracting a linear function (ax + b) from f(x), with the condition that, if

$$h(x) = f(x) - (ax + b)$$

then $h(0) = h(k) = 0$. Defining h(x) in the negative half of the interval of operation as an odd function, then

$$h(-x) = -h(x)$$

yielding $$h'(-k) = h'(k).$$

This gives continuity in both the function and it's first derivative, and minimizes the Gibbs effect. A final zero value is added to the end of each data set. This brings the total number of data points up to 512, or 2^9, as required for input to the interpolation routine.

The actual interpolation of the data was carried out by using the FORTRAN NAG library subroutine C06AAF, to determine the Fourier transform of the now-512 point data set. Enough zeroes were added to the transform such that, upon re-transforming the data from frequency-space, the desired number of data points exist between the original data. For most purposes, eight-fold interpolation was achieved (ie. seven points interpolated between each pair of original data points), whereas for dispersion measurements, only two-fold interpolation was necessary to give the required accuracy.

After interpolation, one-half of the spectrum is then re-orientated by the addition of the original linear function (ax + b) to restore the features to their correct intensities. The other half of the interpolated spectrum is ignored as it is only an inverted reflection of the spectrum.

Convolution of the interpolated spectra with window functions was found not to have much effect, and therefore not carried out.

Where possible, several stellar spectra were collected with the grating angle altered, by less than 0.1 degree, between exposures. This had the effect of positioning the desired spectral features on different diodes. The interpolated, shifted, spectra were then superposed, and the exposures were averaged. This had the effect of reducing the diode sensitivity noise, and improved the continuum fitting.

5.3 Calculation of Equivalent Widths

A typical stellar line profile is represented by a Voigt function, which is the convolution of a Gaussian profile and a dispersion profile. Most medium strength lines can be approximated by a Gaussian function (Smith and Dominy, 1979), having a strong core and very weak wings. Since these lines are generally not strong enough to have significant wings, the wings can be approximated to triangles (Patchett, et al, 1973).

With the spectral region containing the line of interest interpolated to the desired degree (in this case, eight-fold), two points of inflection are located (I_1 and I_2 in Figure 5.2), one on either side of the line minimum. A search is then made, working in towards the line center, to locate the maximum gradient between adjacent points on each side of the line (indicated by g_1 and g_2). The lines with these gradients are then extended upwards, towards the continuum, to form two right triangles (T_1 and T_2). The area of the spectral line results by adding the areas of T_1 and T_2 to the area between, calculated as the sum of the individual trapezoidal areas. Division of this sum by the height of the continuum yields the equivalent width.

It is best to carry out this determination after a number of slightly off-set stellar spectra have been added together and averaged. This allows

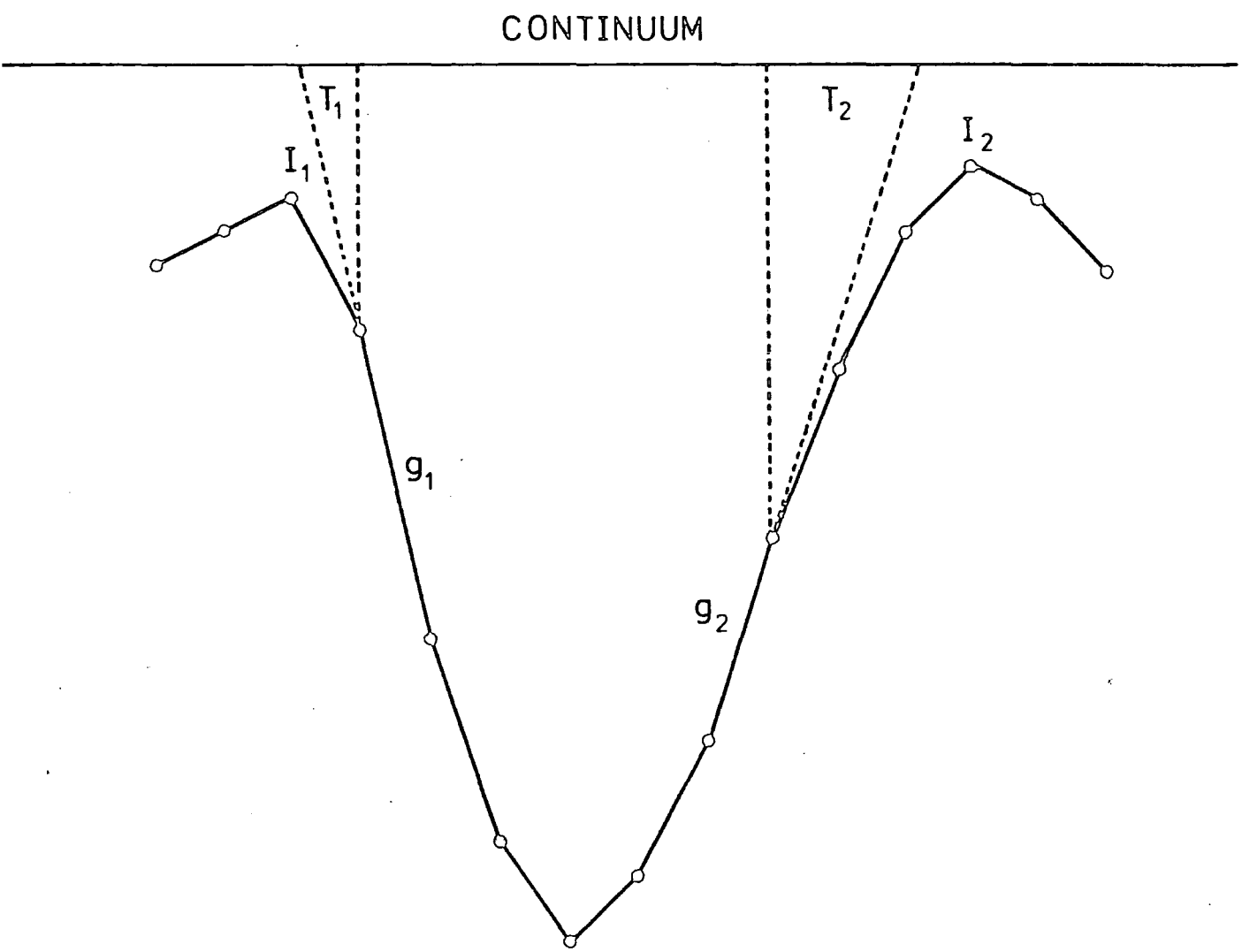


FIG. 5·2: MEASUREMENT OF EQUIVALENT WIDTHS BY THE TRIANGLE-TRAPEZOID METHOD

any diode sensitivity variations to be reduced, and affords a better determination to the continuum level. In these equivalent width measurements, the continuum is defined as the best line drawn through the local high spots on the spectrum (Luck, 1977).

5.4 Detector Dead Space

The 4:1 mark-to-space ratio of the detector array (see Figure 2.2) has been investigated for undesirable effects on the collected spectra. The projected slit width at the detector was 37 μm, being well matched to the diode width of 40 μm. The 37 μm slit width corresponded to one arc second on the sky. For any particular spectral absorption line, the minimum has a finite probability of falling inside the 10 μm wide dead space between adjacent array elements. Also, if the line minimum misses the dead space, but falls near to the edge of a diode, then the shape of the line can still be affected.

This section deals with the possible effects of this 20% dead space on spectral lines. The ideal would have been to scan a narrow (< 10 μm wide) beam of light along the length of the array to observe the variation of response with position. However, at the time the necessary equipment was not available. An idealized computer simulation was carried out to try and investigate this problem. It is important to note that the results obtained here are believed to describe the worst case possible. In reality, the so-called "dead" space cannot be considered truly dead. Charges created by the longer wavelength photons will still tend to be detected (via diffusion through the bulk material) even if the photon misses the sensitive area (see Figure 2.4).

5.4.1 Computer simulation

Computer simulations of the detector array were carried out using

Gaussian curves to represent spectral lines (Kraft et al, 1964, and Smith and Dominy, 1979). The Gaussians were calculated by the formula

$$G(x) = (\sigma\sqrt{2\pi})^{-1} \cdot \exp(-x^2/2\sigma^2)$$

whereby the shape of the line can be completely described by the standard deviation, σ. Different line widths were chosen simply by varying σ in steps between values of 1.0 (ie. 0.2 diode) and 10.0 (ie. 2 diodes). The simulation of the absorption line was realized by subtracting the values of G(x) from a chosen continuum level. Values of x were chosen such that five equi-distant points would cover the detector pitch of 50 μm. The detector simulation computer program plots the shape of the line "seen" by the array, by ignoring every fifth x-value (incident on the dead space), and averaging each remaining group of four x-values. In this way, five different spectra are created for each simulation, with the line minima successively falling on each of the five 10 μm wide bands, and the ignored points progressively stepped through the data set.

Figure 5.3 shows the results of the computer simulation for input lines with standard deviations of 1.0, 2.5, 5.0 and 10.0, labelled a,b,c, and d, respectively. Note that only three examples are indicated for each value of σ. This is due to the symmetry of the simulations, whereby the first and second of the five possible output lines are merely mirror images of the fourth and fifth lines. The right hand side of Figure 5.3 indicates that part of the array upon which the central minimum of the line falls (ie. either between adjacent diodes, on the edge of a diode, or just off the diode centre).

Observation shows that the effect of the line minimum position is only just noticeable in the σ = 10.0 case (d), but for values of σ = 5.0 and less, the effect becomes more pronounced. Table 5.1a shows the actual numbers

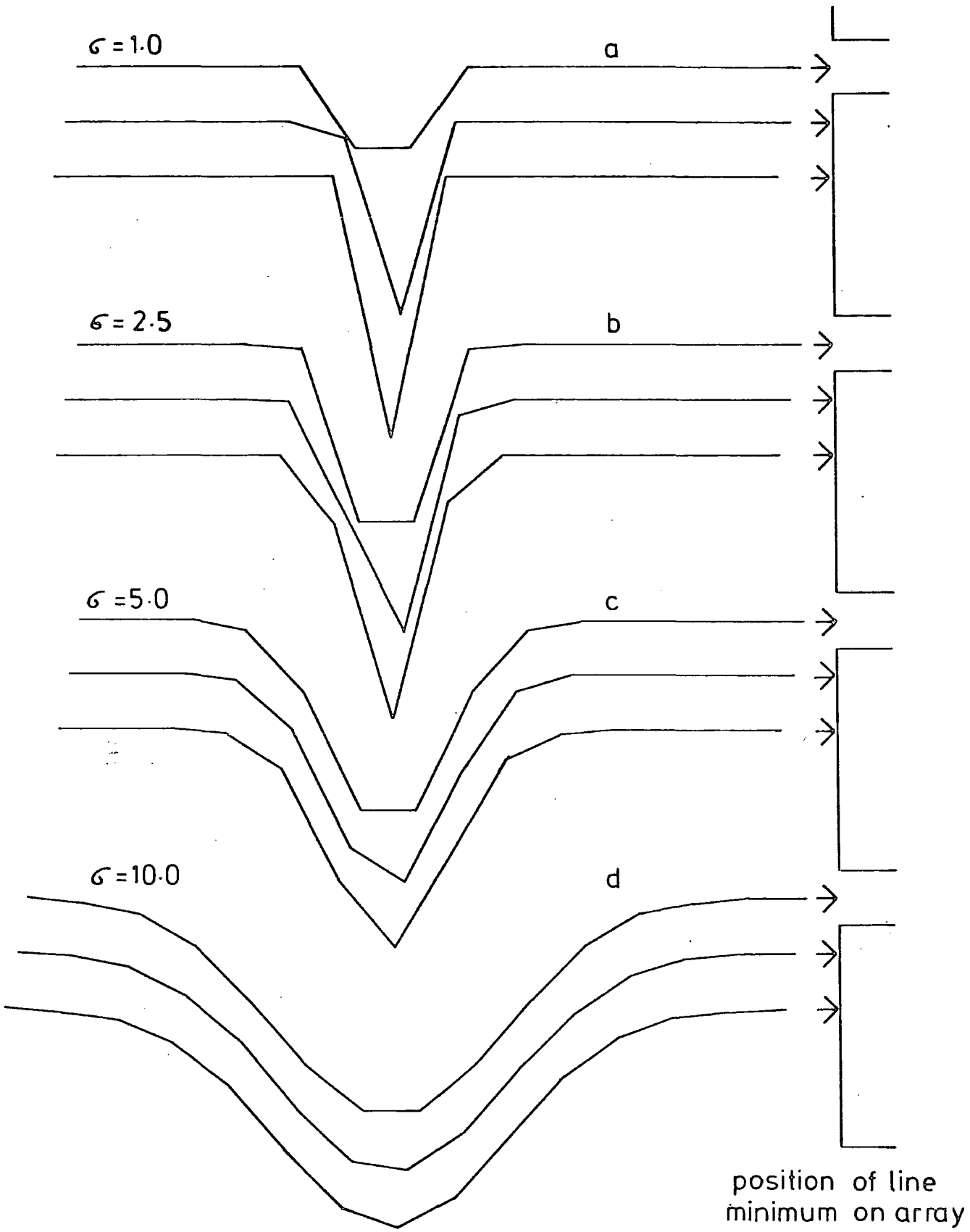


FIG.5·3: SIMULATED DETECTOR OUTPUT LINES

Table 5.1

Measured areas of simulated detector output lines

a

		Line minimum position		
		between diodes	diode edge	diode center
σ	Input	A	B	C
10.0	1.000	0.798	0.797	0.797
5.0	1.000	0.798	0.798	0.798
2.5	1.000	0.797	0.798	0.802
1.0	1.000	0.601	0.758	0.941

b

		Above areas x 5/4			Above lines interpolated		
σ	Input	A'	B'	C'	A(I)	B(I)	C(I)
10.0	1.000	0.997	0.997	0.996	0.931	0.924	0.923
5.0	1.000	0.998	0.998	0.998	0.926	0.925	0.924
2.5	1.000	0.996	0.998	1.002	0.999	0.987	0.995
1.0	1.000	0.751	0.947	1.177	0.768	1.062	1.380

involved in each of these four cases. The input column lists the normalized area of the input Gaussian curve for each value of σ. The next three columns show the measured area of the detected line for each of the three positions of the line centre.

For values of σ of 10.0, 5.0, and 2.5, the measured areas are approximately 20% lower than the input areas. This is due to the simulation program discarding every fifth data point. Multiplying these area values by 1.25 yields the numbers in Table 5.1b, and somewhat compensates for the relatively simplified simulation. However, now it is seen that for one case in the σ= 2.5 line (when the line minimum is nearest the diode centre) the measured area is greater than the input area. This occurs when the bulk of the line is located on the sensitive diode area. The scaling factor of 1.25 is therefore not entirely suitable in this case, when only a small fraction of the line falls on the dead space. This also occurs for the σ =1.0 case, but when the line centre falls between diodes, even the scaling factor cannot compensate for the lost information. It is obvious here that the position of the line on the array is extremely important.

The last three columns of Table 5.1b show the areas of the output lines after the original data (from Table 5.1a) have been interpolated eight-fold, as described in section 5.2. The areas are calculated by the triangle-trapezoid method of Patchett, et al (1973), described in section 5.3. It can be seen that for the wider lines the interpolation of the data can compensate quite well (to within 8%) for the information seemingly lost on the dead spaces. For the line with σ= 1.0, there is too little information for the interpolation program to work on satisfactorily, and errors of up to about 40% can occur. Also, interpolation of very narrow lines results in multiple secondary minima occurring, with corresponding minima being above the continuum level. This will contribute errors to measurements as well.

Inspection of the collected stellar spectra shows that most of the lines of interest have their nearest equivalent to a Gaussian line with $\sigma = 5.0$, and some with even greater values of σ. This means that they will not necessarily have been adversely affected by the detector dead space. Also, the interpolation will tend to constructively smooth out any irregularities in the data.

5.5 Gamma Pegasi

Gamma pegasi (γPeg) is an intrinsically variable star of the beta Cepheid type, with a period of 0.15175 days, a mean magnitude variation of 0.013 (Sareyan, et al, 1976), and a mean radial velocity amplitude of less than 4.5 km.s^{-1} (Ducatel, et al, 1981). Observations of γPeg were carried out in the wavelength range of 6475Å to 6615Å to try to observe the star at H-alpha (6563Å) for one complete period. Due to the lateness of its appearance in the sky, γ Peg was only visible each night for about two hours out of its 3.63 hour period before daylight interfered with observations. Each exposure lasted approximately 20 minutes. Five consecutive runs were collected on the night of 1978 August 14/15 (run numbers 91 to 95), and six runs (numbers 166 to 171) were collected on the night of 1978 August 17/18. Smith and McCall (1978) give the ephemeris

$$RV_{max}\ (JD) = 2434266.854 + n \times 0.15175020$$

for the radial velocity of γ Peg. This means that 77% of the total period was observed (see Figure 5.4).

5.5.1 Radial velocity measurements

The first step in calculating the radial velocity of γ Peg was to determine the wavelengths of particular points on each of the interpolated stellar spectra. Neon calibration runs (see section 5.3.1) yielded, in this case, two reference points on the spectra, namely at 6506Å and 6532Å (see Figure 5.5) Close inspection of selected positions of the spectra with the superpositions of the interpolated neon spectrum enabled the reference wavelength positions to be located. The dispersion and wavelength range of each spectrum was then determined.

The second step was to distinguish between telluric lines and stellar lines. For each of the two sets of consecutive runs, the positions of prominent lines were recorded. Lines that appeared in the same position

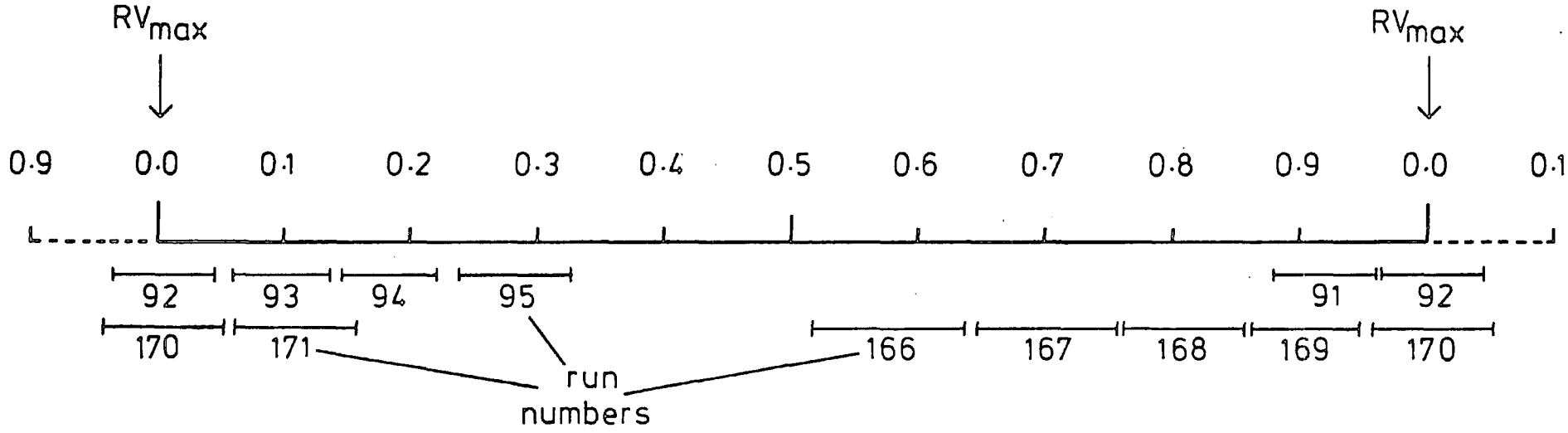


FIG. 5·4: OBSERVED PHASES OF THE RADIAL VELOCITY OF GAMMA PEGASI

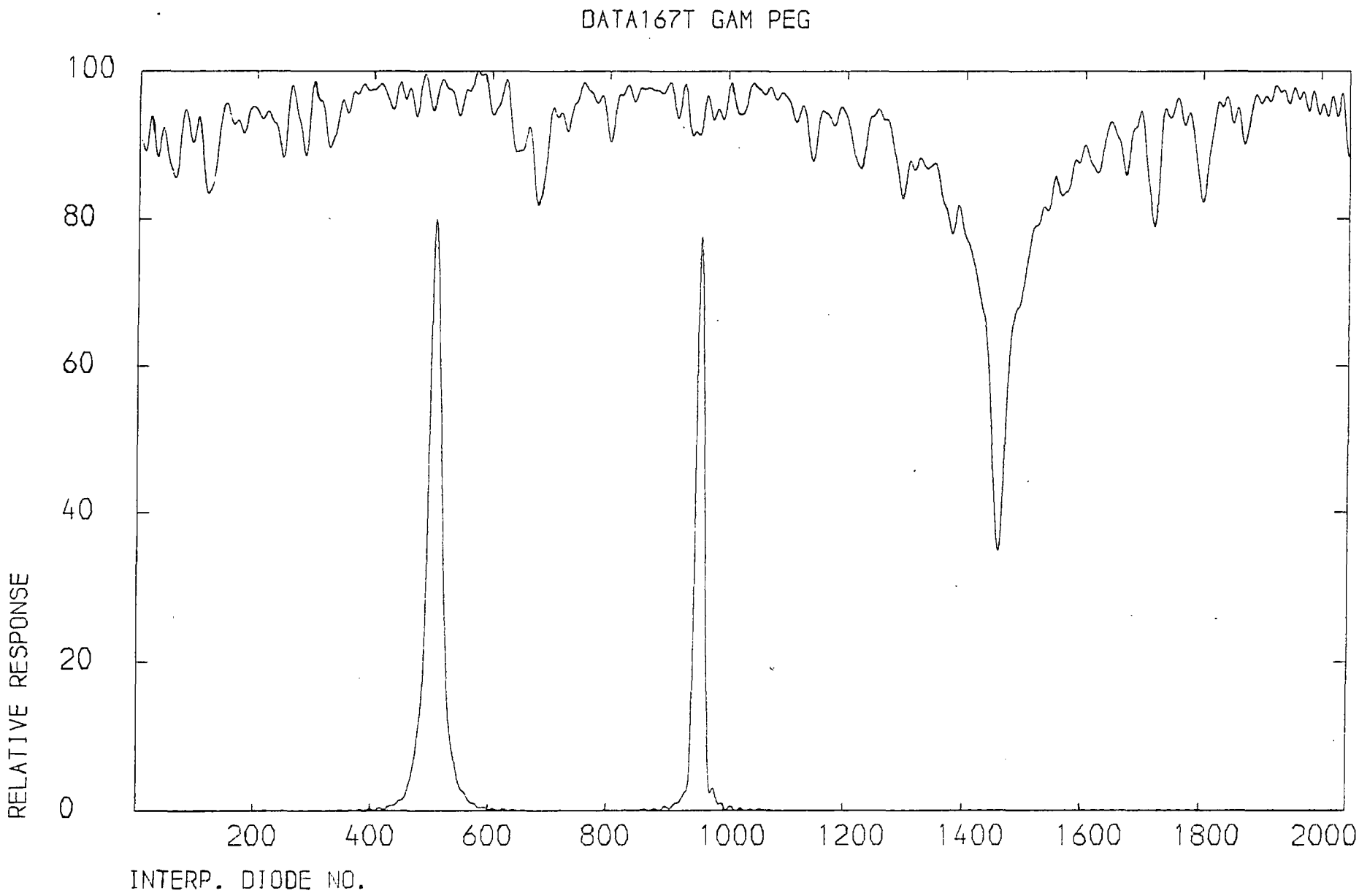


FIG. 5.5 : GAMMA PEGASI SPECTRUM WITH NEON CALIBRATION LINES

in each set were designated telluric lines, since they are formed by absorption in the earth's atmosphere and would be expected to remain fixed. Lines that appeared to move progressively to the left or right (relative to the telluric lines) in successive spectra were assumed to be due to the star, with the apparent movement being due to relative motion of the star along the line of sight. This motion is due to a combination of both the relative movement between the center of the star and the observer, and any expansion or contraction of the observed stellar atmosphere. Using the values calculated for the dispersion and the wavelength range, combined with a list of probable telluric lines, the wavelengths of some of these non-moving lines could be determined. The wavelengths of the stellar spectral lines were then determined by their positions relative to each other, plus the fact that H_{α} is an unmistakeable feature present in these spectra.

The displacement of the stellar lines from their expected positions on the spectrum will give the radial velocity of the star, but only after the correction for the relative motion of the observer has been applied. The method of correction described (Harmer, 1978) permits the determination of the effective positions of the stellar lines relative to positions of the telluric lines.

There are two components to be corrected for, namely the correction for the orbit of the earth around the sun, V_{orb}, and the correction for the diurnal motion due to the rotation of the earth on its axis, V_{di}. If ΔX, ΔY, and ΔZ are the first differences, from the Astronomical Ephemeris, in the rectangular coordinates for the Sun at the time of the observation, then

$$V_{orb} = -1.73 \times 10^{3} \ (l.\Delta X + m.\Delta Y + n.\Delta Z) \quad km.s^{-1}$$

where $l = \cos\delta \cdot \cos\alpha$

$m = \cos\delta \cdot \sin\alpha$

$n = \sin\delta$

and α and δ are the right ascension and declination of the star, respectively. Also,

$$V_{di} = -0.465 \cdot \sin t \cdot \cos \delta \cdot \cos \phi \quad km.s^{-1}$$

where

ϕ = the geocentric latitude of the observatory

t = the hour angle at the time of observing.

The hour angle is determined by

$$t = LST - \alpha$$

where LST is the Local Sidereal Time (in this case Greenwich Sidereal Time), having been converted from the time of the observation (usually measured in Universal Time) using the Astronomical Ephemeris. Finally, the reduced velocity is given by

$$V_{true} = V_{obs} + V_{di} + V_{orb} \quad km.s^{-1}.$$

The calculated variations in the connections for the relative movement of the observer are less than 0.1 $km.s^{-1}$ over the time spanning each of the two sets of observation runs, and 1.2 $km.s^{-1}$ between the first exposure of the first set (run 91) and the last exposure of the second set (run 171). Therefore, the corrections were considered to be negligible over each of the sets of runs, and would not even be detectable between the two sets, since at H_α a variation of 2.7 $km.s^{-1}$ is required to correspond to one interpolated spacing.

As stated before, γ Peg only has a radial velocity variation of less than 9 $km.s^{-1}$ over its period, which is equivalent to only 0.2Å at H_α. The reciprocal dispersion of the detector is approximately 0.47Å . $diode^{-1}$. Therefore, the maximum displacement of a stellar line due to the radial velocity of the star will be less than one half of a diode, or only about $3\frac{1}{2}$ interpolated spacings.

Since a maximum radial velocity change of only 9 $km.s^{-1}$ was being sought, it was realized that the detector array system could not offer sufficient resolution to perform this task. Some movement of the stellar

spectral lines was detected in consecutive runs, but this movement was always in the same direction and of the order of 0.5 to 1 diode spacing overall for each set of runs. The most likely explanation for this varying displacement is movement of the array at the focal plane of the telescope. Movement of this type would arise from thermal instabilities in the cryostat. Since the orientation of the array at the focal plane is such that the line of 256 elements is vertical, slight movement of the array up or down will result in spectral features falling on different array elements. This effect is described more fully in section 5.6. Lines that appeared not to move in successive spectra were probably due to remnants of the odd-even splitting effect (see section 4.3.5) not completely compensated for, and individual diode response variations.

Only one neon spectrum was taken after each of the two sets of stellar runs. In retrospect, a neon spectrum should have been obtained after each and every stellar spectrum in which any line displacements were being investigated. In this way, any short term diode array movement would be detected and could be compensated for.

5.6 Diode Array Instabilities

An investigation was carried out after the observing period in an attempt to understand why the features in some of the stellar spectra which should have been visible were apparently smeared out on very long exposure runs. Also, some lines were seen to experience a displacement along the array during successive runs when this should not have been so (eg. telluric lines). Movement of the diode array inside the cryostat due to thermal instabilities was thought to be the cause of this phenomenon.

Laboratory tests were carried out, with a similar cryostat to the one used on the telescope, to investigate any possible thermal instabilities in the cold finger mounting assembly. The original cryostat described in

section 3.1.1 was used to complete these tests. A hole was cut into the top of the upper cryostat section to allow a dial gauge to rest on the top of the cold finger. As the cryostat could not now be evacuated, suitable insulating material was placed inside the upper body to avoid excessive heat input and reduce icing-up. The cold finger was then cooled by immersing the lower section of the cryostat in liquid nitrogen. Approximately one hour was believed to be sufficient time for the temperature of the cold finger to stabilize before taking any measurements (see Figure 3.7). The coolant level was continually replenished to the maximum during this hour. Recordings of the dial gauge reading were then made (see Figure 5.6) as the coolant evaporated, thereby exposing more and more of the lower cryostat section.

A deflection of less than 40 μm (ie less than the width of one diode) was observed on the dial gauge over the first 60 minutes of observation. However, during the next 30 minutes a sharp increase in movement was detected, equivalent to approximately 160 μm, or more than 3 diode spacings. This was believed to be due to the change in length, with temperature, of the outer, stainless steel, part of the cryostat surrounding the cold finger. Since the bottom of the cold finger is firmly imbedded in the base of this jacket, and the protruding part was always submerged in the coolant, any expansion or contraction of this jacket will result in a vertical movement of the cold finger, and hence, the diode array.

These laboratory measurements were carried out during the day, when the ambient temperature is relatively stable, and therefore the results obtained were considered to be minimized. The effect could be worse while observing, as the ambient temperature at the coudé focus of the telescope varies during the night.

For future observations, the array was to be held firmly inside the upper cryostat section by means of ceramic pillars between it and the inner

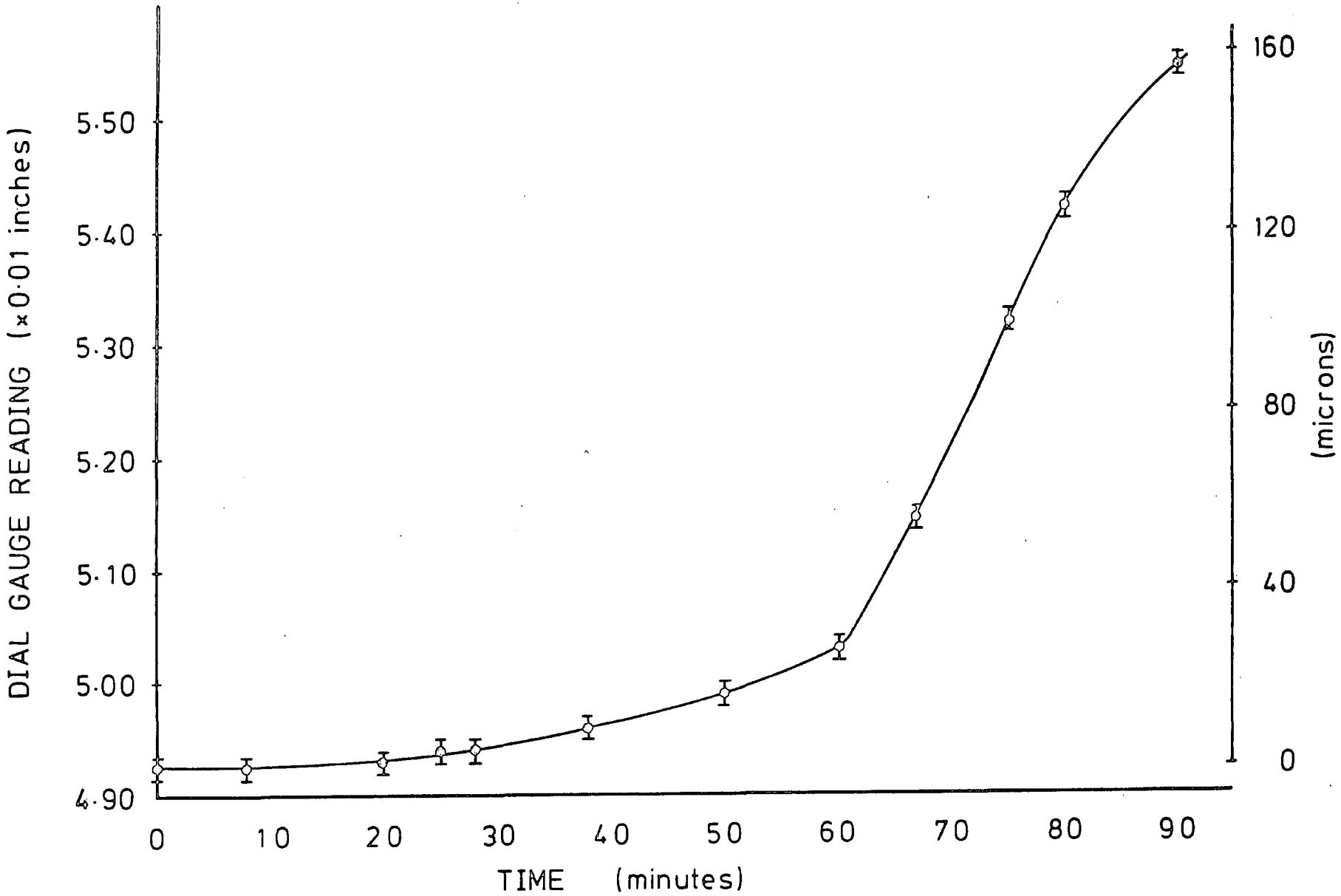


FIG. 5·6: COLD FINGER MOVEMENT AS LN_2 EVAPORATES

wall of the cryostat. Thermal connection of the array to the cold finger was to be obtained by means of copper braid, to absorb any movement of the cold finger itself.

REFERENCES

Ducatel, D., le Contel, J.M., Sareyan, J.P., & Valtier, J.C.: 1981, Astr. Astrophys. Suppl., 43, 359.

Harmer, D.: 1978, personal communication.

Kraft, R.P., Preston, G.W., & Wolff, S.C.: 1964, Astrophys. J., 140, 235.

Lanczos, C.: 1964, "Applied Analysis", ch. 4, Prentice-Hall, New Jersey.

Luck, R.E.: 1977, Astrophys. J., 212, 743.

Patchett, B.E., McCall, A., & Strickland, D.J.: 1973, Mon. Not. R. astr. Soc., 164, 329.

Sareyan, J.P., le Contel, J.M., & Valtier, J.C.: 1976, Astr. Astrophys. Suppl., 25, 129.

Smith, M.A., & Dominy, J.F.: 1979, Astrophys.J., 231, 477.

Smith, M.A., & McCall, M.L.: 1978, Astrophys. J., 221, 861.

CHAPTER SIX

INFRARED STUDIES OF NEUTRAL OXYGEN

In the field of astronomy, the information obtained from the violet end of the visible spectrum can be used to determine the spectral types of stars. However, infrared observations are important when gathering data about stellar luminosity, and they can also give insight into temperature, magnitude, and mass loss (if present). In addition, interstellar absorption is low in the near infrared, thereby facilitating the study of galactic structure. Unfortunately, the investigation of stellar features in the infrared and near infrared regions of the spectrum has been somewhat hindered by the lack of suitable detectors that can operate efficiently in this low photon-energy range. With the relatively recent development of semi-conductor materials for use as solid state detectors, a great store of new information has been (and will continue to be) made available to astronomers and astrophysicists.

One particularly interesting area of astronomical research is concerned with the observation and measurement of stellar absorption lines due to neutral oxygen (OI) atoms. OI transitions between excited states exist within a range of wavelengths from 1300Å to well over 1μm. The strongest unblended OI lines (except possibly in late F-type or early G-type stars) with wavelengths below 1μm are at 7774Å and 8446Å. The feature at λ7774 is actually a blend of three lines, as is the λ8446 feature. However, the components of the λ8446 triplet are separated by less than 0.4Å, so they cannot be resolved in any except very high resolution spectrometry.

6.1 The oxygen multiplet of 7774Å

The neutral oxygen feature at λ7774 was first observed in stellar spectra by Merrill (1925), and has since been recognized as being a resolvable triplet. The line centers of the three components are at 7771.94, 7774.17, and 7775.39Å. The ratio of equivalent widths of the components is very nearly 1:1:1 in supergiants, since the lines are situated on the flat portion of the curve of growth.

λ7774 is one of the most prominent stellar features available both for separating supergiants from other classes of stars, and for determining absolute stellar magnitudes. Due to the high excitation potential (9.15 eV) of the lower energy level of transition ($3s^5S$), λ7774 is also very temperature sensitive. It occurs in an otherwise clear region of the spectrum, remaining relatively free from blending with either telluric or other stellar lines (λ8446 can be contaminated with blends of FeI lines). It appears much stronger than λ8446, with an absorption strength ratio of λ8446:λ7774 in the region of 0.6 (Keenan and Hynek, 1950). Even low dispersion equipment (down to 200 $Å.mm^{-1}$) can be used for observations (Parsons, 1964).

6.1.1 An alternative photometric index

Equivalent widths are not the only measurements that can be made on spectral lines. A photometric index called r(01) can be used to determine the absolute magnitudes of F-type supergiants. The derivation of this index is described fully in Sorvari (1973). The value of r(01) can be approximately expressed as a function of the equivalent width of λ7774 by

$$r(01) = 0.0741 \times W(7774)_{KH} + 1.304 \qquad 6.1$$

where $W(7774)_{KH}$ is the measurement of the equivalent widths of $\lambda 7774$ (in Å) taken from Keenan and Hynek (1950), hereafter denoted KH. Equation 6.1 only holds true provided that $W(7774)_{KH}$ is less than 1.2Å. To overcome this upper limit on the equivalent widths, the utilization of a more extensive set of data in this formula yields an expression for the photometric index of

$$r(OI) = 0.0855 \times W(7774)_o + 1.309 \qquad 6.2$$

where $W(7774)_o$ is the equivalent width of $\lambda 7774$ taken from the data of Osmer (1972a).

6.2 $\lambda 7774$ as an indicator of temperature

The intensities of the OI lines vary with stellar temperature. Beginning with the hotter, early B-type stars (where W(7774) is negligible), the OI line strengths increase with decreasing temperature until a maximum value is reached in the A-type stars (see Figure 6.1). From here, the intensities decrease almost linearly through the F-type stars, and, except in supergiants, the lines disappear in the G-type stars (KH). As an example, for late-type giants experiencing local thermodynamic equilibrium, the temperature dependence is given by

$$\frac{d(\ln N_1)}{d(\ln T)} = \frac{X}{kT}$$

(from Eriksson and Toft, 1979), where

N_1 = population of the lower level

X = excitation potential

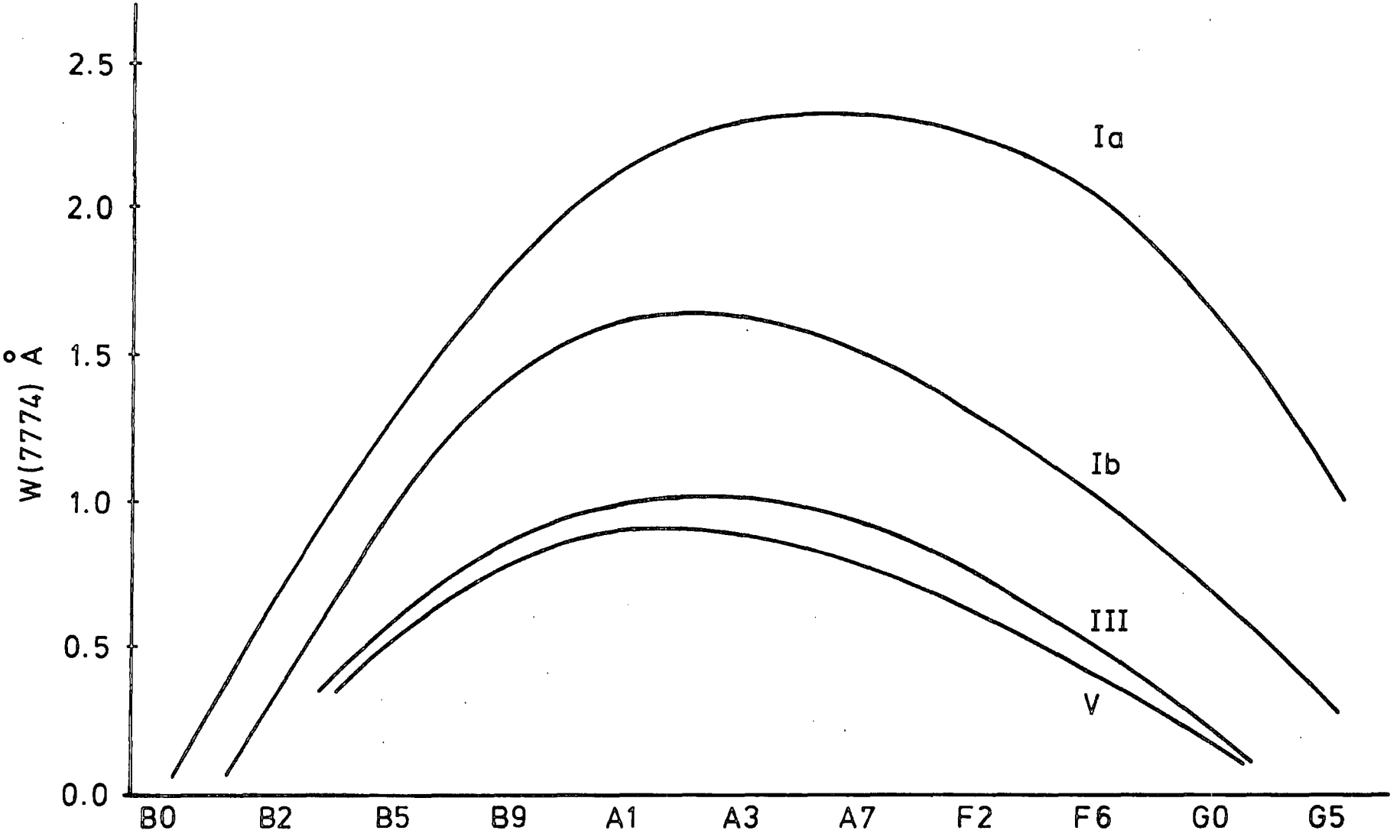


FIG. 6.1: LUMINOSITY LEVELS FOR λ7774 (from KEENAN & HYNEK, 1950)

T = temperature

k = Boltzmann's constant.

So for a change in stellar temperature of only 3%, the lower level population, N_1, will change by a factor of 2.

The relationship between W(7774) and temperature is best explained by separating A-type stars from F and G-type stars. Consider first the latter of these two groups.

In cooler G-type stars, the negative hydrogen ion is the dominant source of continuous opacity. The simple proton-electron combination of the hydrogen atom is highly polarized, and can combine with a free electron, emitting a photon of 0.754 eV.

$$H + e^- \longrightarrow H^- + h\nu$$

These extra electrons are supplied by the ionized metals. With plenty of negative hydrogen ions present, the continuous opacity of the star will be high, and the lines will appear weak.

The negative hydrogen ion can be ionized to a neutral hydrogen atom, plus a free electron, by any photon having a wavelength less than 16,450 Å.

$$H^- + h\nu \longrightarrow H + e^-$$

In the hotter F-type and early G-type stars, this process takes place. The opacity due to the negative hydrogen ion decreases, which reveals itself as an increase in the line strengths. Also, the increased excitation with temperature will strengthen the lines.

In the former group, in particular the B8 to A6 stars, the variation

of W(7774) with temperature can be explained in simple terms as an energy level population effect. The population of the lower, $3s^5S$ level is an inverse function of temperature. An elevation of temperature results in a greater electron pressure, which, in turn, creates an increase in the continuous opacity of H^-. As the temperatures of the stars decrease progressively through spectral type, the lower level populations increase, giving rise to stronger values of W(7774).

Consequently, due to the existence of these different temperature effects, it is best to separate A-type stars from F and G-types when correlating between absolute magnitude or luminosity and λ7774 line strengths.

Rao and Mallik (1978) have discovered that the strength of the FeI line at λ7748 is also a function of stellar temperature. The line first appears in F5 stars and gets stronger in later types. Since W(7748) increases with decreasing temperature, the product

$$E = W(7748) \times W(7774) \qquad 6.3$$

should be sensitive only to luminosity in late-type stars. A relationship between E and the absolute visual magnitude, M_V, was deduced by Rao and Mallik for these stars, and can be expressed as

$$M_V = -10.98 \times E - 1.79 \qquad 6.4$$

with an accuracy of M_V to be $\pm$ 0.5 magnitudes.

6.3 λ7774 as an indicator of absolute magnitude

Many authors have discovered the relationship between the strength of λ7774 and the absolute magnitudes of stars, among them KH, Osmer (1972a) and Sorvari(1973, 1974). The determination of the absolute magnitudes of stars in nearby galaxies is of great importance to astronomers, as it allows these stars to be used as distance indicators.

The maximum luminosity of stars in our galaxy is believed to be approximately 10^6 times that of our own sun. This value corresponds to an absolute magnitude of about -9. If stellar characteristics are identical (or nearly identical) throughout this region of space, then one would expect this to be an upper limit on the luminosity of stars in other galaxies as well.

Among the most luminous stars known are the F-type supergiants. For these stars, Osmer (1972a) has determined a close linear relationship between the equivalent width of λ7774 and absolute visual magnitude, in the magnitude range of -9 to -4. This relationship can be expressed by the equation

$$M_V = -2.62 \times W(7774)_o - 2.55 \qquad 6.5$$

where $W(7774)_o$ is the total equivalent width of all three components of λ7774, measured in Å. Between the limits stated, Osmer's results do not depart from equation 6.5 by more than ± 0.5 magnitudes. Baker (1974) has also derived a relationship for F-type supergiants, given by

$$M_V = \underset{(\pm 0.228)}{-2.711} \times W(7774)_B \underset{(\pm 0.379)}{-2.472} \qquad 6.6$$

which agrees well with Osmer, within the errors given.

For supergiants, Sorvari (1974) has also deduced a relationship between M_v and r(01), namely,

$$M_v = -36.61 \times r(01) + 46.46 \qquad 6.7$$

This can be combined with equation 6.1, using the data of KH, to yield

$$M_v = -2.713 \times W(7774)_{KH} - 1.279 \qquad 6.8$$

which, as stated before, is only believed reliable for equivalent widths of up to 1.2 Å. However, using Osmer's data and combining equations 6.2 and 6.7, one gets

$$M_v = -3.42 \times W(7774)_o - 0.96 \qquad 6.9$$

Once the absolute magnitude of a star is known, it's distance can be determined by measuring the apparent visual magnitude, m_v, and applying the formula

$$m_v - M_v = 5 \log_{10} d - 5$$

where d is the distance in parsecs.

Figure 6.2 shows a plot of equations 6.5, 6.6, 6.8, and 6.9, in the magnitude range of $-2 < M_v < -9$. Agreement between all four sets of results is reasonably good. Sorvari's equation using Osmer's data exceeds Osmer's $\pm$ 0.5 magnitude error bar for stars fainter than about $M_v = -6$. The equation Sorvari uses with the data of KH has a similar slope to both Osmer's and Baker's results, but lies approximately 1.5 magnitudes brighter.

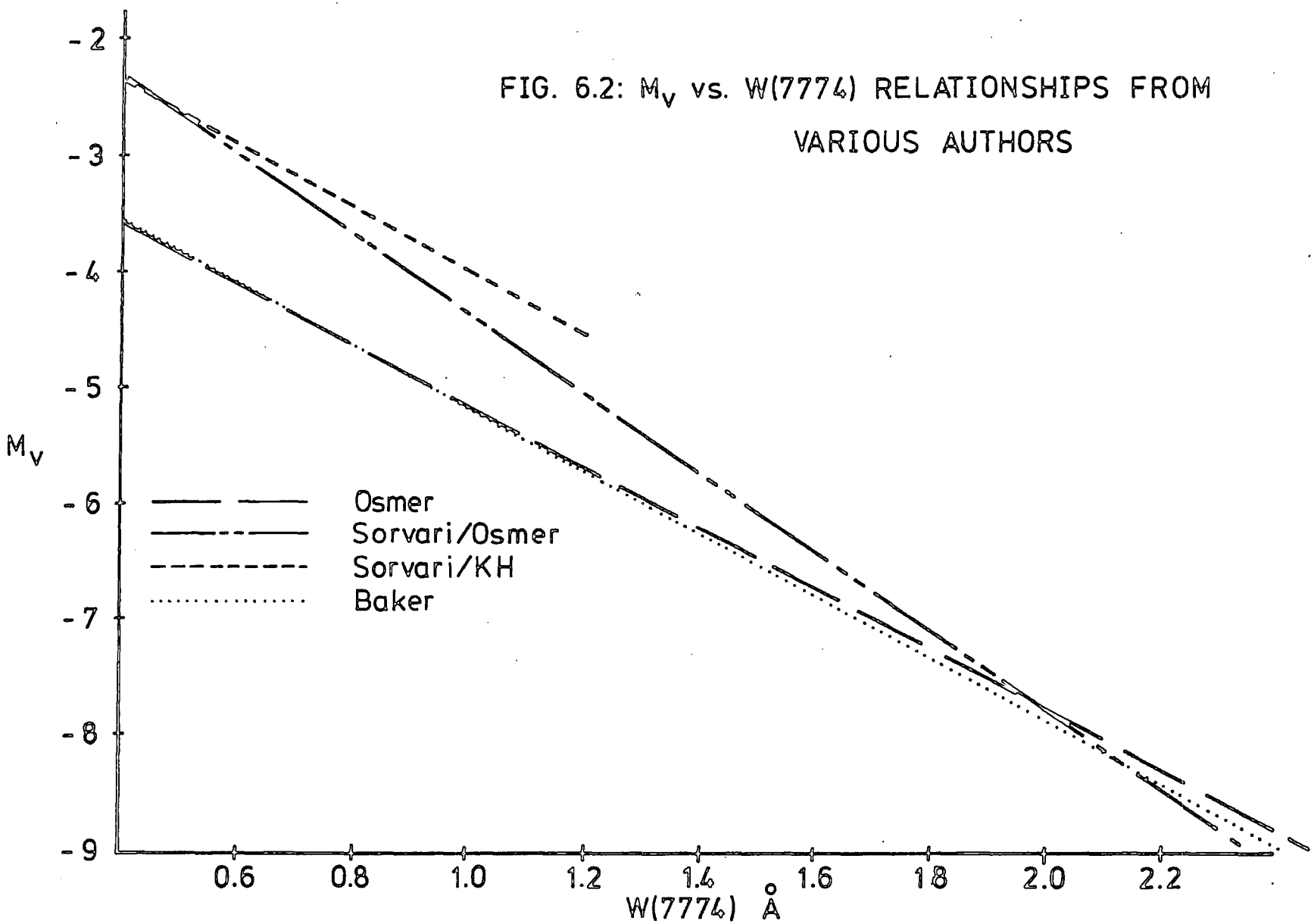


FIG. 6.2: M_V vs. W(7774) RELATIONSHIPS FROM VARIOUS AUTHORS

The actual values taken for M_V could contribute to these observed differences, along with variations in equivalent width determinations resulting from different instruments and observers (Griffin, 1969). Comparisons to the results obtained in this work are given in section 6.6 of this chapter.

6.4 λ7774 as an indicator of luminosity

The λ7774 OI feature is also a reliable criterion for the assignation of luminosity classes. For stars in the spectral range of B5 to G2, many authors have shown that the strength of λ7774 can be used to separate supergiants from other stars. Osmer (1972b) also found that plotting W(7774) against spectral type distinguishes between luminosity classes within the -4 to -9 magnitude range for spectral types A0 to G2 (see Figure 6.3). Class Ia stars are easily distinguished from the rest, having stronger λ7774 lines throughout the spectral range. Class Ib stars are intermediate, with line strengths decreasing with decreasing temperature. Classes II to V all tend to lie on the lower section of the graph, with W(7774) staying below 1Å for all spectral types in this range. The last two groups tend to blend together after spectral type F5, with the Ia stars remaining higher throughout.

The easy separation of supergiants from other stars is not entirely surprising. Supergiants form a relatively homogenous group of stars with reference to their age and composition. Since the stronger spectral lines lie on either the flat or the square root regions of the curve of growth, abundance differences will not seriously affect the line strengths. The OI lines are strong because of the low surface gravity of supergiants. However, no explanation has yet been found as to why there is such a large difference in the W(7774) values between Ia and Ib stars. Note

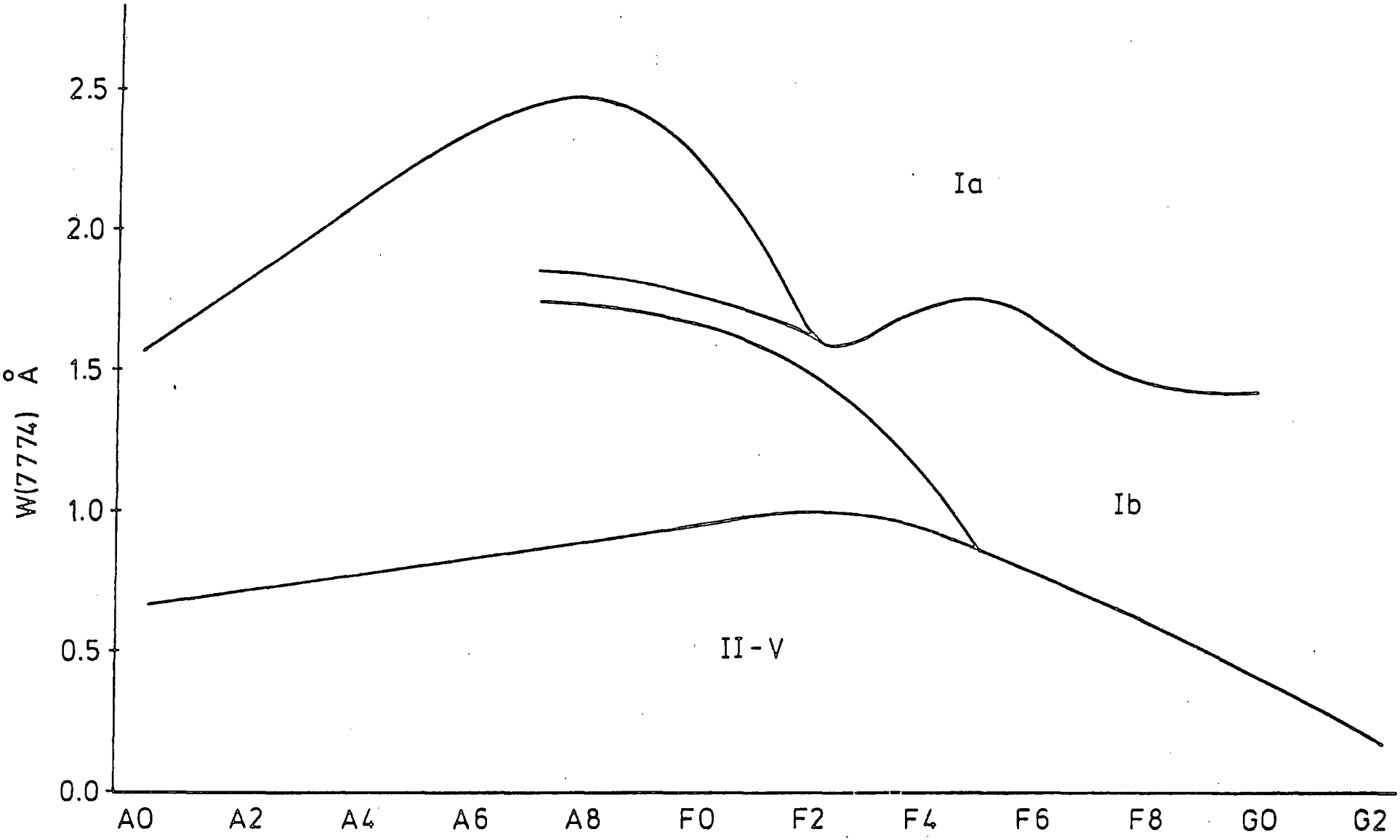


FIG. 6.3: STRENGTH OF W(7774) WITH SPECTRAL TYPE

that in stars later than F5, W(7774) is also a strong function of temperature (see section 6.2 of this chapter). This temperature dependence must therefore be taken into consideration when determining the luminosities of late-type stars from the strength of W(7774).

Sorvari's index, r(01), can also be used as a supergiant discriminant (Sorvari, 1974). In B8 to G3 stars, a plot of r(01) against spectral type effectively separates luminosity classes Ia and Ib from classes II to V (see Figure 6.4), and from each other after type A6. In the A-type stars, r(01) is a strong function of temperature - note the relatively uniform increase of r(01) with spectral type on the left hand side of the diagram. In supergiants of this spectral type, r(01) seems to be unaffected by either microturbulance (W(7774) increases with microturbulance) or absolute magnitude, and therefore cannot differentiate between Ia and Ib luminosity classes.

In the A7 to G3 spectral range, r(01) decreases with spectral type (as the stars become cooler), thereby allowing r(01) to strongly indicate the absolute magnitude. However, the Ia stars are still seen to exhibit much higher values of r(01) than the Ib to V classes (shown on the right hand side of Figure 6.4), with few normal stars in between.

The relationship between W(7774) and luminosity is explained by Thomas, et al. (1979). Referring to the Grotrian diagram in Figure 6.5, the lower level of the 7774 transition is the $3s^5S$ level, and the upper is $3p^5P$. The lower level is metastable, and only has a significant lifetime in the relative absence of collisional de-excitation. This condition exists in large diameter stars of low surface gravity, having low density atmospheres and, therefore, high luminosities (eg. supergiants). A decrease in the strength of W(7774) will therefore result as the luminosities of the stars decrease.

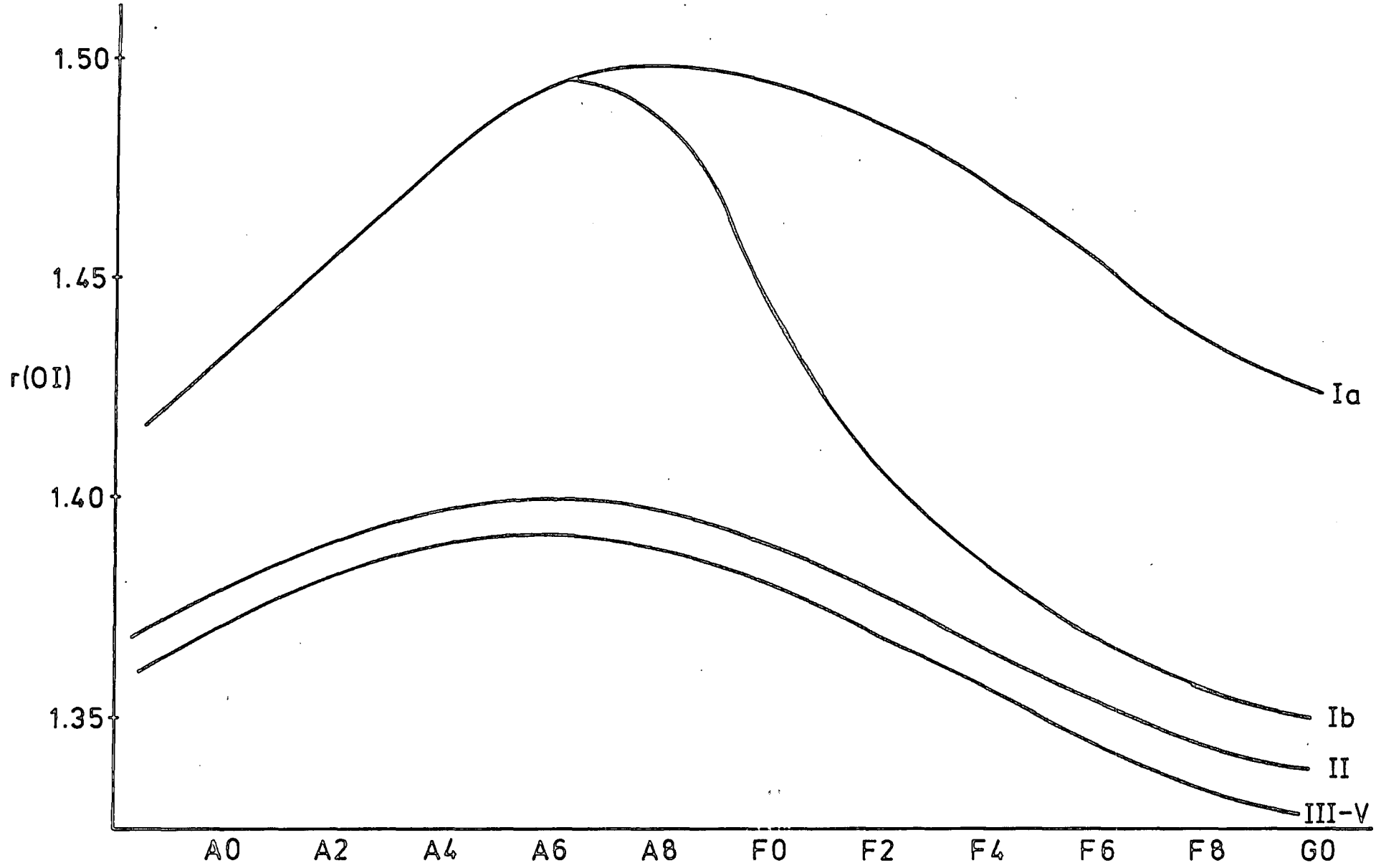


FIG. 6.4: VARIATION OF r(OI) WITH SPECTRAL TYPE AND LUMINOSITY

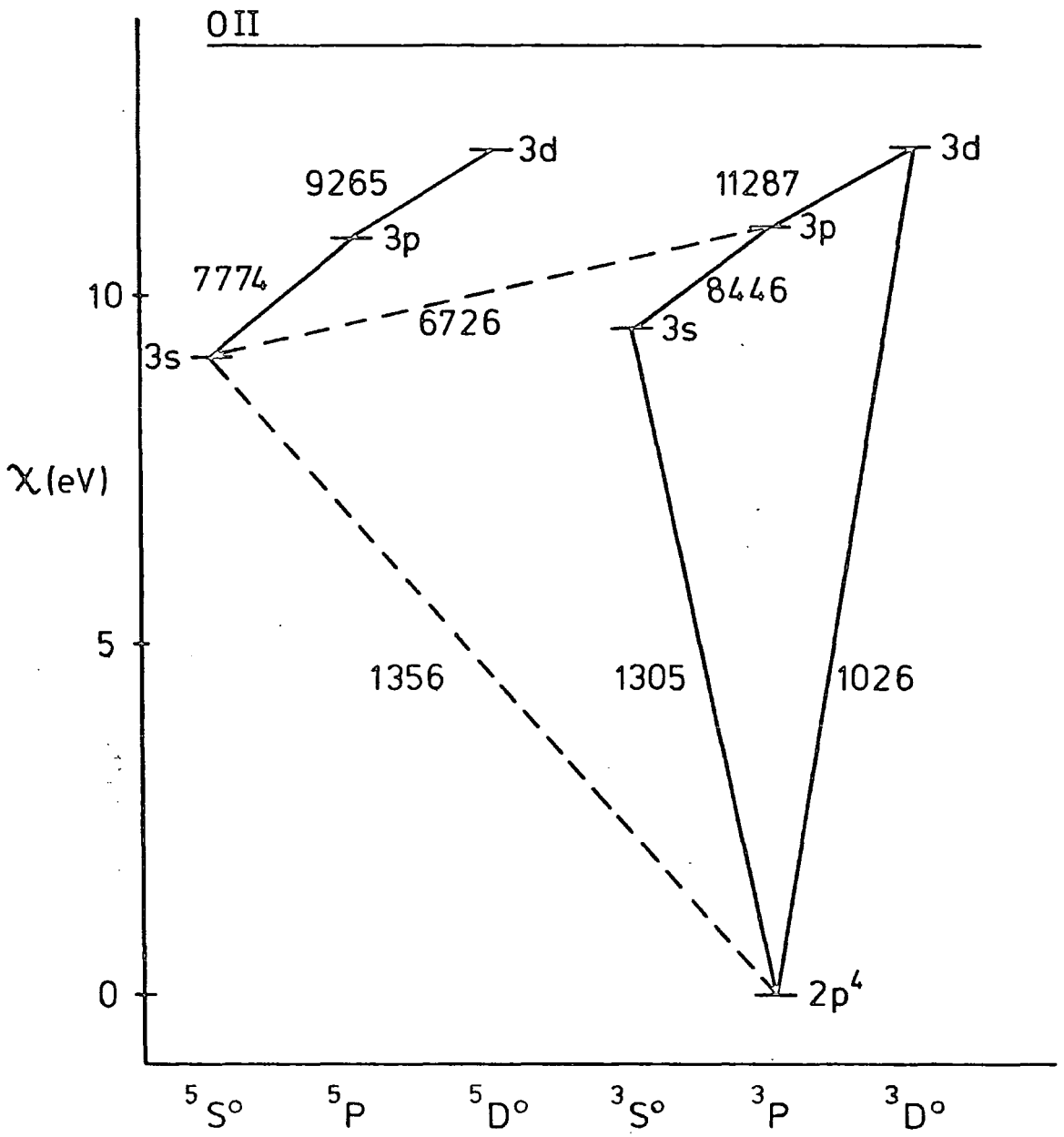


FIG. 6.5: GROTRIAN DIAGRAM FOR NEUTRAL OXYGEN

6.5 Apparent oxygen over-abundance

Several authors have observed an unusually deep absorption of the λ7774 line (eg. Houziaux, 1969), implying an over-abundance of oxygen. However, this can be explained as resulting from a departure from local thermodynamic equilibrium (LTE). In the situation where LTE exists, any oxygen abundance will only have a weak effect on W(7774) because of the strong temperature dependence. A decrease in the oxygen abundance would result in the line becoming weaker. The line is therefore being formed in the deeper layers of the star where the temperature is higher. But if the temperature where the line is formed is higher, then the lower level population is increased, and the line is strengthened.

Non-LTE effects can give rise to the over-population of the metastable $3s^5S$ level. This occurs in the outer layers of the stellar atmospheres where the density is low (Johnson, et al, 1974)., and there is a lack of collisional de-excitation. Non-LTE effects are enhanced where the surface gravity is low (Eriksson and Toft, 1979), thereby allowing fewer collisions which increases non-LTE. Fewer collisions mean a smaller electron pressure, and, in cooler stars, less H^-. Thus, the continuous opacity is decreased which seems to strengthen the lines. The effect is especially conspicuous in the extended atmospheres of supergiants, where the low density results in high luminosities.

In the LTE regime, the depth of the line core cannot go below a certain level. This level is defined by the surface temperature, with the star behaving as a blackbody radiator (Gray, 1976). The core of the line is formed in the outer layers of the star, and the wings formed in the deeper layers. Looking deeper into the star, where LTE dominates, serves to raise the line core above the minimum level. When non-LTE effects are in evidence, the entire line is formed in the outer,

low density, layers of the star. Thus, the core is formed further out and the line minimum, set by the Planck function for LTE, is disregarded.

Although the wings of the spectral lines are not noticeably affected by non-LTE effects, the deepening of the line cores can cause the equivalent widths to be increased by a factor of two to three. In addition, microturbulent effects serve to raise the flat part of the curve of growth. The lines of the F-type supergiants lie on this part of the curve of growth, so microturbulence would give the appearance of an increase in the equivalent widths of these stars. However, in the mid-B to early A-type supergiants, the OI lines lie on the square root portion of the curve of growth, and are therefore insensitive to microturbulence.

6.6 Description of results

The results of the observations of the OI triplet at λ7774 obtained with the Plessey linear diode array are shown in Table 6.1 and the reduced spectra are illustrated in Figure 6.6. Only six stars were found to have spectra from which any astronomical inferences could be drawn. The remaining stars observed at λ7774 had features which were almost completely washed out due to thermally induced geometrical instabilities in the cold finger of the cryostat over very long integration times ($\gtrsim$ 60 minutes) (see Chapter 5). Of the six stars analyzed, only the data from νPer consisted of a single exposure. The remainder individually consisted of results averaged over three spectra, each of the three exposures being slightly shifted along the array of detector elements. As described in Chapter 5, this technique serves to reduce responsivity variations and improve on the Nyquist rate of sampling.

The first column of Table 6.1 lists the names of the stars, with

Table 6.1

Equivalent widths (in Å) of observed stars compared with other measurements

Name	n	W(7774) meas	W(7774)				$W(7774)_{LC}$	W(7748)	Sp
			KH	O	B	RM			
α Cyg	3	2.27	2.19	>2	-	-	2.18	~0	A2Ia
ε Aur	3	1.94	2.34	2.51	-	-	1.85	~0	A8Ia
α Per	3	1.26	1.15	0.95	1.05	1.16	1.14	0.18	F5Ib
γ Cyg	3	1.37	1.16	1.18	1.15	-	1.26	0.24	F8Ib
β Aqr	3	0.60	0.54	-	-	-	0.47	0.32	G0Ib
ν Per	1	0.70	-	-	-	-	0.58	0.10	F5:I

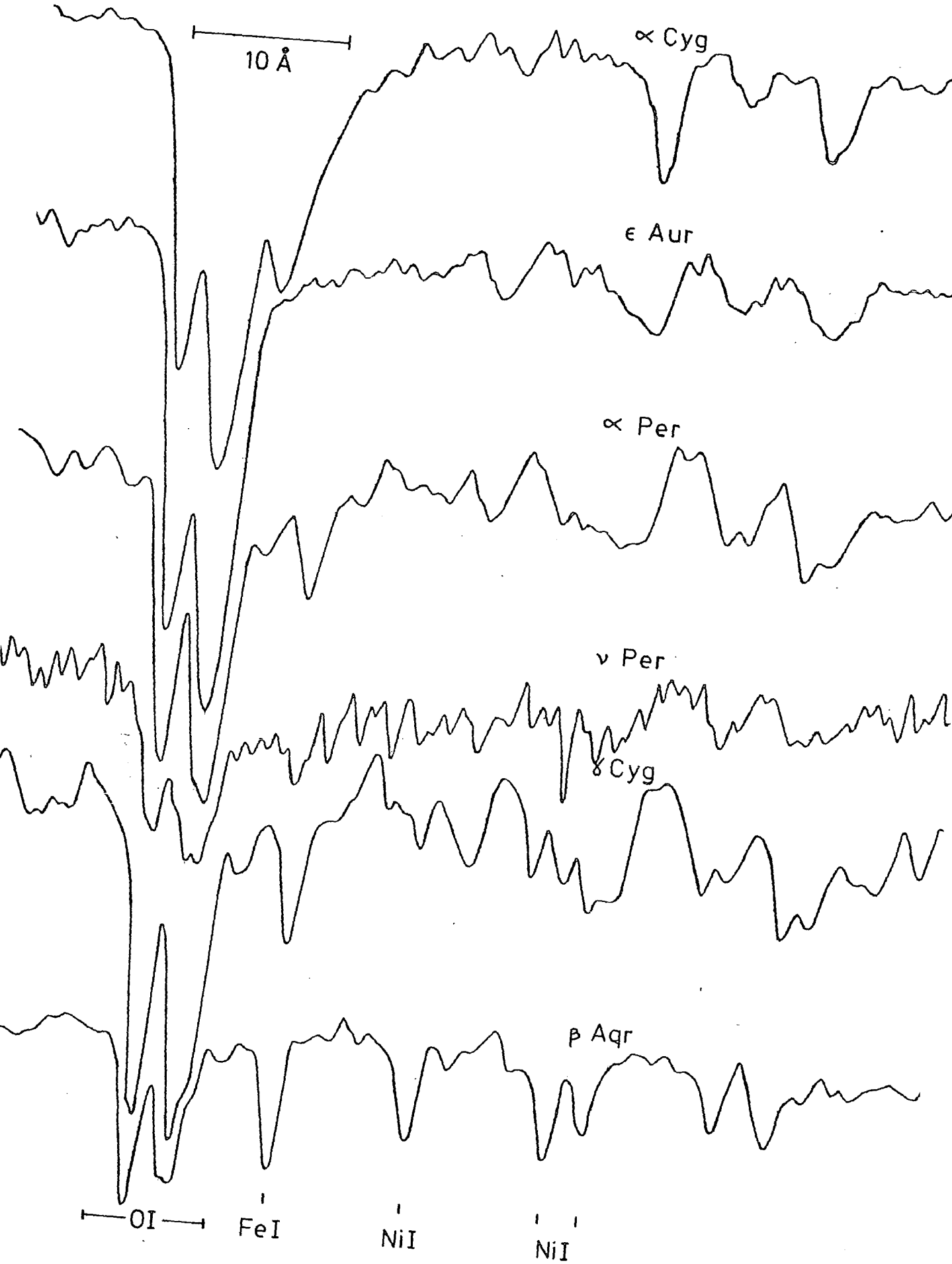


FIG. 6.6: COMPARISON OF OI(λ7774) LINE STRENGTHS FOR OBSERVED STARS

the number of exposures in column two. Column three shows the equivalent width of the λ7774 triplet, measured by the triangle-trapezoid method described in Chapter 5. Columns four to seven indicate, for comparison, the values of W(7774) obtained by KH, Osmer, Baker, and Rao and Mallik, respectively. Column eight shows the calculated equivalent width of λ7774 for a continuum level 3% lower than that in column three. Column nine lists the measured equivalent width of the FeI line of λ7748, and column ten the spectral type.

Inspection of columns three to seven shows that the results calculated are, with one exception, consistantly higher than the previously published results of the other authors. According to Osmer (1972a), variations in equivalent widths measured using different spectrometers can be as high as 20%. In some cases, a factor of three discrepancy in the equivalent widths can be found (Minnaert, 1934). These differences can be attributed to scattered light, resolving power, continuum fitting etc. With the exception of the result for αPer obtained by Osmer, all of the comparison values are within 18% higher than those measured. For the determination of equivalent widths, the continuum was drawn at a height such that it best intersected the local high points of the spectrum near the OI triplet (Luck, 1977). Since the equivalent widths were calculated from averaged sets of data (except for νPer), and the continuum was only being fit to a waveband of about 15Å, this method was considered sufficiently accurate. However, an important factor in the determination of equivalent widths is, obviously, the level at which the continuum is drawn on the data to be analyzed. This dependence is reflected by the fact that a continuum level drawn 3% lower in the data collected alters the variations from the published values significantly. Again, except for Osmer's result on αPer, the measured values now lie mainly within 9% of the published values.

The spectra of εAur were recorded at a time when considerable drift was being experienced by the detector system. As the presence of drift serves only to impart a DC-type level to the output, the resultant effect is a reduction in the calculated equivalent widths. This could explain why the data from εAur alone gave a lower value of W(7774) compared with the other, published results.

Table 6.2 shows the calculated values of absolute visual magnitude using the equivalent widths measured and the equations illustrated in Figure 6.2. Columns one and two give the star and the measured value of the λ7774 equivalent width. Columns three to six indicate the calculated M_V values from the measured data and the equations of Osmer (3), Baker (4), Sorvari/KH (5), and Sorvari/Osmer (6). Column seven lists the values of absolute magnitudes from the references stated.

Table 6.3 is presented in an identical fashion to Table 6.2, but here the measured W(7774) values are those calculated with a continuum level 3% lower than above. Figure 6.7 shows the relationship between the published absolute magnitudes of the stars and the measured W(7774) values, with a least squares line fitted to the data. Also shown is the relationship with the 3% lower continuum level. Putting the measured values of W(7774) back into each of the respective equations for the lines drawn through the data yields absolute magnitudes shown in column eight of Tables 6.2 and 6.3. The similarity of these columns indicates that the lowering of the continuum does not have that strong an effect after all on the determination of absolute magnitude. Therefore, for the stars investigated, a good fit to the data is given by

$$M_V = 2.58 \times W(7774) - 1.81 \qquad 6.10$$

Table 6.2

Calculated values of M_V from measured values of W(7774)

Name	W(7774) meas	M_V				M_V (ref)	M_V (calc)
		O	B	S/KH	S/O		
∝ Cyg	2.27	-8.5	-8.6	-7.4	-8.7	-7.5 a	-7.5
ε Aur	1.94	-7.6	-7.7	-6.5	-7.6	-7.0 b	-6.7
∝ Per	1.26	-5.9	-5.9	-4.7	-5.3	-4.5 a,c -4.6 d,e -5.3 f	-4.9
ɣ Cyg	1.37	-6.1	-6.2	-5.0	-5.6	-5.6 f -4.7 g	-5.2
β Aqr	0.60	-4.1	-4.1	-2.9	-3.0	-4.5 h,i,j	-3.2
υ Per	0.70	-4.4	-4.4	-3.2	-3.4	-2.2 g	-3.4

a: Abt (1957)
b: Hack & Selvelli (1979)
c: Sorvari (1974)
d: Osmer (1972a)
e: Rao & Mallik (1978)
f: Baker (1974)
g: Allen (1976)
h: Kraft et al (1964)
i: Allen (1973)
j: Engold & Rygh (1978)

Table 6.3

As above but with 3% lower continuum level

Name	W(7774) meas	M_V				M_V (ref)	M_V (calc)
		O	B	S/KH	S/O		
∝ Cyg	2.18	-8.3	-8.4	-7.2	-8.4	-7.5	-7.5
ε Aur	1.85	-7.4	-7.5	-6.3	-7.3	-7.0	-6.7
∝ Per	1.14	-5.5	-5.6	-4.4	-4.9	-4.5 -4.6 -5.3	-4.9
ɣ Cyg	1.26	-5.9	-5.9	-4.7	-5.3	-5.6 -4.7	-5.2
β Aqr	0.47	-3.8	-3.7	-2.6	-2.6	-4.5	-3.2
υ Per	0.58	-4.1	-4.0	-2.9	-2.9	-2.2	-3.5

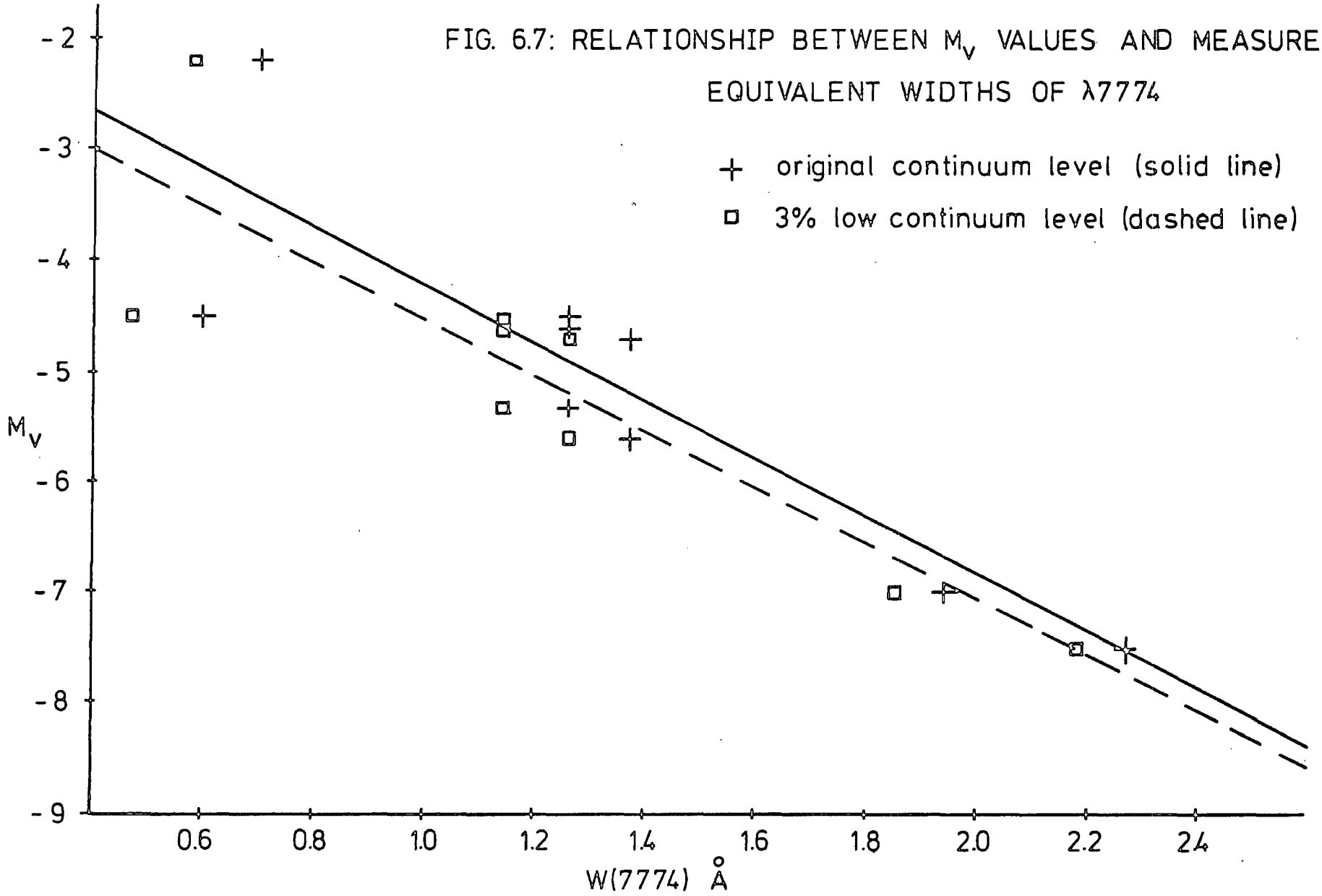


FIG. 6.7: RELATIONSHIP BETWEEN M_V VALUES AND MEASURED EQUIVALENT WIDTHS OF λ7774
original continuum level (solid line)
3% low continuum level (dashed line)
-2
-3
-4
-5
-6
-7
-8
-9
M_V
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
W(7774) Å

This equation can now be compared to the equations illustrated in Figure 6.2. Figure 6.8 shows that the deviation of equation 6.10 from any other equation is, at most, 1 magnitude. The gradient is almost identical to all of the equations, with the exception of Sorvari/Osmer. As these equations are only meant to represent a relationship for F-type supergiants, the larger deviations from the line in Figure 6.7 for the two lower equivalent width value stars, namely ρ Aqu (G-type) and ν Per (bright giant), are not surprising.

The λ7748 (FeI) line was found to be unresolvably blended with a NiI line (RMT 156), but inspection of the data in Table 6.1 does show that an inverse relationship with temperature does exist. Values of the product E from equation 6.3 were calculated and plotted against M_V for the four stars later than F5 (see Figure 6.9). A least squares fit yields the equation.

$$M_V = -7.79 \times E \; -2.70 \qquad 6.11$$

The dashed line on the graph shows the relationship determined by Rao and Mallik (1978). Equation 6.11 agrees fairly well with equation 6.4, with a maximum M_V difference of less than 1 magnitude, considering the uncertainty of W(7748) and the scarcity of data points.

Figure 6.10 shows the six observed stars entered on the graph of Figure 6.3. The graph indicates how W(7774) differentiates between stars with respect to both temperature and luminosity. All of the stars lie in their respective regions, except for ε Aur, being too low. The drift experienced during these runs would account for this discrepancy.

Further investigations to determine the validity of the results obtained from the diode array were carried out on data collected after the period described in this work. Spectra from nine other stars were collected

in the region around $\lambda 7774$ and were combined with the results from the six stars listed here. The investigation was a collaboration between this author and Dr. G.R. Hopkinson (formerly of Durham University), and was published as a paper. A pre-print of this paper is included in Appendix B.

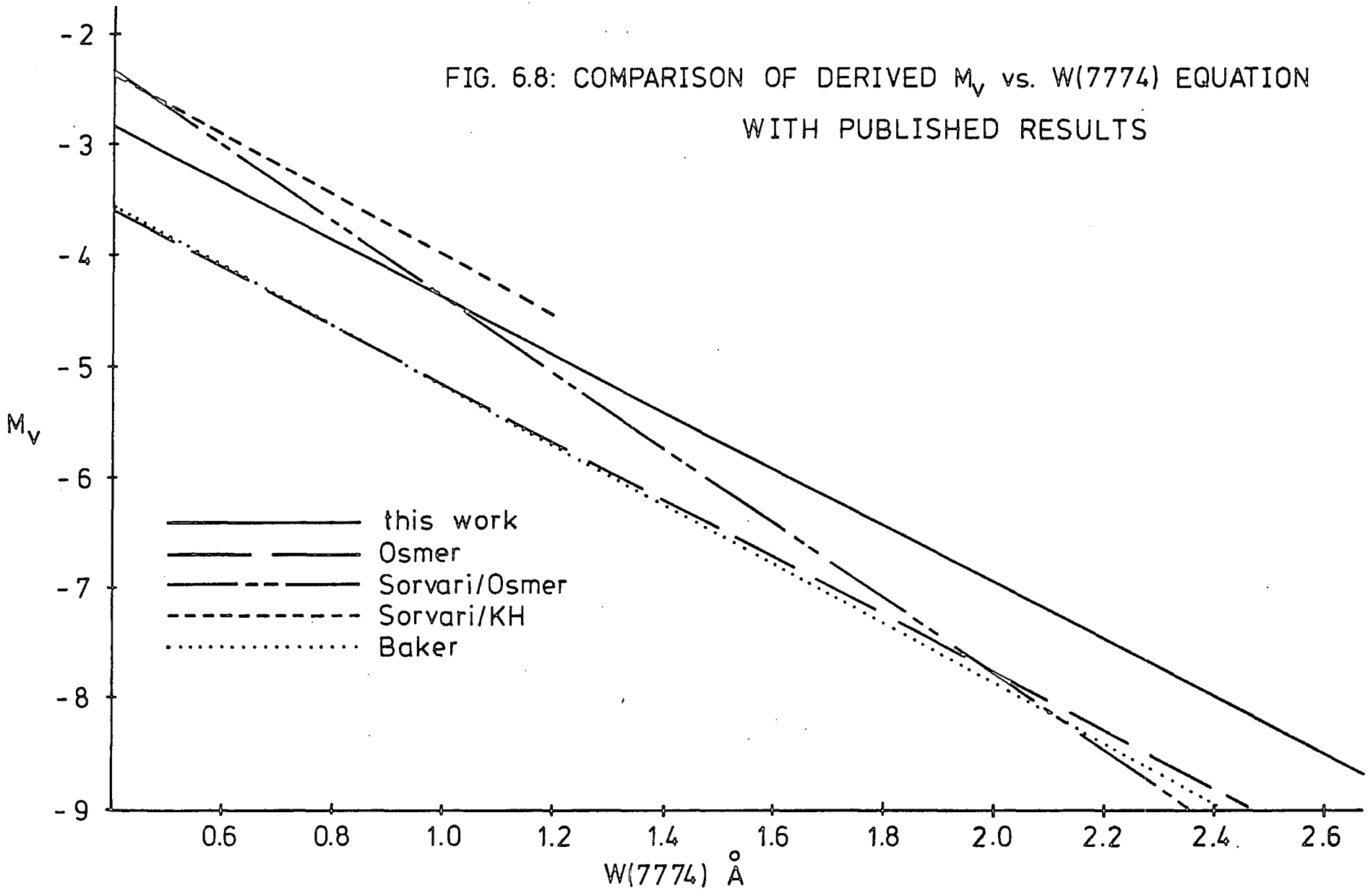


FIG. 6.8: COMPARISON OF DERIVED M_V vs. W(7774) EQUATION WITH PUBLISHED RESULTS
-2
-3
-4
-5
-6
-7
-8
-9
M_V
this work
Osmer
Sorvari/Osmer
Sorvari/KH
Baker
0.6
0.8
1.0
1.2
1.4
1.6
1.8
2.0
2.2
2.4
2.6
W(7774) Å

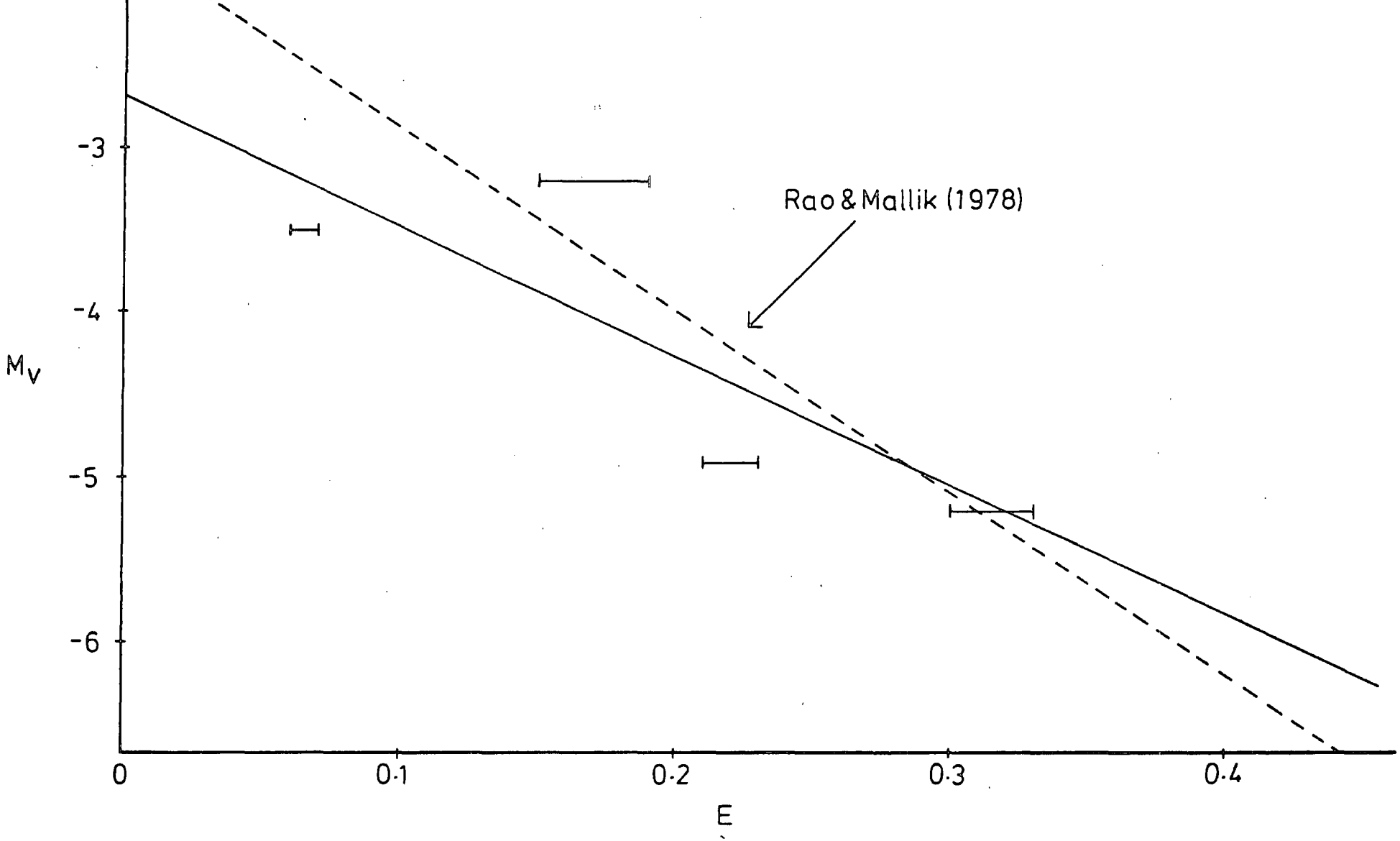


FIG.6·9: RELATIONSHIP BETWEEN M_V AND THE PRODUCT E=W(7774)×W(7748)
(ERROR BARS DERIVED FROM DIFFERENT CONTINUUM LEVELS)

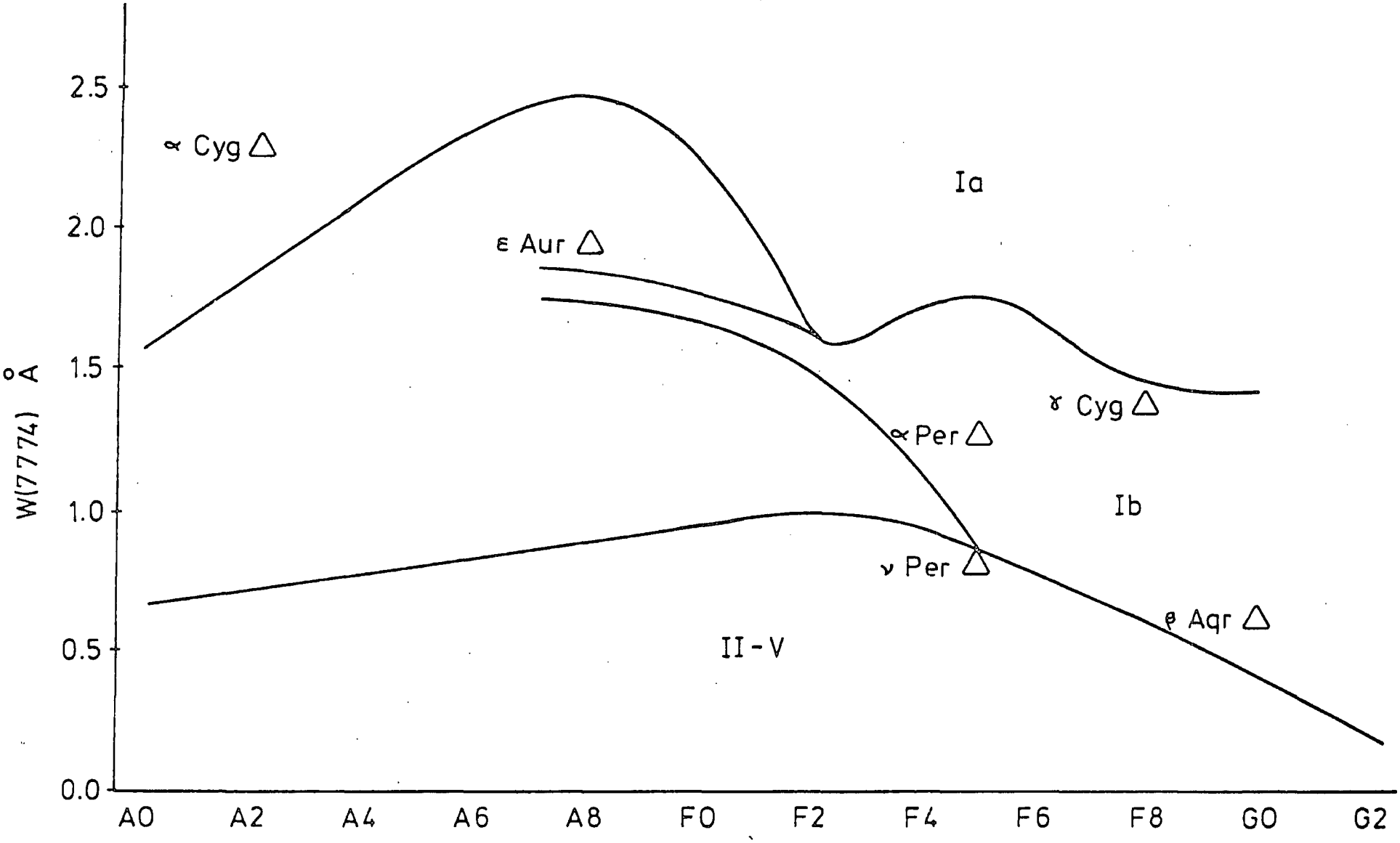


FIG. 6.10: STRENGTH OF W(7774) WITH SPECTRAL TYPE

REFERENCES

Abt, H.A.: 1957, Astrophys. J., 126, 138.

Allen, C.W.: 1973, "Astrophysical Quantities", Athlone Press, London.

Allen, C.W.: 1976, "Astrophysical Quantities", Athlone Press, London.

Baker, P.W.: 1974, Publs. astr. Soc. Pacif., 86, 33.

Baschek, B., Scholz, M., and Sedlmayr, E.: 1977, Astr. Astrophys., 55, 375.

Engold, O., and Rygh, B.O.: 1978, Astr. Astrophys., 70, 399.

Eriksson, K., and Toft, S.C.: 1979, Astr. Astrophys., 71, 178.

Gray, D.F.: 1976, "The Observation and Analysis of Stellar Photospheres", Wiley and Sons, New York.

Griffin, R.F.: 1969, Mon. Not. R. astr. Soc., 143, 319.

Hack, M., and Selvelli, P.L.: 1979, Astr. Astrophys., 75, 316.

Houziaux, L.: 1969, "Theory and Observations of Normal Stellar Atmospheres", ed. O. Gingerich, MIT Press, Cambridge, 305.

Johnson, H.R., Milkey, R.W., and Ramsey, L.W.: 1974, Astrophys. J., 187, 147.

Keenan, P.C., and Hynek, J.A.: 1950, Astrophys. J., 111, 1.

Kraft, R.P., Preston, G.W., and Wolff, S.C.: 1964, Astrophys. J., 140, 235.

Luck, R.E.: 1977, Astrophys. J., 212, 743.

Merrill, P.W.: 1925, Publs. astr. Soc. Pacif., 37, 272.

Minnaert, M.: 1934, The Observatory, 726, 328.

Osmer, P.S.: 1972a, Astrophys. J. Suppl., 24, 247.

Osmer, P.S.: 1972b, Astrophys. J. Suppl., 24, 255.

Parsons, S.B.: 1964, Astrophys. J., 140, 853.

Rao, N.K. and Mallik, S.G.V.: 1978, Mon. Not. R. astr. Soc., 183, 211.

Sorvari, J.M.: 1973, Ph.D. Thesis, University of Rochester, New York.

Sorvari, J.M.: 1974, Astr. J., 79, 1416

Thomas, R.M., Morton, D.C., and Murdin, P.G.: 1979, Mon. Not. R. astr. Soc., 188, 19.

CHAPTER SEVEN

MASS LOSS

7.1 Introduction

The loss of material from stars is no surprising phenomenon to astrophysicists. The cataclysmic ejection of matter can give rise to novae, supernovae, or planetary nebulae (ie dense, gaseous structures surrounding hot, low luminosity stars). In addition, mass loss can take place in most normal stars during much of their lifetimes, though the circumstances might not be so explosive. According to Reimers (1977), all red giants eject mass. Even our own sun continuously sheds its mass in the solar wind, but this amounts to only 10^{-14} solar masses per year, which is negligible compared with nuclear fuel burning. Late-type giants and supergiants are believed to be responsible for approximately one-fifth of the return of matter to the interstellar medium through stellar mass loss (Hagen, 1978). In our own galaxy, Reimers (1975) estimates that red giant mass loss supplies one-half of the gas used for new star formation. So significant amounts of mass loss are expected in the evolution of stars.

7.2 Evolutionary Theory of Mass Loss

The theory of evolution of red giants and supergiants requires the loss of significant amounts of stellar mass to occur. For a star with a mass of less than 1.4 solar masses, once it has exhausted its supply of nuclear fuel, it evolves into a white dwarf. White dwarfs are stars composed of degenerate matter, where the atomic nuclei exist compacted, stripped of their electrons. The density of such matter ranges from 10^9 to 10^{10} $kg.m^{-3}$, resulting in stellar dimensions approximating those of the earth.

Stars with masses greater than the white dwarf (or Chandrasekhar)

limit will undergo collapse after nuclear fuel depletion, and will result in supernova catastrophes. The observed rate of supernova explosions in the entire galaxy is only about one in every century (Shklovskii, 1978). However, the "death" rate of stars with masses greater than 1.4 solar masses is considerably higher. Deutsch (1969) estimates that a star "death" should occur once every three years. Therefore, there must be one or more mechanisms by which these more massive stars can shed sufficient matter to fall below the mass limit for supernova eruptions. Recent theories propose that this mass is lost predominantly by one of two methods - either from radiation pressure, or via a stellar wind driven by the dissipation of acoustic wave energy.

7.3 Mass Loss Mechanisms

Matter is lost from the surfaces of stars via one of two main methods, namely by radiation pressure or by stellar wind.

Radiation pressure acts on the stellar dust grains. It cannot function in metal-deficient stars due to the scarcity of elements which are able to condense into grains (Fusi-Pecci & Renzini, 1975). At temperatures at or below 1000K, dust grains are formed in the relatively cool stellar photospheres. Radiation pressure causes the condensing dust to be accelerated through the stellar gas, transferring its momentum to the gas by collisions. The gas is, in turn, accelerated to supersonic velocities (Gehrz & Woolf (1971) estimate the dust to gas mass ratio as being around 500), forming a circumstellar envelope. The gas velocities can be determined from the circumstellar line velocities.

Mass loss by radiation pressure appears to be the dominant mechanism in hotter stars. Giants of spectral type M3 and earlier have no dust shells (Goldberg, 1979), yet they give optical evidence for the loss of matter. Therefore, this cannot be the only method of shedding stellar material.

In cooler stars (eg. red giants), stellar wind generated by acoustic flux is the dominant mechanism of mass loss. Turbulance in cool rotating stars causes matter to be ejected in the forward direction of rotation, resulting in a rapid loss of angular momentum. If the turbulent elements have root mean squared velocities of the order of one-third of the stellar escape velocity, and have dimensions of only a few per cent of the stellar radius, then a cool silicate disk will form around the equatorial region of the star, and rotational braking will occur (Kafatos & Michalitsianos, 1979). Once grains and silicates are formed around the star from these shock driven stellar winds, this circumstellar material can then be expanded further by radiation pressure. (Due to the dust to gas coupling, radiation pressure is a highly effective accelerating mechanism in late-type stars as well as the hotter, early-type stars, but it is more likely to increase the rate of expansion than initiate it (Goldberg, 1979)).

7.4 Mass Loss Detection

The first proposals for the loss of mass from red giants were put forward by Deutsch (1956). The presence of blue-shifted absorption cores observed in stellar spectra supports the theory that a cool expanding envelope surrounds these stars, transferring mass to the interstellar medium. These circumstellar (CS) envelopes constitute the material from which new stars are to be created. Mass loss in red giants and supergiants should be observed via the CS envelope. Hagen (1978) observes an infrared excess in late-type stars, due to thermal radiation from CS envelope dust giants.

Weymann (1962) also found metal lines in M-type stars with deep, blue-shifted absorption cores, or even emission components, indicating some form of mass loss. Evidence exists (Rosendhal, 1973) that all of the most luminous OB stars exhibit mass flows and outwardly accelerating envelopes.

Many stars (including cool giants and supergiants) undergoing mass

loss exhibit spectra with P Cygni profiles. In its simplest form, a P Cygni profile is an absorption line with a sharp blue side and a winged red side (see section 7.4.1). It may also consist of double or triple absorption lines, due to line formation in different shells of the stellar atmosphere. This characteristic indicates (de Groot, 1969) an extensive expanding atmosphere with a uniformly decreasing density (ie. inversely proportional to r^2).

All cool stars experiencing mass loss exhibit strong Fraunhofer lines (with excitation potentials of less than 1ev) having blue shifted components. Also, it is possible that if a star consists of hot, rising and cool, descending elements, then there should be higher velocities observed for lines of small excitation potential (Goldberg, 1979).

7.4.1 P Cygni line profiles

There are many different types of P Cygni (P Cyg) profiles (Beals, 1950), and to describe each one, a highly complicated explanation would be required. One of the basic P Cyg profiles can be described in more or less simple terms.

If the stellar atmosphere is extensive and expanding, then there is a redward emission component which (unlike the absorption component) is found to originate high in the atmosphere. The envelope must be large with respect to the photosphere to yield a significant amount of scattered light. If V equals the expansion velocity of the atmosphere, then the emission component exhibits all velocities in the range of -V to +V, while absorption shows only -V. However, electron scattering (ie the increase in the continuum opacity) reduces the red wing from the receding side of the envelope, and increases the absorption component. Also, the central disk attenuates the redward component somewhat. These combined effects yield a line profile with a blue shifted absorption core, having a sharp blue side and a winged red side, with a possible redward emission visible.

7.4.2 H-alpha feature

The Balmer line of H_α, when in emission with a P Cyg profile, is a very strong indicator of mass loss (Underhill, 1960). The lower limits of absolute visual magnitude where H_α begins to indicate emission is $M_v \sim -5.8$ for B0 to B1 stars, and $M_v \sim -6.8$ for B8 to A3 stars (Rosendhal, 1973). As expected, the location of the emission formation is in the outermost parts of the expanding stellar chromosphere. In K-type supergiants, this circumstellar originating component is superposed on an absorption core, shifted by -50 to -80 $km.s^{-1}$, or about 1 to 2Å (Kraft, et al, 1964). No circumstellar absorption lines are found in any K or G supergiants, except possibly CaI intercombination lines at $\lambda 6572$ (Reimers, 1977).

7.5 Circumstellar Features

Observation of circumstellar (CS) features in the near-infrared and infrared regions of the spectrum can give evidence for the loss of matter from the surfaces of stars. All stars later than F5 are believed to have chromospheres, coronae, and stellar winds, and therefore must experience some sort of mass loss (Deutsch, 1969). Most G-type and K-type supergiants and all giants and supergiants later than M0 exhibit strong metal lines of low excitation with blue displaced absorption cores (Reimers, 1975, 1977). This core displacement has been measured to be around 10 $km.s^{-1}$ (Hagen, 1978), and is direct observational evidence for a CS envelope.

This blue displacement of the absorption cores is most strongly observed in supergiants, but very high luminosity is not necessarily a requirement. Stars with this characteristic are all mass ejectors, having cool expanding CS envelopes (Boesgaard & Hagen, 1979), which can reach out to as far as several hundred stellar radii. The CS line strengths tend to increase with later spectral types and with increasing luminosity. Early measurements of the CS envelope velocities ranged from 5 to 25 $km.s^{-1}$

(Reimers, 1969). This is less than the photospheric escape velocity of around 100 $km.s^{-1}$, but at the heights to which the envelopes extend, the escape velocity is exceeded. The expansion velocity is directly determined from measurements of the core displacement.

The strongest CS lines are the CaII H and K lines at λ 3968 (H) and λ 3934 (K). The CS component of the CaII is a sharp, deep, blue-shifted absorption core (k_4) superposed on the chromospheric emission core (k_2). The chromospheric absorption component is k_3 (Goldberg, 1979). Other lines of interest are the Na D lines ($\lambda\lambda$5890, 5896), CaI (λ4227), SrII (λ4077), and H_α(λ6563), all of which can be readily observed in supergiants (Reimers, 1975).

7.5.1 Supergiants

Among the largest and most luminous stars known are the supergiants. They are young, rapidly evolving stars located high up on the Hertsprung-Russell diagram. Mose are light variables and radial velocity variables, with extended atmospheres exhibiting high turbulance, and thus have broadened spectral features. Although they are massive stars, they maintain high luminosity-to-mass ratios. Since the group of supergiant stars includes some of the brightest stars known (with absolute magnitudes of up to $M_V \sim -9$), they can very effectively be used as distance indicators.

Late-type supergiants have low densities and low surface gravities. According to Deutsch (1956), each is surrounded by an envelope of escaping material which can extend out to a radius of 1.5×10^{11} km (compare the maximum distance between Pluto and the Sun of only 7.3×10^9 km). They therefore exhibit circumstellar lines with blue displaced absorption cores. Large scale shock waves in the cool extended atmospheres can cause matter to be lost to the interstellar medium. Mass may also be sporadically ejected by the acceleration of material from turbulence in the atmosphere. In fact, Reimers (1975) states that all M-type giants and supergiants

continuously eject matter into the interstellar medium. Deutsch (1956) claims that photospheric (reversing layer) lines in all M-type supergiants are broadened "as though by turbulance" and Herbig & Spaulding (1955) found that luminous supergiants up to F8 have broad lines. However, Kafatos & Michalitsianos (1979) deduce that late-type supergiants (and giants) are slow rotators because the lines are sharp. Clearly, this area bears some investigation.

7.6 Mass Loss Rates

The densities and dimensions of stellar dust shells can be deduced from CS emission line profiles. Mass flow rates must be determined from the region in which grains are found and accelerated (Gehrz & Woolf, 1971). Infrared observations refer to the region just outside this, where the density of the dust is high, and therefore give better information on mass flow rates than on the mechanisms in operation.

The rates of loss of material from stars has been calculated in many different ways by different investigators. Goldberg (1979) finds that the mass loss rate, $\dot{M}$, is proportional to the stellar surface area, such that

$$\dot{M} = 2.6 \times 10^{-12} \cdot R^2 \quad M_\odot \cdot yr^{-1}$$

and is not necessarily proportional to the absolute magnitude. Mullen (1978) had determined a relationship of

$$\dot{M} \propto M.R^{\frac{1}{2}}$$

which agrees well with Goldberg for a constant of proportionality of 1.6×10^{-9}. Fusi-Pecci & Renzini (1975) have developed a formula for mass loss, such that

$$\dot{M} = \eta \cdot \left({}^{-R}/_{G.M}\right) \cdot L_{ac} = \eta \cdot \left({}^{Lac}/_{g.R}\right) \quad M_\odot \cdot yr^{-1}$$

where η = efficiency factor

L_{ac} = total acoustic energy generated in the convective region.

For particular types of stars, mass loss rates have also been calculated. For example, Reimers (1975) shows that in red giants

$$\dot{M} \propto {}^{L}/_{g.R} \quad \text{or} \quad \dot{M} \propto g^{-1.6}$$

where the empirical formula is

$$\dot{M} = 4 \times 10^{-13} \cdot ({}^{L}/_{g.R}) \quad M_{\odot} .yr^{-1}$$

For late-type supergiants, Deutsch (1956) gives a particular value of rate of mass loss of

$$\dot{M} \sim 3 \times 10^{-8} \quad M_{\odot} .yr^{-1}$$

and for K-type supergiants, Reimers (1977) has

$$\dot{M} \sim 10^{-7} \quad M_{\odot} .yr^{-1}$$

both of which agree with the results of Goldberg.

In the genral case, Goldberg, Bernat (1977), Weymann (1962), and Hagen (1978) all use the formula

$$\dot{M} = 4\pi r^2 \cdot \rho(r) \cdot V(r)$$

and if V(r) is a constant, then

$$\rho(r) \propto {}^{1}/r^2$$

yielding

$$\dot{M} = 4\pi \cdot (\mu . m_H) \cdot N_H \cdot R_i \cdot V$$

where V = expansion velocity

R_i = inner shell radius

N_H = density of hydrogen atoms in the shell

$(\mu . m_H)$ = mass per hydrogen atom

This formula then relates the rate of mass loss to the inner CS shell radius, ie.

$$\dot{M} \propto \text{(column density)} \cdot \text{(inner radius)} \cdot \text{(expansion velocity)}$$

or

$$\dot{M} \propto R_i$$

and herein lies the chief uncertainty, the determination of the inner shell

radius. Gehrz & Woolf follow the same fundamental formula, with the density and velocity pertaining to the gas, ie.

$$\dot{M} = 4 \pi R^2 \cdot \rho_g \cdot V_g$$

Finally, de Loore (1977) quotes the values of mass loss rates for particular stellar types. These values are listed in Table 7.1.

7.7 Observations

During the entire observation program, many stars were observed in the H_α region of the spectrum. Since emission of H_α is a criterion for stellar mass loss, this is the feature that was searched for.

Inspection of the collected stellar spectra indicated that most of the stars observed exhibited normal absorption lines. Only one star was found to strongly indicate the loss of matter from its surface. The F-type supergiant 89 Herculis (89 Her) shows a definite emission feature on the redward side of the H_α line (see Figure 7.1). Also, there may be some information in the structure of the deepest part of the absorption component.

In order to make any measurements on emission features, it is important to closely observe the profile of the line under investigation. Ideally, several stellar spectra are collected over the same approximate spectral region, with each one slightly shifted with respect to all of the others. This ensures that the feature or features of interest fall on different areas of the diode array. The collected spectra are then interpolated and superposed. With the data shifted to align the spectral features, the spectra are then averaged. In this way, errors in measurement due to sensitivity variations are minimized, and the line profiles can be more closely scrutinized. However, due to unfortunate circumstances, only one spectrum of 89 Her at H_α was collected. This meant that precise measurement of the line profile could not be carried out. The errors incurred are believed to be due to diode sensitivity variations across the array,

Table 7.1

Rates of loss of mass for particular stellar types

Type	$\dot{M}$ $(M_{\odot} \cdot yr^{-1})$
B, A Supergiants	$10^{-9} - 10^{-8}$
A2 Ia	3.0×10^{-10}
A5	9.6×10^{-14}
F0	4.8×10^{-13}
F2 Ia	10^{-8}
F8 Supergiants	10^{-5}
G0	8.4×10^{-15}
G2	2.2×10^{-14}
G5	4.4×10^{-15}
M0III	2.5×10^{-9}
Red Supergiants	10^{-5}

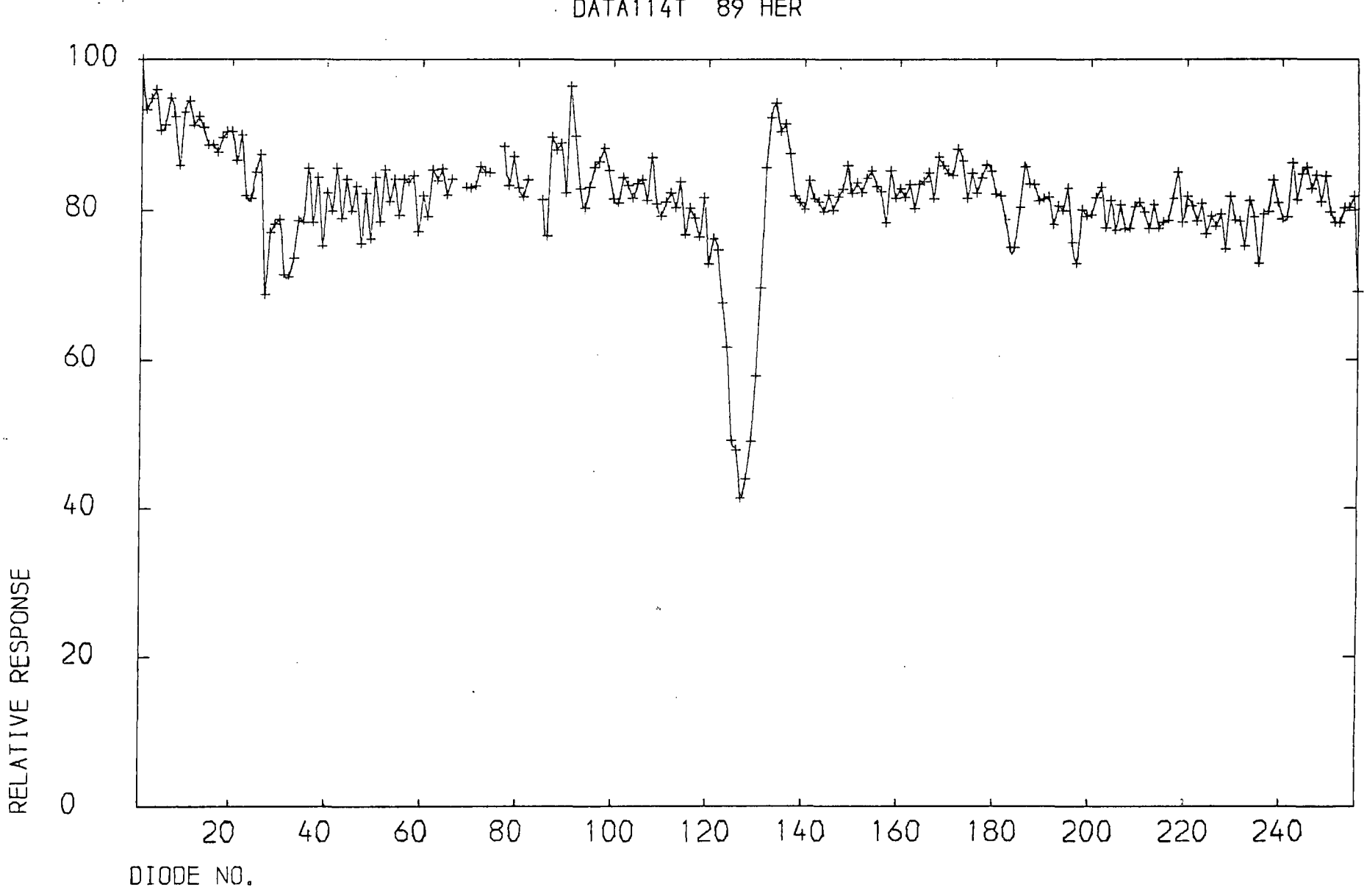


FIG. 7.1 : AN EXAMPLE OF H-ALPHA IN 89 HERCULIS SHOWING EMISSION

and to residual second-order effects from the odd-even splitting correction of the spectrum (see section 4.3.5). However, some indication of the order of magnitude of mass loss from 89 Her can be gained from the displacement of the absorption core of H_α from its rest position. The rest position can be determined by reference to the positions of nearby normal photospheric lines.

Figure 7.2 shows an enlarged portion of the spectrum depicted in Figure 7.1. The data points were interpolated eight-fold prior to magnification to avoid end effects. Also shown is the locus of the midpoints of the core of the line near to the line minimum. These points are used to indicate the measured position of the H_α absorption core. The calculated rest position of H_α (at 6563 Å) is also displayed on this spectrum.

The position of the absorption component of H_α is measured as being shifted 1.89 Å from the rest position, towards the blue end of the spectrum. This displacement is equivalent to a velocity of -86 $km.s^{-1}$, with an error believed to be $\pm 7 km.s^{-1}$. It is not known if the observed asymmetry of the core of the line is real or due to undesirable imaging effects. However, the stated velocity error is sufficient to take any such unsure results into account.

This value of the velocity of the core of H_α agrees well with the value of Sargent and Osmer (1968) of $-80 km.s^{-1}$, certainly within the limits of error. This indicates that, following the calculation of Sargent and Osmer, the rate of loss of material from the surface of 89 Her is indeed in the region of 10^{-8} $M_\odot$ $.yr^{-1}$.

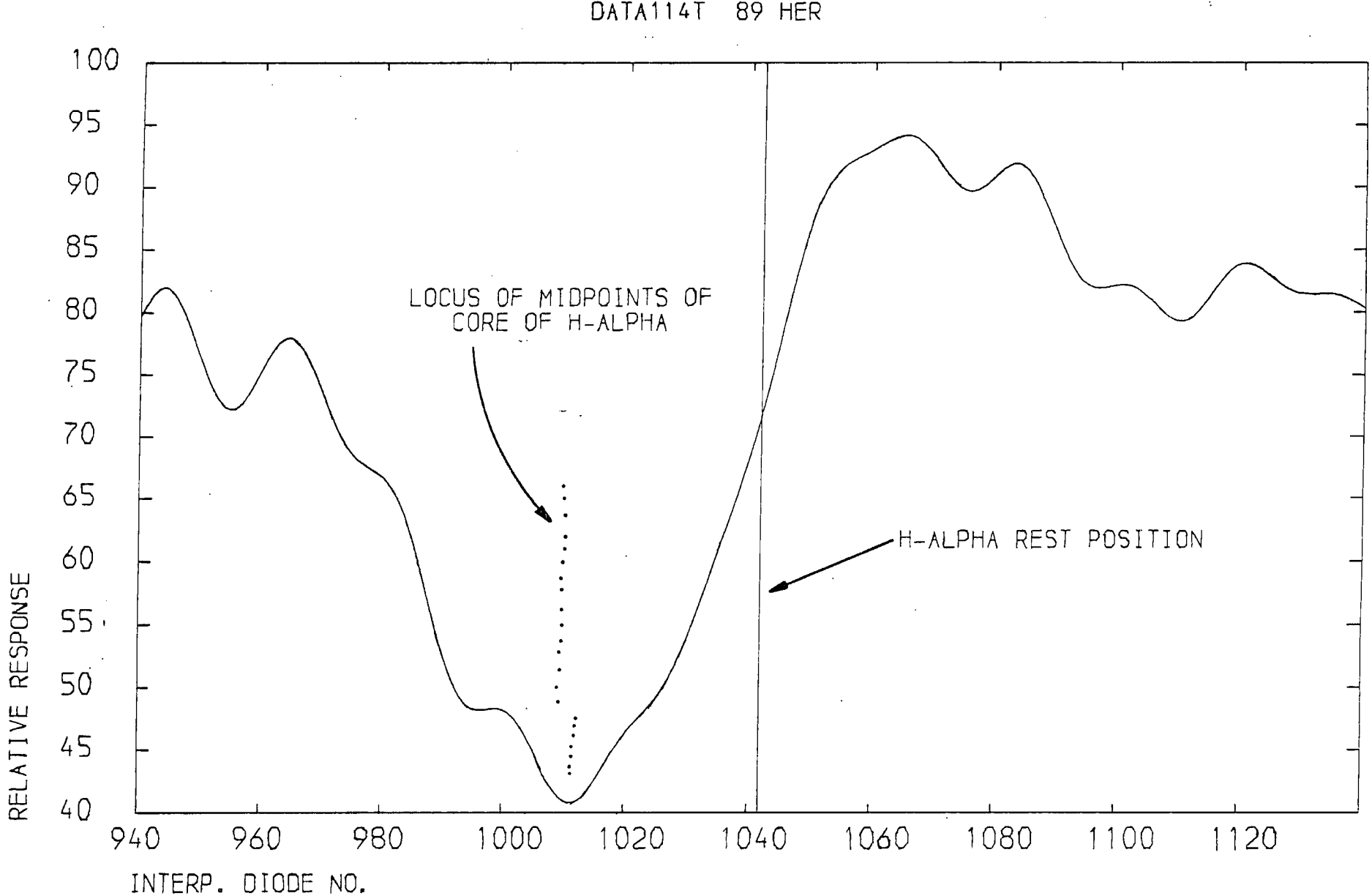


FIG. 7.2 : MAGNIFICATION OF H-ALPHA REGION OF PREVIOUS SPECTRUM

REFERENCES

Beals, C.S.: 1950, Publ. Dom. Astrophys. Obs., 9, 1.

Bernat, A.P.: 1977, Astrophys. J., 213, 756.

Boesgaard, A.M., & Hagen, W.: 1979, Astrophys. J., 231, 128.

Deutsch, A.J.: 1956, Astrophys. J., 123, 210.

Deutsch, A.J.: 1969, "Mass Loss from Stars", ed. M. Hack, Reidel, Dordrecht, 1.

Fusi-Pecci, F., & Renzini, A.: 1975, Astr. Astrophys., 39, 413.

Gehrz, R.D., & Woolf, N.J.: 1971, Astrophys, J. 165, 285.

Goldberg, L.: 1979, Q. Jl. R. astr. Soc., 20, 361.

de Groot, M.: 1969, "Mass Loss from Stars", ed. M. Hack, Dordrecht, Reidel, 26.

Hagen, W.: 1978, Astrophys. J. Suppl., 38, 1.

Herbig, G.H., & Spaulding, J.F.: 1955, Astrophys. J., 121, 118.

Kafatos, M., & Michalitsianos, A.G.: 1979, Astrophys. J., 228, L115.

Kraft, R.P., Preston, G.W., & Wolff, S.C.: 1964, Astrophys. J., 140, 235.

de Loore, C.: 1977, "Highlights of Astronomy", vol. 4, part II, ed. E.A. Muller, IAU, Reidel, Dordrecht, 155.

Mullen, D.J.: 1978, Astrophys. J., 226, 151.

Reimers, D.: 1969, Astr. Astrophys., 52, 457.

Reimers, D.: 1975, "Problems in Stellar Atmospheres and Envelopes", ed. B. Baschek, et al., Springer, Berlin, 229.

Reimers, D.: 1977, Astr. Astrophys., 57, 395.

Rosendhal, J.D.: 1973, Astrophys. J., 186, 909.

Shklovskii, I.S.: 1978, "Stars, Their Birth, Life and Death", Freeman & Co., San Francisco, ch. 15.

Underhill, A.B.: 1960, "Stellar Atmospheres", ed. J.L. Greenstein, Univ. of Chicago Press, Chicago, ch. 10.

Weymann, R.: 1962, Astrophys. J., 136, 844.

Sargent, W.L.W. & Osmer, P.S.: 1968, "Mass Loss from Stars", ed. M. Hack, Dordrecht, Reidel, 57.

DURHAM UNIVERSITY
12 MAY 1983
SCIENCE LIBRARY

CHAPTER EIGHT

CONCLUSION

Solid state detectors are among the newest and most important type of detector presently under development. Their response to light in the visible and especially the infrared regions of the spectrum render them vastly superior to photographic plates and the like. Added to their more sensitive longer wavelength response, the output signals obtained are highly suited for almost immediate reduction and analysis by means of either small, on-line or larger off-line computers.

In particular, the Plessey linear photodiode array has many advantages. The on-chip electronics minimize noise and pick-up, being in close proximity to the detector elements themselves. Also, four independent arrays of 256 elements each are located in the same package in two parallel lines. This geometry affords the ability to collect both a stellar spectrum and a reference spectrum simultaneously, or even a stellar spectrum and a flat field spectrum, for responsivity corrections. But overwhelmingly, the main advantage of the Plessey array is that it has a non-destructive read-out facility. This is of enormous benefit, especially when observing fainter objects for long periods of time.

On the other hand, several disadvantages and problems were encountered with the experimental set-up used. Firstly, movement of the array inside the cryostat during an exposure is a very serious problem, considering the size and high geometric stability of the individual array elements. For exposures of more than one hour, array displacement can, and in some instances did, take place, as discussed in Chapter Five. This displacement, if much greater than the pitch of the array elements, will effectively smear out any spectral features into obscurity. Laboratory measurements indicated that this movement could amount to as much as three diode spacings over a period of 90 minutes, clearly disastrous. At the time of finishing

the experimental work described here, plans had been put forward and the construction started to remedy this problem. A rigid mount for the detector array was fabricated out of insulating material, and copper braid was used to act as a flexible cold finger.

Secondly, the idea of on-chip electronics is a very good one. Unfortunately, the components themselves are light sensitive. If the light from the spectrometer (or any other source) does not fall precisely onto the array, but instead overspills onto the electronic circuitry, then unusual effects occur in the collected spectrum. Chapter Four described this problem, as well as an off-line computer software remedy.

In retrospect, several areas of research would be desirable to investigate with the Plessey diode array. The spectral response of the array would prove useful, especially to search for any variations between individual elements. Chapter Two did describe some spectral investigations, but these were only carried out on a "typical" element and at only two temperatures. A more comprehensive study of the spectral response as a function of diode temperature could be important, since some kind of balance needs to be found between an increased infrared response, and higher thermal noise and leakage drawbacks.

The actual change of diode response across and between the sensitive areas of the elements bears investigation. Chapter Five dealt with the theoretical case of a 20% dead space between adjacent array elements. If close observations of line profiles and weak line observations are ever hoped to be carried out, a better understanding of this effect needs to be completed.

The occurance of odd-even splitting is another serious problem. Off-line software compensation can remove this effect somewhat, but obviously not entirely, as there was no way of knowing exactly where the overspilled light was doing the "damage". Masking of the non-diode area is one solution; a cylindrical lens mounted over the array is another. Both of

these methods were attempted, but without much success. A cylindrical lens was manufactured and tried, but it could not be aligned parallel to the array quite accurately enough. The problem with the mask was that it could not be brought close enough to the silicon material, due to the physical construction of the array package. On the whole, it seems that a cylindrical lens would remedy the situation best, provided that it could be attached to the outside of the array package with sufficient accuracy.

In order to reduce the likelihood of unsuitable results, which had appeared to be satisfactory at the time of the observation, it would be desirable to reduce the data before the end of the observation program. Since this was the first test of the diode array on a telescope, many problems with data reduction were not envisaged until the results were more closely investigated. Hence, the development of the data reduction software only took place well after the observing time. Although each of the spectra could be observed almost immediately after a flat field compensation, the odd-even splitting effect could not be removed without more powerful computing facilities. This can have the effect of an apparently good spectrum not really being totally satisfactory. Also, if radial velocity measurements are being carried out, it is essential that elemental arc calibration runs are taken more frequently. Even if the diode array movement inside of the cryostat is remedied, any mechanical displacement at the focus of the telescope of even fractions of a millimetre could upset calculations. Therefore, if accurate radial velocities are to be determined, a calibration exposure should be collected after every stellar exposure.

Finally, it was thought that on future observations, a more well planned, precise astronomical program should be developed first. Admittedly, the observation program described here was mainly a proving ground for the Plessey diode array. However, as described in Chapter Six and Chapter Seven, many more results are required for each region of astronomical interest

observed. Perhaps only the oxygen triplet should have been investigated and not mass loss at all, or vice versa. But, as such, some results were obtained and many difficulties were ironed out; not forgetting that one of the biggest advantages of this system is the non-destructive readout capability. So it appears that this device, with a few minor refinements, can be quite a powerful tool in visible and near-infrared astronomy.

APPENDIX A

FIG. A-1 TEMPERATURE REPRODUCIBILITY I

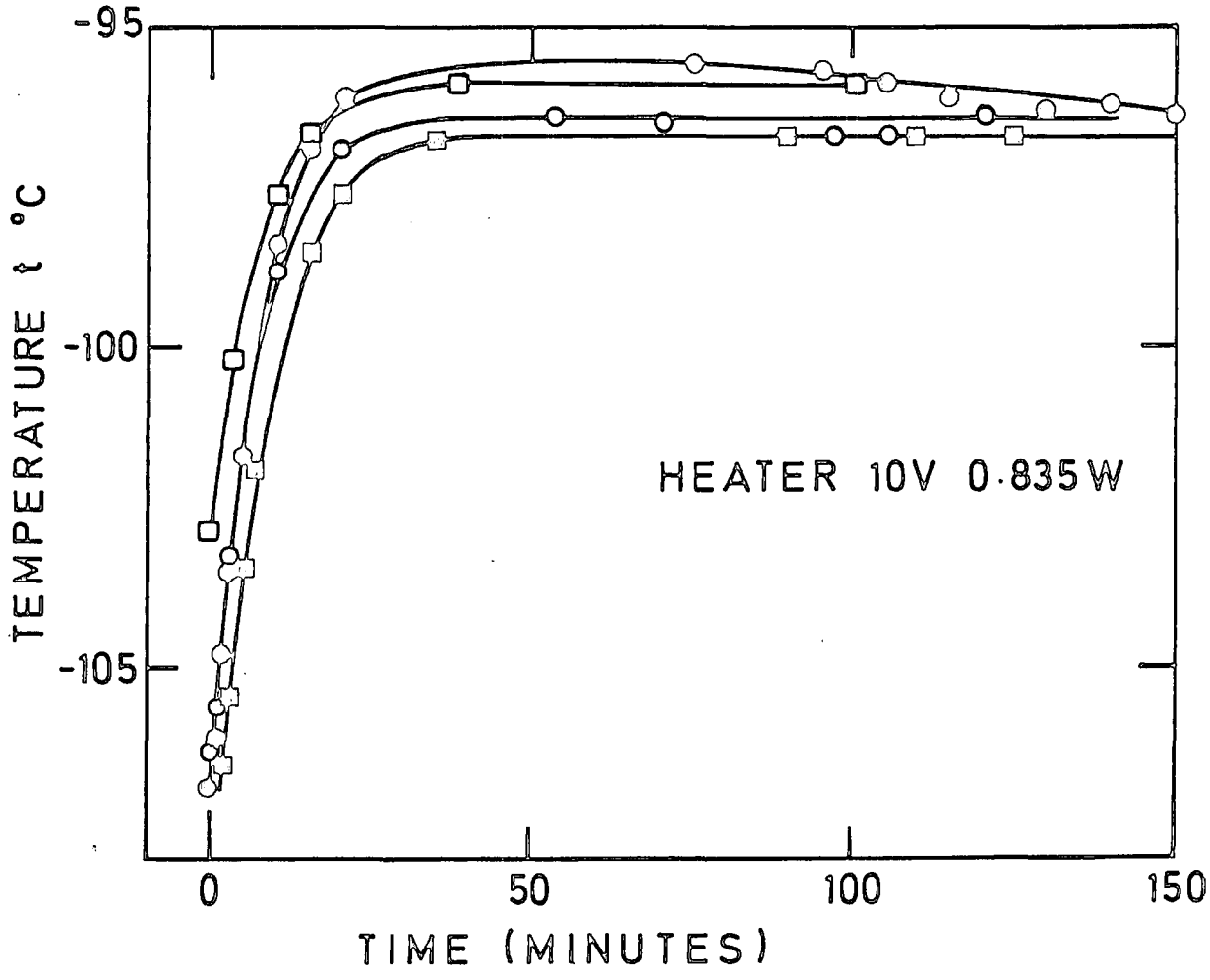


FIG. A·2 TEMPERTURE REPRODUCIBILITY

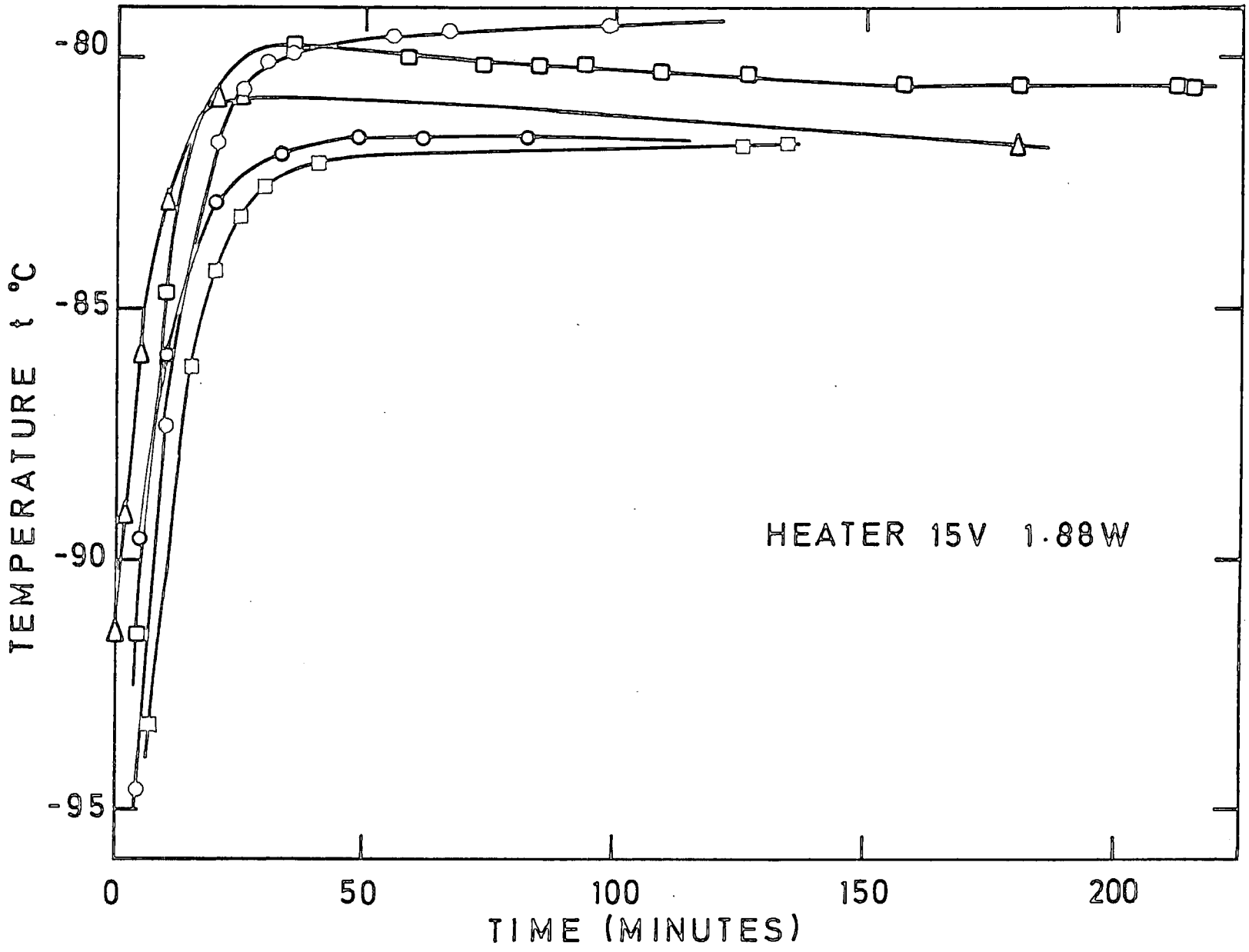


FIG. A-3
TEMPERATURE REPRODUCIBILITY II

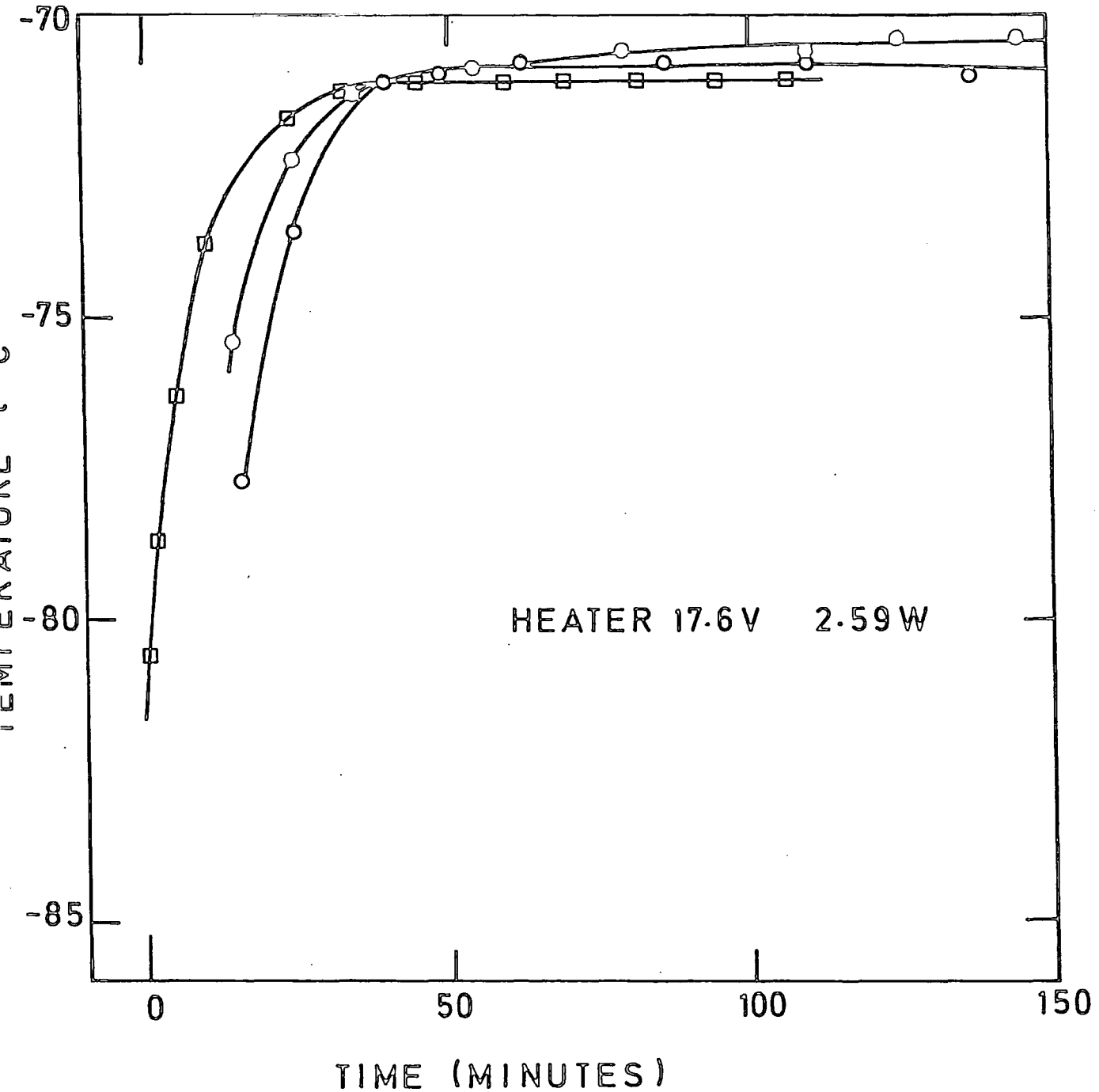


FIG.A-4 TEMPERATURE REPRODUCIBILITY Ⅳ

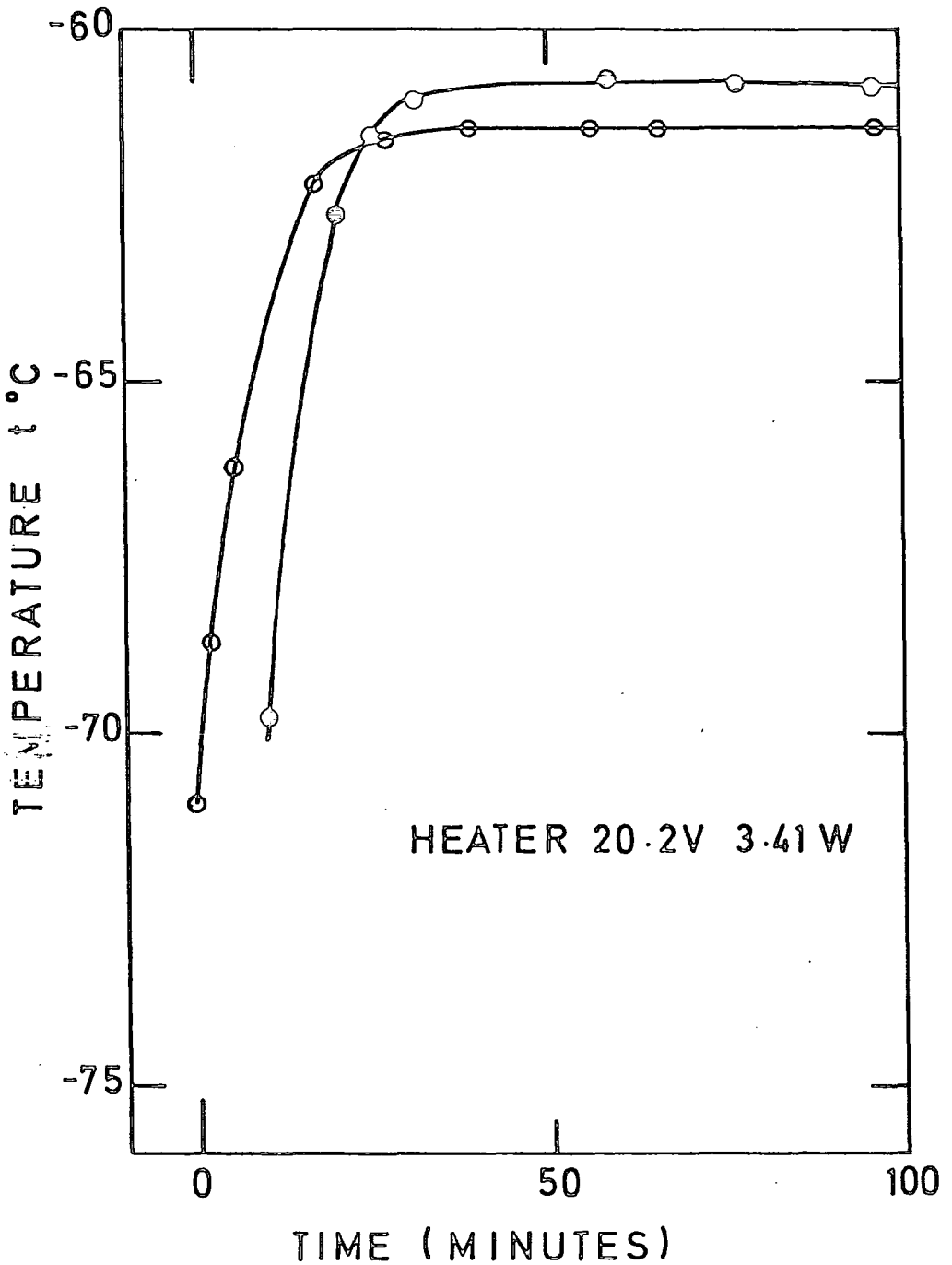


FIG.A·5 TEMPERATURE REPRODUCIBILITY V

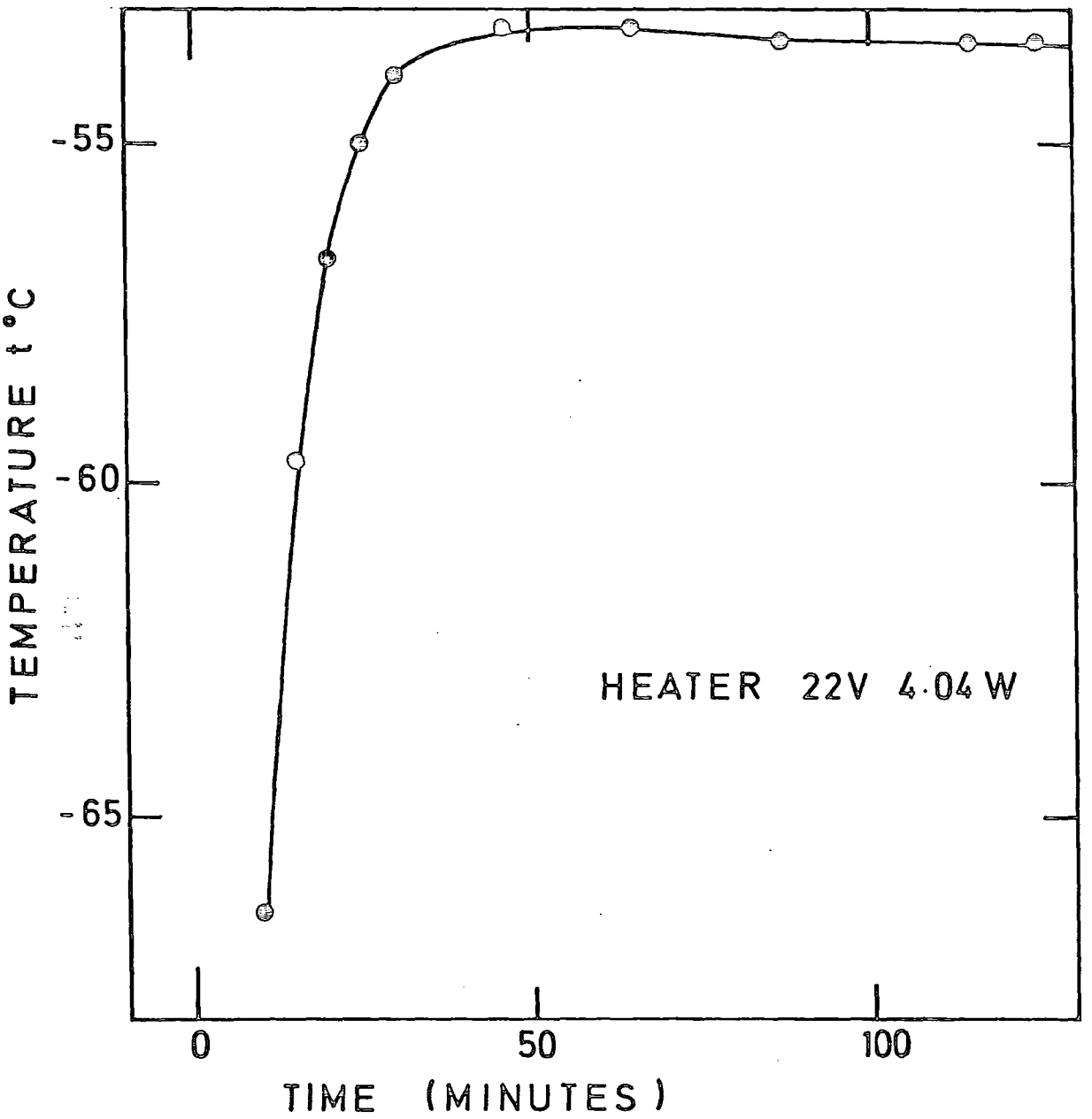


FIG. A·6 TEMPERATURE REPRODUCIBILITY VI

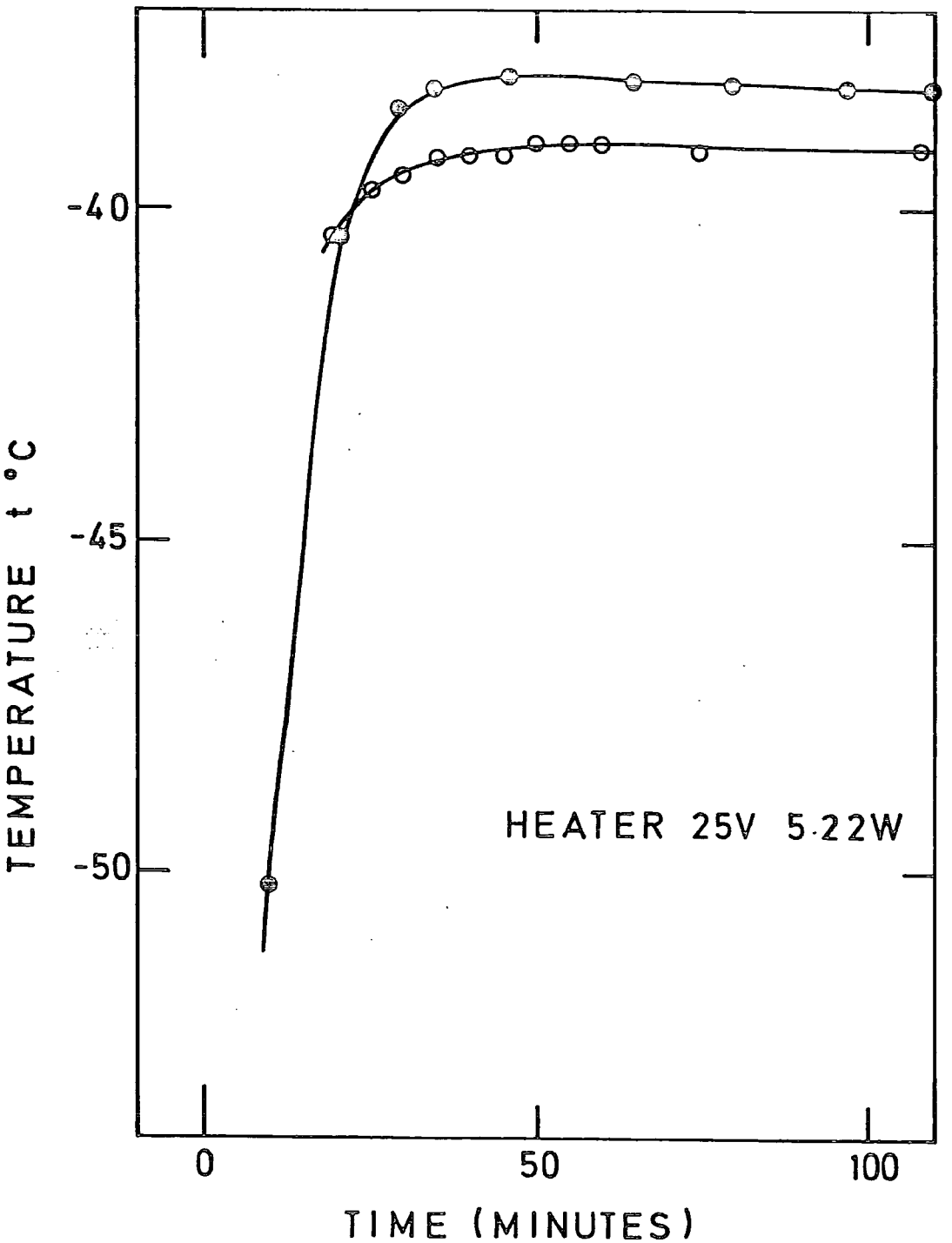
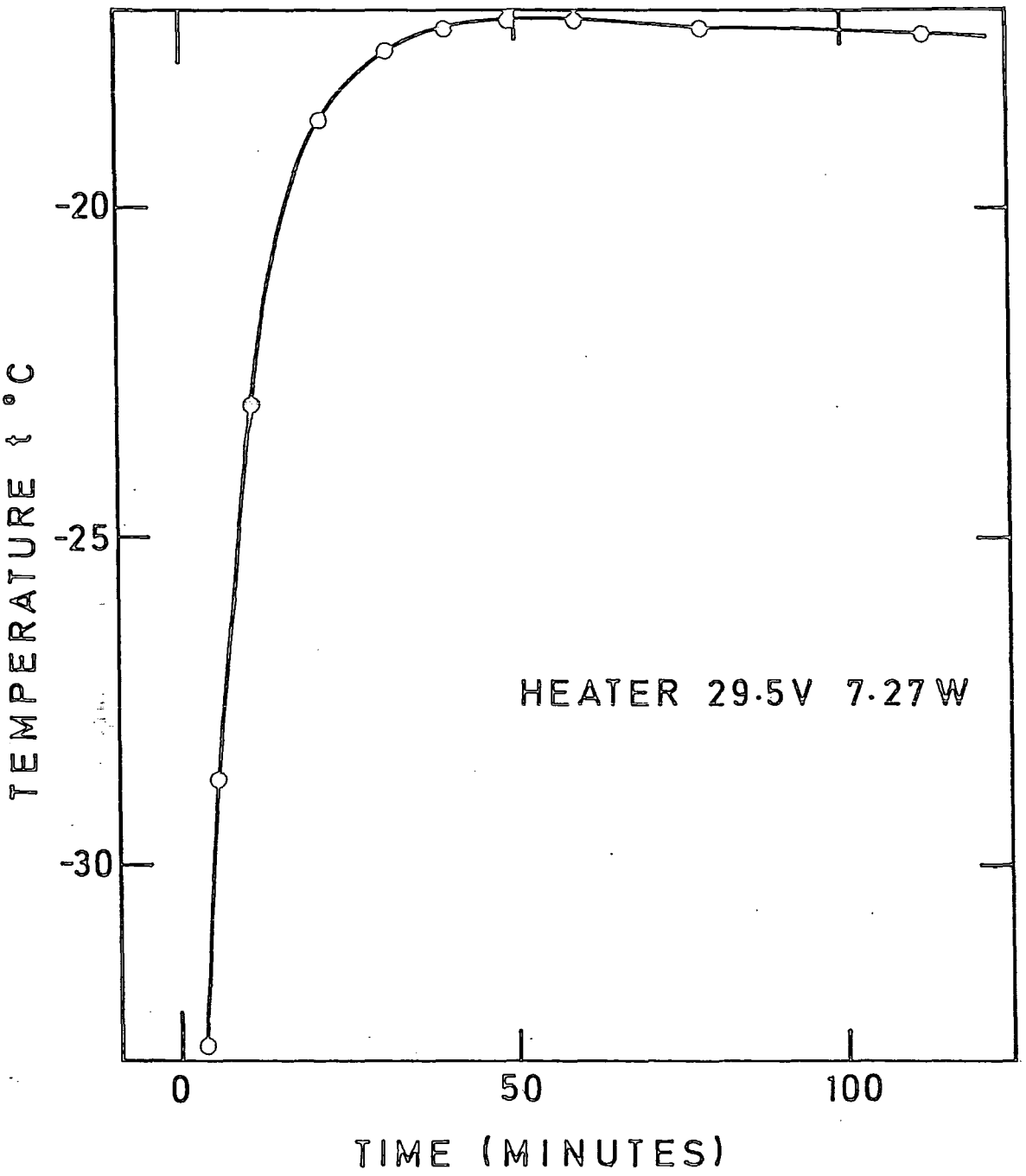


FIG. A·7 TEMPERATURE REPRODUCIBILITY VII



APPENDIX B

Mon. Not. R. astr. Soc. (1981) 195, 661–668

Observations of the O I λ 7773 triplet in intermediate-type supergiants using a linear photodiode array

G. R. Hopkinson* and A. Humrich* *Department of Physics, Science Laboratories, South Road, Durham DH1 3LE*

Received 1980 October 10; in original form 1980 July 24

Summary. Partially resolved spectra of the infrared oxygen triplet in A–G supergiants have been obtained with a new self-scanned photodiode array system. Curve of growth analyses indicate that the lines are formed in non-LTE. A line is identified at λ 7777.9 which is strong in A supergiants and which will complicate the analysis of low resolution spectra. At a resolution of 0.45 Å/diode the CN lines which appear in G8 and later stars are blended with the O I triplet rendering its equivalent width unreliable as a luminosity indicator.

1 Introduction

The study of the neutral oxygen λ 7773 triplet is interesting for several reasons. The combined strength of the lines at λλ 7771.94, 7774.17 and 7775.39 has been established as a measure of stellar luminosity. The relationship was first noted by Keenan & Hynek (1950) for B5–G2 stars and has since been confirmed by several authors. The coverage of spectral types has recently been extended to B1 by Thomas, Morton & Murdin (1979). A knowledge of luminosity class or absolute magnitude of course allows stars to be used as distance indicators.

The strength of the λ 7773 feature is very temperature sensitive due to the high excitation potential (9.15 eV) of the metastable $3s^5S$ lower level. The lines are particularly strong in supergiants where the extended atmospheres allow low collisional de-excitation rates. In this situation, however, the lines are likely to be formed in conditions of non-LTE (Johnson, Milkey & Ramsey 1974; Eriksson & Toft 1979). In fact Baschek, Scholz & Sedlmayr (1977) have suggested that the lines are formed in non-LTE even in main-sequence models; the effect being always to increase the equivalent widths compared with the LTE case. Note that for cool supergiants the lines can be enhanced, also in LTE models, because the continuous opacity (due to H^-) is lowered.

The O I triplet may also be useful for the determination of oxygen abundance if the non-LTE effects can be reliably estimated. In the later spectral types the [O I] magnetic dipole

* Present address: X-Ray Astronomy Group, University of Leicester, Leicester LE1 7RH.

transitions at λλ 6300 and 6363 can be used (e.g. Lambert & Ries 1977); the only other alternatives are [O I] λ 5577 and the λ 8446 triplet, both of which are seriously blended (with C_2 lines in the first case and Fe I lines in the latter).

Until recently the lack of suitable detectors has tended to restrict observations to the measurement of equivalent widths. The recent availability of solid state arrays means that the feature can be examined in more detail. With this in mind, a study of several intermediate-type supergiants was made during the first observing runs with a silicon photodiode array system built at Durham University.

2 The observations

The observing system was based on a prototype self-scanned photodiode array chip manufactured by the Plessey Co. Ltd in conjunction with the Royal Greenwich Observatory. Each chip contains four linear arrays of 256 diodes which are paired to form two parallel lines of 512 elements. The pixel size is 200 × 40 μm with a centre-to-centre spacing of 50 μm. A voltage sampling readout method was used followed by digitization to 12 bits. The device has a non-destructive readout capability which allowed an exposure to be monitored. The devices are not available commercially but, at present, a batch of 22 are on loan to the group (four of which have all 1024 elements operational). Data acquisition was controlled via CAMAC by a PDP 11/03 microcomputer and data stored on paper tape under the control of a CATY 2 program. The system has since been upgraded to include a DEC RXO2 floppy disc unit. Further details of device and system characteristics have been given elsewhere (Campbell *et al.* 1979; Hedge *et al.* 1978).

The observations were carried out during 1978 August and 1980 January using the coudé spectrograph of the 30-inch telescope at the Royal Greenwich Observatory. Spectra were taken with a single line of 256 diodes at a reciprocal dispersion of 0.45 Å/diode (~ 10 Å mm^{-1}) to give a spectral range of 115 Å. The detector was cooled to ~ −110°C using a cold finger dipped into liquid nitrogen. Micrometer adjustments provided horizontal, vertical and rotational alignment of the cryostat and tungsten and neon lamps were used for flat field and calibration. The projected slit width was 37 μm at the detector (1 arcsec on the sky) and to achieve sampling at the Nyquist rate or higher, two or more spectra were taken at slightly differing grating angles.

With a 12-bit ADC the noise was, at that time, digitization limited. Recent laboratory calibrations indicate that the step size was ~ 130 detected photons with a rather low quantum efficiency of ~ 20 per cent (at 8000 Å) and that the readout noise (with correlated double sampling) can be as low as 40 electrons. It is likely that variations in the fixed pattern noise (caused, for example, by drifts in the chip temperature) also produced some contribution to the noise. As an example of the overall performance, a spectrum of an $m_v = 2.3$ star could be obtained with a SNR of ~ 100 in 15 min. Note that poor weather and seeing conditions limited the observations to stars brighter than $m_v = 5$.

3 Data analysis

Several stages of data manipulation were necessary after the spectra had been extracted from the raw video levels. In some of the earlier runs misalignments resulted in a fall-off in signal to one end of the array which was not removable on division by a flat field spectrum. This was corrected by fitting clear regions of the continuum to a fifth-order polynomial.

Some problems were encountered in aligning the array. Odd–even effects were often produced since the switching transistors are sensitive to light and those associated with odd and

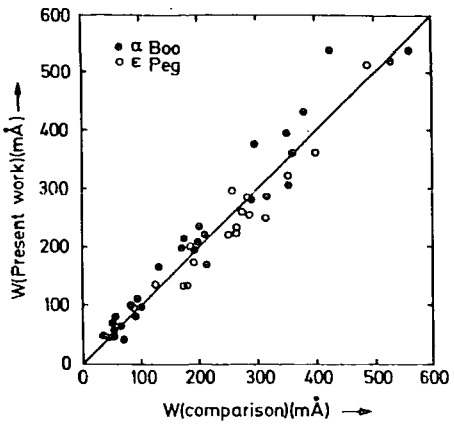


Figure 1. A comparison of equivalent widths.

even diodes are situated on either side of a row. Fortunately, these extra signals were removed satisfactorily in the data analysis. The magnitude of the correction was assumed to be proportional to the true signal and, to a first approximation, the effect of a slight misalignment was to produce a linear variation in the effect along the array.

Finally, when summing spectra taken at slightly different grating angles the data were first interpolated to 2048 points by adding zeros to the Fourier transform. The grating shifts could then be determined to 1/8 of a diode. The method of interleaved sampling is discussed by Bracewell (1965) and, before adding, the spectra should be convolved with window functions. In practice this was found to be unnecessary and a simple addition was performed. Care was taken however to minimize Gibbs-effect errors using the techniques described in Keating (1978) and Lanczos (1964).

To establish that the observing system could provide reliable equivalent widths, spectra of ϵ Peg around Hα were compared with the results of Harmer, Lawson & Stickland (1978) and a spectrum of α Boo around λ 5400 with the Arcturus atlas (Griffin 1968). The results are shown in Fig. 1. For α Boo our results are on average 5 per cent too large and for ϵ Peg 5 per cent too small, although this is not a true comparative test because of the differences in resolution and spectrograph characteristics. An upper limit of 5 per cent was placed on the combined contributions to the continuum of scattered light and drift in the acquisition electronics from observations of the residual signal in saturated lines of the atmospheric A band. This is an upper bound since the lines were not fully resolved.

4 Results

The spectra of the programme stars are shown in Fig. 2. The baselines and estimate of the continuum positions are also marked. An underexposed spectrum of α Lep is also included. The oxygen triplet can be seen partially resolved in all except the late-type stars. In α Cyg and η Leo there is a strong component to the blend at λ 7777.9; this is present to a lesser extent in the other supergiants and is tentatively assigned to a predicted line of Si I (as discussed later). Gaussian fits were made to this, the oxygen triplet and the nearby Fe I line at λ 7780.56 using a least squares technique. The equivalent widths of these and the Fe I/Ni I λ 7748 blend are given in Table 1. The errors are estimated to be ~ 10 per cent and are mainly due to random errors in the data and to inaccuracies in the placement of the

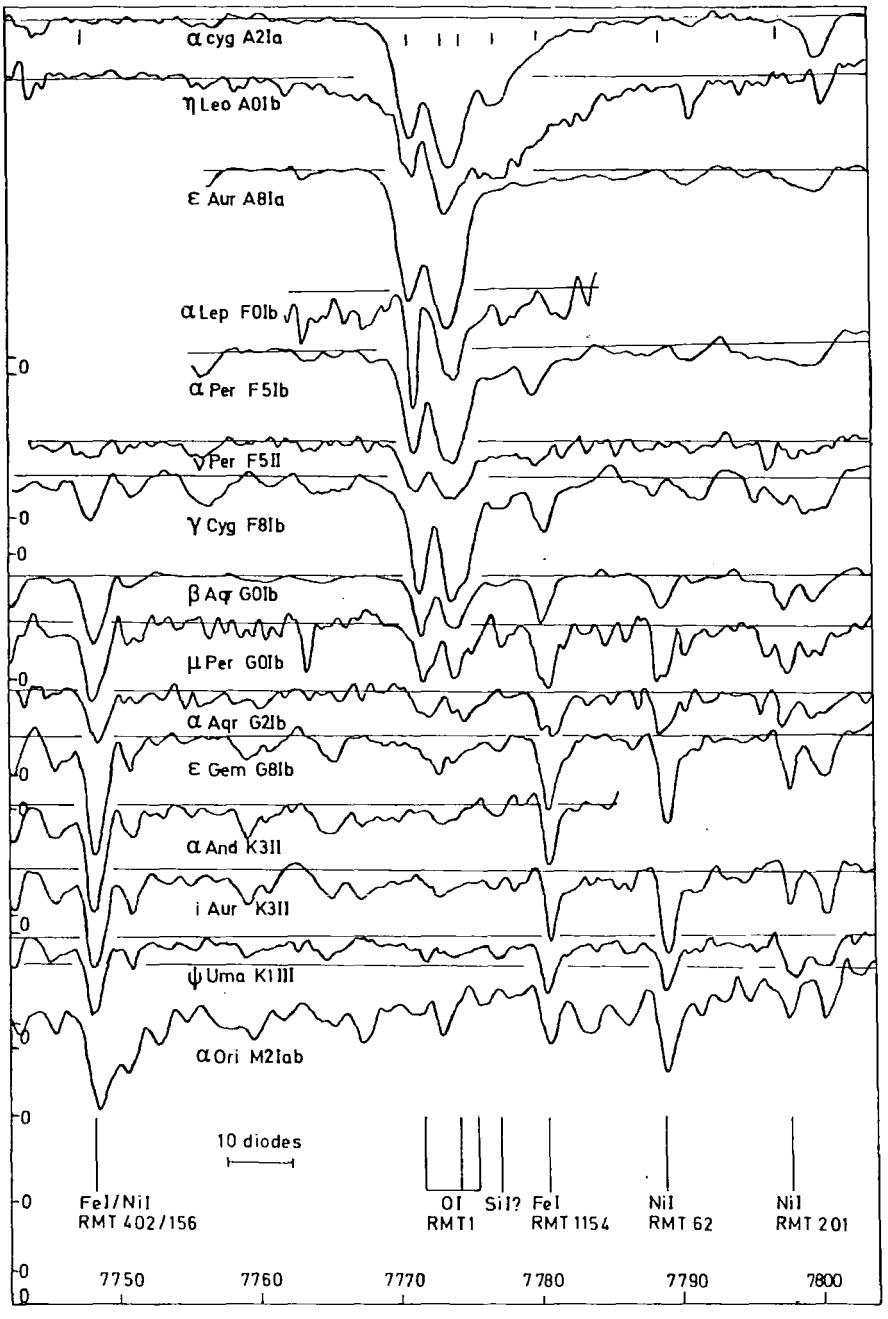


Figure 2. Photodiode array spectra in the region of λ 7770.

continuum level. W(O I) indicates the total equivalent width of the triplet and this is compared with previous studies in Table 2. According to Sorvari (1974) his results are not reliable for $W > 1.2$ and these are not included; nor are Parsons' (1964) since these only give a central depth. Other measurements of the triplet have been made by Sargent & Searle (1962) and Slettebak (1951) on Ap stars and Albers (1969) on southern Milky Way stars.

Table 1. Equivalent widths (in A).

Star	Type	Number of Observations	W(OI)	OI W(7772)	OI W(7774)	OI W(7775)	W(7777.9)	W(7748)	W(7780.6)
αCyg	A2Ia	3	1.9	0.75	0.59	0.57	0.74	∿0	0.31
εAur	A8Ia	4	2.4	0.98	0.84	0.58	0.19	∿0	0.07
ηLeo	A0Ib	2	2.0	0.76	0.42	0.87	0.82	∿0	0.93
αLep	F0Ib	1	1.2	0.40	0.39	0.39	-	-	-
αPer	F5Ib	3	1.56	0.54	0.49	0.54	0.19	0.18	0.32
γCyg	F8Ib	3	1.57	0.66	0.48	0.44	0.22	0.24	0.30
βAqr	G0Ib	3	0.64	0.31	0.22	0.11	0.09	0.32	0.19
μPer	G0Ib	2	0.66	0.32	0.22	0.11	0.12	0.44	0.34
αAqr	G2Ib	1	0.41	0.17	0.11	0.13	0.03	0.27	0.28
εGem	G8Ib	2	∿0.3	-	-	-	∿0	0.60	0.30
νPer	F5II	1	0.79	0.28	0.24	0.27	0.09	0.10	0.17
ιAur	K3II	2	-	-	-	-	∿0	0.45	0.21
γAnd	K3II	2	-	-	-	-	∿0	0.58	0.21
ψUMa	K1III	2	-	-	-	-	∿0	0.33	0.22

Table 2. A comparison with other measurements of W(O I).

Star	Present work (±10%)	a	b	c	d
αCyg	1.9	2.19	-	-	-
εAur	2.4	2.34	2.51±0.05	-	-
ηLeo	2.0	1.84	-	-	-
αLep	1.2	0.96	1.7±0.25	1.17±0.06	-
αPer	1.56	1.15	0.95±0.04	1.05±0.03	1.16
γCyg	1.57	1.16	1.18±0.03	1.15±0.04	-
βAqr	0.64	0.54	-	-	-
αAqr	0.41	0.6	-	-	-

a: Keenan & Hynek (1950)
b: Osmer (1972a)
c: Baker (1974)
d: Rao & Mallik (1978)

The spectra of α Per and γ Cyg (Fig. 2) are cases where a large number of nearby blended lines make the local continuum rather hard to define; this may explain the discrepancies between our values of W(O I) and previous results.

5 Discussion

5.1 THE BEHAVIOUR OF THE OXYGEN TRIPLET

A correct interpretation of the oxygen triplet in giants and supergiants demands an understanding of non-LTE effects (Eriksson & Toft 1979). In non-LTE the depth of a line is increased substantially because the lack of collisions in the low density atmospheres allows the lower level to become overpopulated. According to Baschek *et al.* (1977) the departure

coefficients are increased by a factor of 4–10. The effects may extend even to main-sequence stars but Sneden, Lambert & Whitaker (1979) have found none in dwarfs.

In layers close to the surface the source function will be less than the Planck function and the cores of lines formed in this region will be deeper. There will also be curve of growth effects as strong lines (for instance those in A and F supergiants) will be either on the square root or flat parts of the curve of growth and for these the enhancement in equivalent width will be less. The usual microturbulence effect will result in a raised flat part of the curve of growth (and hence the equivalent widths will increase). Baschek *et al.* (1977) point out that the detailed effect of non-LTE is difficult to assess since a large microturbulence results in the cores being formed in deeper layers where thermalization by collisions is more important; on the other hand, the reduced core depth produces a mean radiation field which is more non-locally determined. The radiative transition rates are also changed. There is some indication from the data that the ratio of $W(7772)$ to $W(\mathrm{O\,I})$ increases with spectral type, as would be expected on moving from the flat to the linear region of the curve of growth.

If the equivalent widths of the oxygen triplet are entered (as ordinates) on to the correct curve of growth for a star (i.e. one which takes into account non-LTE, microturbulence and other effects) then the mantissas should give the ratio of the f values, i.e. 0.431:0.301:0.184 (from Wiese, Smith & Glennon 1966). However, if the curve of growth has been estimated from measurements of lines formed in LTE, then the derived f values will be wrong. For the stars in this study the triplet equivalent widths lie on the linear or flat parts of the curve of growth; these regions are raised and steepened in non-LTE since lines can grow deeper before saturation occurs. This results in a larger spread of the equivalent widths.

One simple, though approximate, method of obtaining the LTE curve of growth is to plot experimental widths against empirical strengths, log X, i.e. to perform a differential analysis. This was done for the stars ϵ Aur, α Per, γ Cyg, α Aqr and β Aqr, using equivalent width data from D. J. Stickland (private communication) for ϵ Aur and α Per, Greenstein (1948) for α Per and Osmer (1972b) for α Per, γ Cyg, α Aqr and β Aqr. Empirical strengths were taken from Catchpole, Pagel & Powell (1971). It is admitted that there are several sources of error in this kind of analysis – both in the data and due to the presence of such effects as stratification of the stellar atmosphere. However, the present results indicate that non-LTE effects are important in that the equivalent width ratios are less like 1:1:1 than would otherwise be expected.

5.2 THE BEHAVIOUR OF NEARBY LINES

The use of broadband photometry for measurement of the equivalent width of the O I triplet relies on the absence of strong nearby lines. We see from Fig. 2 that, in fact there is a strong feature in the A and F supergiants to the red of the triplet (at ~ λ 7777.9). This is likely to be the Si I $4p^3P–7s^1P^o$ transition which is listed as a predicted line in Moore, Minnaert & Houtgast (1966), Moore (1967) and Swensson *et al.* (1970). The line seems to be misplaced at λ 7797.56 in Lambert & Warner (1968). The equivalent width of the feature decreases sharply with temperature until the line disappears at G8 – this is consistent with the high excitation (6.08 eV) of the Si I transition. CN lines are strong in the spectra of the cooler stars. In ϵ Gem the blend of $(6–3)Q_2(48)$ and $(7–4)R_1(37)$ at λ 7772.9 is seen between the first two components of the oxygen triplet and pronounced CN lines (together with TiO) appear in α Ori (M2Iab). However, the absence of these and other CN lines in the hotter stars indicates that the likelihood is small of the feature being due, for instance, to the λ 7777.52 $(7–4)Q_1(30)$ CN transition. There is also a suggestion that the Cr I λ 7771.74 line appears in the blue wing of the λ 7772 O I component in α Cyg, ϵ Aur and η Leo.

Finally, we discuss the behaviour of some of the nearby metal lines which can be seen clearly in the diode array spectra. Rao & Mallik (1978) measured the equivalent width of the λ 7748 Fe I line and used the product of W(O I) and W(7748) as a new luminosity indicator, the intention being to eliminate temperature effects as W(O I) should decrease with decreasing temperature whilst W(7748) increases. In our spectra the Fe I line is unresolvably blended with Ni I RMT 156 but the equivalent width of the blend shows the same correlation with temperature. The λ 7780.6 Fe I line has a higher excitation potential (4.45 compared to 2.94 eV) and is not temperature sensitive.

6 Conclusions

It is seen that with a photodiode array good quality spectra can be obtained with a small telescope and poor seeing conditions. The sensitivity in the near infrared enables spectral features to be examined in detail where previously only broadband photometry has been possible. For the O I triplet the existence of nearby lines, both in the intermediate and late type stars, means that these photometric results must be interpreted with care.

It is possible (Athay & Skumanich 1968) that many near infrared lines are formed in conditions of non-LTE and it would seem that, with their high quantum efficiency, solid state detector arrays are eminently suitable for their study.

We note that, although the Plessey arrays are prototype devices and the number of pixels is rather small (512), the good noise performance makes them competitive with other solid state detectors. So far, the array most used for spectroscopy in the near infrared has been the Reticon; this has a higher readout noise (~ 800 electrons) and cannot be read out non-destructively (Vogt, Tull & Kelton 1978).

Acknowledgments

We express our gratitude to the Director and Staff of the RGO for their hospitality during the observations; in particular to Mrs D. L. and Mr C. F. W. Harmer and Dr D. McMullan. Professor A. W. Wolfendale and Dr J. M. Breare are thanked for their interest and support. A. R. Hedge and A. W. Campbell are acknowledged for the design of the system electronics. AH was supported by an SRC studentship. We also wish to thank the referees of this paper for some helpful suggestions.

References

Albers, H., 1969. *Astrophys. J.*, 156, L37.
Athay, R. G. & Skumanich, A., 1968. *Astrophys. J.*, 152, 211.
Baker, P. W., 1974. *Publs astr. Soc. Pacif.*, 86, 33.
Baschek, B., Scholz, M. & Sedlmayr, E., 1977. *Astr. Astrophys.*, 55, 375.
Bracewell, R., 1965. *The Fourier Transform and its Applications*, McGraw-Hill, New York.
Campbell, A. W., Hedge, A. R., Hopkinson, G. R., Humrich, A. & Breare, J. M., 1979. *Adv. Electron. electron Phys.*, 52, 431.
Catchpole, R. M., Pagel, B. E. J. & Powell, A. L. T., 1971. *R. Obs. Bull. No. 154.*
Eriksson, K. & Toft, S. C., 1979. *Astr. Astrophys.*, 71, 178.
Greenstein, J. L., 1948. *Astrophys. J.*, 107, 151.
Griffin, R. F., 1968. *A Photometric Atlas of the Spectrum of Arcturus*, Cambridge Philosophical Society.
Harmer, D. L., Lawson, P. A. & Stickland, D. J., 1978. *Observatory*, 98, 250.
Hedge, A. R., Breare, J. M., Campbell, A. W., Hopkinson, G. R. & Humrich, A., 1978. In *ESO/SRC Conf. on Applications of CAMAC to Astronomy*, p. 210.
Johnson, H. R., Milkey, R. N. & Ramsey, L. W., 1974. *Astrophys. J.*, 187, 147.
Keating, P. N., 1978. *IEEE Trans. on Acoustics, Speech and Signal Processing*, ASSP-26, 368.

BIBLIOGRAPHY

Heywood, R.B.: 1962, "Designing Against Fatigue", Chapman and Hall Ltd., London.

Kaye, G.W.C., & Laby, T.H.: 1960, "Tables of Physical and Chemical Constants," Longmans, London.

Kittel, C.: 1971, "Introduction to Solid State Physics", Wiley and Sons, Inc., New York.

Moss, T.S., Burrell, G.J. & Ellis, B.: 1973, "Semiconductor Opto-Electronics", Butterworths, London.

Polakowski, N.H., & Ripling, E.J.: 1966, "Strength and Structure of Engineering Materials, "Prentice-Hall, Inc., New Jersey.

Weast, R.C. (ed.): 1977, "CRC Handbook of Chemistry and Physics," 58th Ed., CRC Press, Inc., Cleveland.

Keenan, P. C. & Hynek, J. A., 1950. *Astrophys. J.*, **111**, 1.
Lambert, D. L. & Ries, L. M., 1977. *Astrophys. J.*, **217**, 508.
Lambert, D. L. & Warner, B., 1968. *Mon. Not. R. astr. Soc.*, **139**, 35.
Lanczos, C., 1964. *Applied Analysis*, ch. 4, Prentice-Hall, Englewood Cliffs.
Moore, C. E., 1967. *Selected Tables of Atomic Spectra*, NSRDS-NBS3.
Moore, C. E., Minnaert, M. G. J. & Houtgast, J., 1966. *NBS monogr. No. 61.*
Osmer, P. S., 1972a. *Astrophys. J. Suppl.*, **24**, 247.
Osmer, P. S., 1972b. *Astrophys. J. Suppl.*, **24**, 225.
Parsons, S. B., 1964. *Astrophys. J.*, **140**, 853.
Rao, N. K. & Mallik, S. G. V., 1978. *Mon. Not. R. astr. Soc.*, **183**, 211.
Sargent, W. L. & Searle, L., 1962. *Astrophys. J.*, **136**, 408.
Slettebak, A., 1951. *Astrophys. J.*, **113**, 437.
Sneden, C., Lambert, D. L. & Whitaker, R. N., 1979. *Astrophys. J.*, **234**, 964.
Sorvari, J. M., 1974. *Astr. J.*, **79**, 1416.
Swensson, J. W., Benedict, W. S., Delbouille, L. & Roland, G., 1970. *The Solar Spectrum from* λ *7498 to* λ *12016*, Soc. r. des Sciences de Liège, Liège.
Thomas, R. M., Morton, D. C. & Murdin, P. G., 1979. *Mon. Not. R. astr. Soc.*, **188**, 19.
Vogt, S. S., Tull, R. G. & Kelton, P., 1978. *Appl. Opt.*, **17**, 574.
Wiese, W. L., Smith, M. W. & Glennon, B. M., 1966. *Atomic Transition Probabilities*, NSRDS-NBS 4.