IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE

UNIVERSITY OF LONDON

# ATMOSPHERIC TURBULENCE MEASUREMENTS RELEVANT TO ADAPTIVE OPTICS

**Thomas William Nicholls**

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Supervised by Professor John Christopher Dainty,
Applied Optics Group,
Department of Physics

March 1996

# Abstract

This thesis describes the results of an investigation into the effects of atmospheric turbulence on the quality of an astronomical image formed by a large-aperture telescope. Starting with the simple Kolmogorov model of refractive index fluctuations, a well known procedure is reviewed for deriving the spatial and temporal properties of an electromagnetic wave affected by turbulence in its propagation path. The effects are found to manifest themselves mainly in the form of distortions in the phase of the wave incident on the telescope aperture. The variations in the phase perturbations over different temporal and spatial scales can be characterised in terms of the spatio-temporal correlation function of the phase using a very small number of measurable parameters. Relations between the statistics of the output of a Shack-Hartmann wavefront sensor and the statistics of the phase itself are predicted using the Kolmogorov model and the Taylor hypothesis of frozen turbulence; these allow the parameters in question to be measured. Also presented is an evaluation of the experimental methods involved with the use of the wavefront sensor.

Samples of data have been collected using Shack-Hartmann sensors at Mauna Kea (Hawaii), La Palma (Canary Islands) and Calar Alto (Spain). A novel form of the sensor which allows rapid sampling of wavefronts using infrared detectors is described. The data have been used to investigate the properties of the phase distribution in the aperture at both visible and near-infrared wavelengths. The results obtained do not appear to be altogether consistent with the established Kolmogorov-Taylor theory. An approach for characterising non-Kolmogorov turbulence is reviewed and a new method for measuring the extent of deviations from Kolmogorov theory using a Shack-Hartmann sensor is proposed. Theoretical predictions for different models are compared with simulations and experimental studies.

# Acknowledgements

There are two people whose contributions to the work described in this thesis have been invaluable. First of all, I must thank Chris Dainty, without whose willingness always to make time within a busy schedule, whose clear grasp of the science, and whose hard work in securing financial support, much of the good work in the group simply would not happen. Secondly, I thank Nick Wooder, for his endless work in keeping the computer systems running, for his general help and advice, and, particularly, for maintaining a degree of sanity in the proceedings during those long nights in the JKT dome and control room.

I acknowledge the direct contributions of others to this work, of which the largest share comes from Nick Wooder, for building the data aquisition systems used at the JKT, Calar Alto and UKIRT (the latter assisted by Ian Munro) and for writing the centroiding algorithm and routines for accessing data from permanent media. I also thank Martyn Wells at the University of Edinburgh, for the original conception of the coordinate-remapping Shack-Hartmann configuration; Glenn Boreman, for collaboration in the investigation and characterisation of non-Kolmogorov effects; Jean Vernin, for permission to reprint meteorological data; Fred Reavell, for much of the optical design (including the 'Fred widget'); and Paul Brown, Dave Mountain, John Bean and Martin Dowman in the optics workshops, for bringing my rather unorthodox mechanical drawings to life.

On the personal side of things, I shall also take this opportunity to thank my parents and brothers Robert and Richard (for Saturday afternoons at Selhurst Park); my closest friends, Tim, Helen, Ginny, Alastair; all at Holy Joes; the good people of the Optics Subbuteo League: Pedro, for taking the championship back to Mexico, Ian (for helping me rediscover this fine game in Hawaii), Robin, John Hamilton, Keith, Hayden, Dave, John Hodkinson and many others (who would probably prefer to remain nameless); the world-beating optics football team of 1995; Fred, for support after relegation at Newcastle last year; my Mexican friend and fellow beer-critic Victor; and the others who share a desire to stop the stars twinkling: Vince, Miles, Pedro, Ian and Liam.

Finally, I acknowledge the financial contribution of the Royal Observatory, Edinburgh (initiated by Colin Humphries) to this project over the three years of its duration.

# Contents

**List of Figures** 9

**1 The Problem of Atmospheric Turbulence in Astronomy** 11

1.1 Adaptive Optics in Astronomy . . . . . . . . . . . . . . . . . . . . . . . . 14
1.2 Seeing Measurements Relevant to Adaptive Optics . . . . . . . . . . . . 16
1.3 Overview of Thesis . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 18
1.4 Publications . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 19

**2 The Kolmogorov Model of Turbulence** 21

2.1 The Statistics Of Refractive Index Fluctuations . . . . . . . . . . . . . . 25
2.2 Optical Effects: Formulation of Ideas . . . . . . . . . . . . . . . . . . . 29
2.3 Propagation Through a Thin Layer . . . . . . . . . . . . . . . . . . . . . 30
2.4 Effect Of A Thin Layer On The Statistics At Ground Level . . . . . . . 32
2.5 Coherence Function Through Many Layers . . . . . . . . . . . . . . . . . 33
2.6 The Limit Of Atmospheric Resolution . . . . . . . . . . . . . . . . . . . 34
2.7 The Small Perturbation Approximation . . . . . . . . . . . . . . . . . . 37
2.8 The Near-Field Approximation . . . . . . . . . . . . . . . . . . . . . . . 39
2.9 The Restrictions of the Geometric Approach . . . . . . . . . . . . . . . . 40
2.10 The Time Dependence of Seeing . . . . . . . . . . . . . . . . . . . . . . 45
2.11 The Significance of the Greenwood Frequency . . . . . . . . . . . . . . . 48
2.12 The Validity of the Kolmogorov Spectrum . . . . . . . . . . . . . . . . . 49
2.13 Summary of the Major Results of Kolmogorov Theory . . . . . . . . . . 49

**3 Wavefront Sensors** 51

3.1 The Phase as a Wavefront . . . . . . . . . . . . . . . . . . . . . . . . . 51
3.2 Modal Wavefront Representation . . . . . . . . . . . . . . . . . . . . . . 52
3.3 Sensing the Wavefront . . . . . . . . . . . . . . . . . . . . . . . . . . . 54
3.4 Shack-Hartmann Wavefront Sensor . . . . . . . . . . . . . . . . . . . . . 55
3.5 Use of the Shack-Hartmann Sensor in Seeing Measurements . . . . . . . 57
3.6 Angle of Arrival Statistics . . . . . . . . . . . . . . . . . . . . . . . . . 58
3.6.1 The Centroid and Angle-of-Arrival . . . . . . . . . . . . . . . . . . 58
3.6.2 Spatial Covariance . . . . . . . . . . . . . . . . . . . . . . . . . . 59

3.6.3 Single Aperture Variance . . . . . . . . . . . . . . . . . . . . . . . 60
3.6.4 Two-Aperture Differential Variance . . . . . . . . . . . . . . . . . 63
3.6.5 Single Aperture Temporal Power Spectrum . . . . . . . . . . . . 64
3.6.6 Differential Centroid Motion Power Spectrum . . . . . . . . . . . 68
3.7 Performance of the Shack-Hartmann Sensor . . . . . . . . . . . . . . . . 69
3.7.1 Performance Limitations . . . . . . . . . . . . . . . . . . . . . . . . 69
3.7.2 Design of a Shack-Hartmann Sensor with a CCD Array . . . . . 71
3.7.3 Determination of Centroid Accuracy . . . . . . . . . . . . . . . . 74
3.7.4 Error Calculation: A Case Study . . . . . . . . . . . . . . . . . . 75
3.7.5 The Effect Of A Finite Number Of Measurement Samples . . . . 76
3.8 Wavefront Curvature Sensor . . . . . . . . . . . . . . . . . . . . . . . . . 80
3.9 Curvature Sensing vs Shack-Hartmann: A Comparison . . . . . . . . . . 83

**4 The Shack-Hartmann Wavefront Sensor: Algorithms for Measuring Turbulence 86**

4.1 Seeing Measurements Relevant to Adaptive Optics . . . . . . . . . . . . 86
4.2 Determination of Centroid Positions . . . . . . . . . . . . . . . . . . . . 88
4.2.1 Short Exposures as a Method of Sampling . . . . . . . . . . . . . 88
4.2.2 Cleaning Images . . . . . . . . . . . . . . . . . . . . . . . . . . . . 89
4.2.3 Centroiding Algorithm . . . . . . . . . . . . . . . . . . . . . . . . 89
4.3 Statistical Analysis of Centroids . . . . . . . . . . . . . . . . . . . . . . 91
4.3.1 Implementation of Algorithms . . . . . . . . . . . . . . . . . . . 92
4.3.2 Statistical Error Analysis . . . . . . . . . . . . . . . . . . . . . . 93

**5 Seeing Measurements 95**

5.1 Overview of Experiments . . . . . . . . . . . . . . . . . . . . . . . . . . 95
5.2 High-Precision Measurement Of Spatial Wavefront Information . . . . . 96
5.2.1 Objectives of the Visible Shack-Hartmann Observations . . . . . 96
5.2.2 Instrumentation . . . . . . . . . . . . . . . . . . . . . . . . . . . . 97
5.3 Temporally Resolved Infrared Wavefront Measurements . . . . . . . . . 99
5.3.1 The Difficulties of Infrared Seeing Measurements . . . . . . . . . 99
5.3.2 A Modified Shack-Hartmann Sensor . . . . . . . . . . . . . . . . 101
5.3.3 Data Acquisition System . . . . . . . . . . . . . . . . . . . . . . . 104
5.4 Temporally Resolved Visible Wavefront Measurements . . . . . . . . . . 105
5.4.1 Description Of The System . . . . . . . . . . . . . . . . . . . . . . 105
5.5 Summary of Data Acquired . . . . . . . . . . . . . . . . . . . . . . . . . 107
5.6 Analysis of Correlation of Measurements . . . . . . . . . . . . . . . . . . 108
5.7 $r_0$ Variation over the Aperture . . . . . . . . . . . . . . . . . . . . . . 111
5.8 $r_0$ Evolution Over Time . . . . . . . . . . . . . . . . . . . . . . . . . . 117
5.9 $r_0$ Measured for Different Sky Positions . . . . . . . . . . . . . . . . . 122

5.10 Centroid Power Spectra . . . . . . . . . . . . . . . . . . . . . . . . . 124

**6 Non-Kolmorogov Turbulence** **131**

6.1 Evidence for a non-Kolmogorov Structure Function . . . . . . . . . . . 131

6.2 Consequences of a non-Kolmogorov Power Law . . . . . . . . . . . . . 132

6.3 Measurement of a non-Kolmogorov Power Law . . . . . . . . . . . . . 133

**7 Concluding Remarks** **139**

**A Error Analysis Formulae** **141**

A.1 Error Resulting from Independent Estimates . . . . . . . . . . . . . . . 141

A.2 Error Resulting from Correlated Estimates . . . . . . . . . . . . . . . . 141

A.3 Error Propagation Formulae . . . . . . . . . . . . . . . . . . . . . . . . 143

# List of Figures


1.1 Examples of long- and short-exposure images. . . . . . . . . . . . . . . . 12

1.2 Schematic diagram showing the principal components of an adaptive optics system. . . . . . . . . . . . . . . . . . . . . . . . . . . 15

2.1 Graphs showing the turbulence profile $C_n^2$ verses height measured at four different times at the La Palma site. . . . . . . . . . . . . . . . . . . . . 28

2.2 OTF for a circular aperture and for the atmosphere under various conditions of seeing. . . . . . . . . . . . . . . . . . . . . . . . . . . 36

3.1 Principle of the Shack-Hartmann Wavefront Sensor. . . . . . . . . . . . 56

3.2 A comparison between different methods of calculating the two-aperture differential variance ratio for different aperture separations. . . . . . . . 65

3.3 Graph showing the form of the temporal power spectrum of the centroid motions of a single circular aperture for Kolmogorov turbulence. . . . . 66

3.4 Graph showing the form of the temporal power spectrum of the differential centroid motions of pairs of circular apertures with different separations for Kolmogorov turbulence. . . . . . . . . . . . . . . . . . . 69

3.5 Diagram showing how the telescope pupil is reimaged to a size which is appropriate for the pixel size of the detector. . . . . . . . . . . . . . . . 72

3.6 Diagram showing the array of pixels which make up a CCD detector and suitable separations for neighbouring Shack-Hartmann images. . . . . . 73

3.7 Diagram showing the process involved in simulating a Gaussian image formed on a discrete array of detector elements. . . . . . . . . . . . . . . 76

3.8 Graph showing the relative centroid accuracy as a function of the number of photons detected per subaperture image by the Princeton CCD camera, calculated for photon noise and readout noise. . . . . . . . . . . 77

3.9 The procedure for estimating the probability distribution of the variance of a Gaussian variable calculated using a finite number of independent measurements. . . . . . . . . . . . . . . . . . . . . . . . . . . . 78

3.10 Diagram showing the reasons why a measurement of a centroid variance leads to a bias in the value of $r_0$. . . . . . . . . . . . . . . . . . . . . . 79

3.11 Expected bias on $r_0$ through measurement using a finite number of frames. 80

3.12 Ray optics representation of a single defocused image. . . . . . . . . . . 81
3.13 Principle of the curvature sensor. . . . . . . . . . . . . . . . . . . . . . . . 83

4.1 Modular diagram showing the major parts of the centroiding algorithm. 90
4.2 Modular diagram showing the centroid statistics processing algorithm. . 93

5.1 Diagram of the simultaneous Shack-Hartmann wavefront sensor and speckle imager used at the JKT in La Palma. . . . . . . . . . . . . . . . . . . . 98
5.2 A short-exposure image acquired with the Shack-Hartmann/Speckle imaging system at the JKT. . . . . . . . . . . . . . . . . . . . . . . . . . . . 99
5.3 Principle of the standard Shack-Hartmann wavefront sensor, showing the unused space between subaperture spots. . . . . . . . . . . . . . . . . . 100
5.4 Principle of the coordinate-remapping Shack-Hartmann wavefront sensor. 101
5.5 Scale diagram of the coordinate remapping Shack-Hartmann sensor and corresponding locations of subaperture images on the detector array. . . 103
5.6 Multiple-exposure image and a single image acquired at UKIRT using the 256x48 pixel readout mode. . . . . . . . . . . . . . . . . . . . . . . . 104
5.7 Multiple-exposure image and a single image acquired at UKIRT using the 64x64 pixel readout mode.. . . . . . . . . . . . . . . . . . . . . . . . 105
5.8 Multiple-exposure image acquired at Calar Alto. . . . . . . . . . . . . . 106
5.9 Temporal covariance of the centroid position difference for one selected pair of subaperture images from data acquired at UKIRT. . . . . . . . . 109
5.10 Temporal covariance of the centroid position difference for one selected pair of subaperture images from data acquired at Calar Alto. . . . . . . 110
5.11 Histograms of $r_0$ calculated for all pairs of subapertures in the pupil for one data set acquired at La Palma. . . . . . . . . . . . . . . . . . . . . . 112
5.12 Histograms of $r_0$ calculated for all pairs of subapertures in the pupil for one data set acquired at UKIRT. . . . . . . . . . . . . . . . . . . . . . . 113
5.13 Histograms of $r_0$ calculated for all pairs of subapertures in the pupil for one data set acquired at Calar Alto. . . . . . . . . . . . . . . . . . . . . 113
5.14 $r_0$ plotted against subaperture separation (La Palma). . . . . . . . . . . 115
5.15 $r_0$ plotted against subaperture separation (UKIRT). . . . . . . . . . . . 116
5.16 $r_0$ plotted against subaperture separation (Calar Alto). . . . . . . . . . . 116
5.17 $r_0$ and temperature plotted against time for a data set acquired at La Palma. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 118
5.18 $r_0$ plotted against time for a second set acquired at La Palma. . . . . . . 119
5.19 $r_0$ plotted against time for three consecutive data sets acquired at UKIRT.120
5.20 $r_0$ plotted against time for three data sets acquired at different times on the same night at Calar Alto. . . . . . . . . . . . . . . . . . . . . . . . . 121

5.21 Seeing around the sky measured at UKIRT over 90 minutes on 3rd August 1994, also shown compensated for the declination angle assuming Kolmogorov turbulence. . . . . . . . . . . . . . . . . . . . . . . . . . 122
5.22 Seeing around the sky measured at Calar Alto over 120 minutes on 13th July 1994, also shown compensated for the declination angle assuming Kolmogorov turbulence. . . . . . . . . . . . . . . . . . . . . . . . . . 123
5.23 Centroid position in two directions in the sky (parallel and orthogonal to tracking) plotted against time (set E: UKIRT). . . . . . . . . . . . . . 125
5.24 Temporal Power Spectrum (low frequency end) of centroid motions in the direction of tracking and in the orthogonal direction (set E: UKIRT). 126
5.25 Temporal power spectra of centroid motions measured in the direction orthogonal to tracking for 4 different data sets acquired at UKIRT. . . . 127
5.26 Temporal power spectra of differential centroid motions for 4 different data sets acquired at UKIRT. . . . . . . . . . . . . . . . . . . . . . . . 128
5.27 Temporal power spectra of centroid motions measured in an arbitrary direction for 3 different data sets acquired at Calar Alto. . . . . . . . . . 129
5.28 Temporal power spectra of differential centroid motions for 3 different data sets acquiredat Calar Alto. . . . . . . . . . . . . . . . . . . . . . . 130

6.1 Theoretical dependence of the ratio of transverse to longitudinal differential angle-of-arrival variance on the exponent of the power spectrum of phase. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 135
6.2 Comparison of the simulated centroid variance ratio with theoretical calculations. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 136
6.3 The ratio of differential variances found in experimental data acquired at the JKT compared with theoretical curves for varying phase power exponent. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137
6.4 The ratio of differential variances found in experimental data acquired at Calar Alto compared with theoretical curves for varying phase power exponent. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 137

# List of Tables

3.1 Relation between Zernike numbering coefficient i and the radial degree and azimuthal coefficient. . . . . 53

5.1 Table of observations made and their respective sites. . . . . 96
5.2 Summary of data sets for which centroids have been found. . . . . 108
5.3 Summary of the spatial correlation factors found for various sets of data. 111
5.4 Errors estimated for $r_0$ measurements taken with individual subaperture pairs compared with the expected values. . . . . 114

# Chapter 1

# The Problem of Atmospheric Turbulence in Astronomy

For some time, astronomers have attempted to understand and correct for the effects of turbulence on the passage of light through the atmosphere. It has long been known that the effects of the atmosphere impose a limit on the maximum resolution obtainable with a large telescope, a limit which cannot be removed simply by increasing the diameter of the telescope. Even under the very best conditions available, the World's largest telescope is reduced to the performance, in terms of resolution, of a 10-20cm diameter telescope, when a long-exposure image, obtained at visible wavelengths, is examined.

Astronomy has continually demanded the construction of larger telescopes because the study of faint objects requires the maximum light collecting capability. However, the higher resolving power which ought to accompany this has not been available due to degradation of the image quality caused by the atmosphere, which manifests itself in two ways. The image is seen to 'wander' about a central position; over the course of an exposure taken over a few seconds of time, the effect of this motion is to blur out the image into a characteristic 'seeing disc'. If a short enough exposure is taken (typically a fraction of a second), then this motion may be 'frozen'. Higher order image-degradation effects become increasingly more significant for telescopes with larger apertures. These are generally referred to as 'image blurring', their effect being to cause the wandering image to break up further into a rapidly fluctuating speckle pattern. Examples of long- and short- exposure images are shown in Fig. 1.1. Turbulence is also responsible for the so-called 'twinkling' of stars. Over small sized apertures (*e.g.* the human eye), there is a temporal variation in the overall incident intensity due to the focusing effects of turbulence in the higher part of the atmosphere.

A long-exposure image of a point source (*e.g.* an unresolvable star), is found to be $(\lambda/r_0)$ in angular extent, where $\lambda$ is the wavelength of the light detected and $r_0$ is a

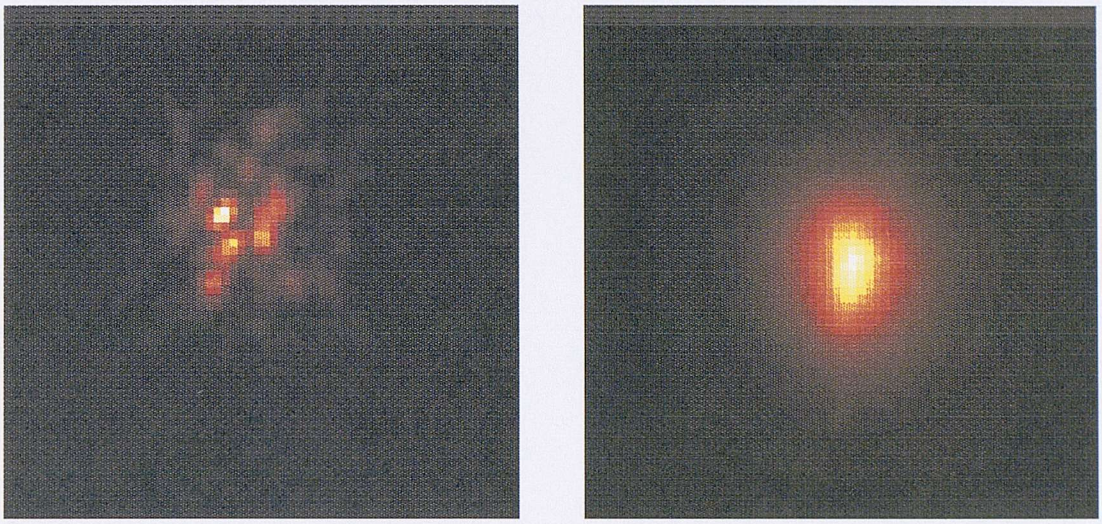


**Figure 1.1**: Examples of long- and short-exposure images acquired at the 3.8-metre UK Infrared Telescope at Mauna Kea, Hawaii (at infrared wavelengths). Left: a single short-exposure image, exhibiting a speckle structure; right: the sum of 2500 short-exposure images, in which the seeing disc is visible.

wavelength-dependent atmospheric parameter which fluctuates in time and can reach 20cm at visible wavelengths at the World's best sites. $r_0$ is a measure of the strength of the turbulence, or, alternatively, it can be thought of as the maximum diameter of telescope which can support diffraction-limited imaging (with an appropriate definition; see section 2.6) under the atmospheric conditions encountered at the time in question. This image size compares very poorly with the $(\lambda/D)$ theoretical diffraction limit, where $D$ is the diameter of the telescopic aperture, and is 10 metres at some telescopes currently under construction. Correction of the effects of turbulence could then, potentially, lead to a fifty-fold increase in the resolution of astronomical images.

A very short exposure (typically a few milliseconds in length) would be capable of freezing the speckle image, the size of these speckles being of the order of the diffraction limit of the telescope. Labeyrie [1] first recognised that this was high spatial frequency information which could be used to improve on the seeing. The long-exposure image (which is effectively the integration of a large number of speckle patterns produced by different realisations of the atmosphere) is a blurred-out speckle pattern in which all of the high frequency information is lost. However, the study of astronomical objects, in general, requires the collection of photons over long periods of time due to the faint magnitudes involved.

The image quality, which is characterised in terms of the spatial and temporal properties of the degradation introduced into the optical wave traversing the atmosphere, is known to vary over time. This variation is over and above the rapid fluctuations already described which cause image motion and blurring, and which can be considered to be the result of different realisations of a particular statistical description of the atmosphere. The strength of the turbulence over a period of time which gives rise to a corresponding image quality is generally referred to as the 'seeing'. Good seeing generally corresponds to a long-exposure image with a relatively high resolution; poor seeing refers to the opposite case.

An obvious method of tackling these problems is to place a telescope above the atmosphere; the Hubble Space Telescope has demonstrated that this is viable but also extremely expensive and involves technical risk. Furthermore, only a limited number of observations can be made during its lifetime, and the technologies involved often lag behind due to the delay between design and launch of such a system.

Many attempts have been made to reduce the effects of turbulence using conventional ground-based telescopes, some of which are now widely used for enhancing images. They can be broadly divided into two categories: *post-detection processing* and *real-time* techniques. The former involve taking short-exposure images (short enough to 'freeze' the turbulent image), although the resulting lack of photons will inevitably lead to a low signal-to-noise level in each image. By taking large numbers of such images, the intention is to use the small amount of diffraction-limited information available in each of them to try to reconstruct an image. Perhaps the simplest of these techniques involves freezing only the image motion in each exposure, finding the centre-of-gravity of the intensity distribution, re-centering each image and, finally, superimposing to increase the light level. Such centroiding techniques cannot correct for the higher order blurring effects, however (although 'shift-and-add' techniques, where images are centred to their brightest speckle, can eliminate some of these effects). Other techniques, such as speckle interferometry and bispectral imaging [2] [3] [4] [5], employ statistical methods in an attempt to reconstruct the object more accurately. If some extra information can be obtained about the phase of the turbulence-degraded wave (the 'wavefront') at the same time (by use of a wavefront sensor for example), the processing of these images is assisted [6]. Alternatively, pupil masking techniques ('aperture synthesis' [7] [8]), which are immune to the effects of speckle noise, can be used to extract information about the object.

Post-detection techniques are generally limited to objects of bright magnitudes, since the fast exposure rate limits the photon count in each image. Furthermore, they are applicable only for simple imaging of objects concerned. Astronomical tools such as spectroscopy require a high-resolution image to be formed at the time of detection, so that the light from a part or parts of the object can be studied in detail. Techniques which correct for atmospheric turbulence in real-time are referred to, in general, as adaptive optics. A very simple example would involve performing the shift-and-add technique in real-time by, say, recentering the images using a tilting mirror. Indeed, adaptive optics systems are often designed to correct for image motion independently of image blurring.

## 1.1 Adaptive Optics in Astronomy

Adaptive optics relies on a technique known as *phase conjugation*; that is, the process of removing distortions in the phase geometry which have been introduced by the turbulence. Many of the world's observatories are currently embarking on programmes to install adaptive optics (AO) systems on new or existing large-aperture telescopes. Although the fundamental designs of such systems vary considerably from one project to another, each system must perform certain essential tasks. The first is to sample the turbulence, *i.e.* to measure the degradations in the phase caused by it, using a *wavefront sensor*. This requires a *reference* wave, whose phase geometry before encountering the turbulence is, to some extent, known. If the astronomical object (the 'science object') is too dim to provide this, a bright star may be used (whose wavefront is effectively flat before entering the atmosphere) to generate the reference wave. If no such star is available, an artificial beacon (manufactured by firing a laser into the atmosphere and observed by collecting scattered or resonantly emitted light) may be used; this is a rather expensive option due to the extremely high-power laser which is required. In any case, there will always be a finite supply of light from the reference source which must be used to estimate the phase distortions. The turbulence information is then immediately used to perform the compensation. In real time, a medium, such as a deformable mirror, must be altered in some way to introduce distortions which to some degree cancel those observed in the reference wave. Currently, only phase aberrations can be corrected for. Hence the four major components of every AO system are: the reference star, the wavefront sensor, the processing algorithm to convert wavefront sensor output to control over the phase corrector, and the phase corrector itself (see Fig. 1.2).

Ideally, one would wish to provide sufficient compensation of the phase distortions so as to achieve perfect diffraction limited imaging at a large telescope under all seeing conditions which occur at the site. However, there are certain fundamental limitations to the degree of compensation which can be performed, as will be discussed below. The performance of an adaptive optics system is often measured in terms of the *Strehl ratio*, which is the ratio between the maximum intensity (strictly the *central* intensity) obtained in a compensated image and that which would be expected in a diffraction limited image. The Strehl ratio achievable by a proposed system can be estimated by modelling its component parts, if the typical seeing quality encountered at the site is known. Clearly, during poor seeing a greater degree of compensation is necessary to

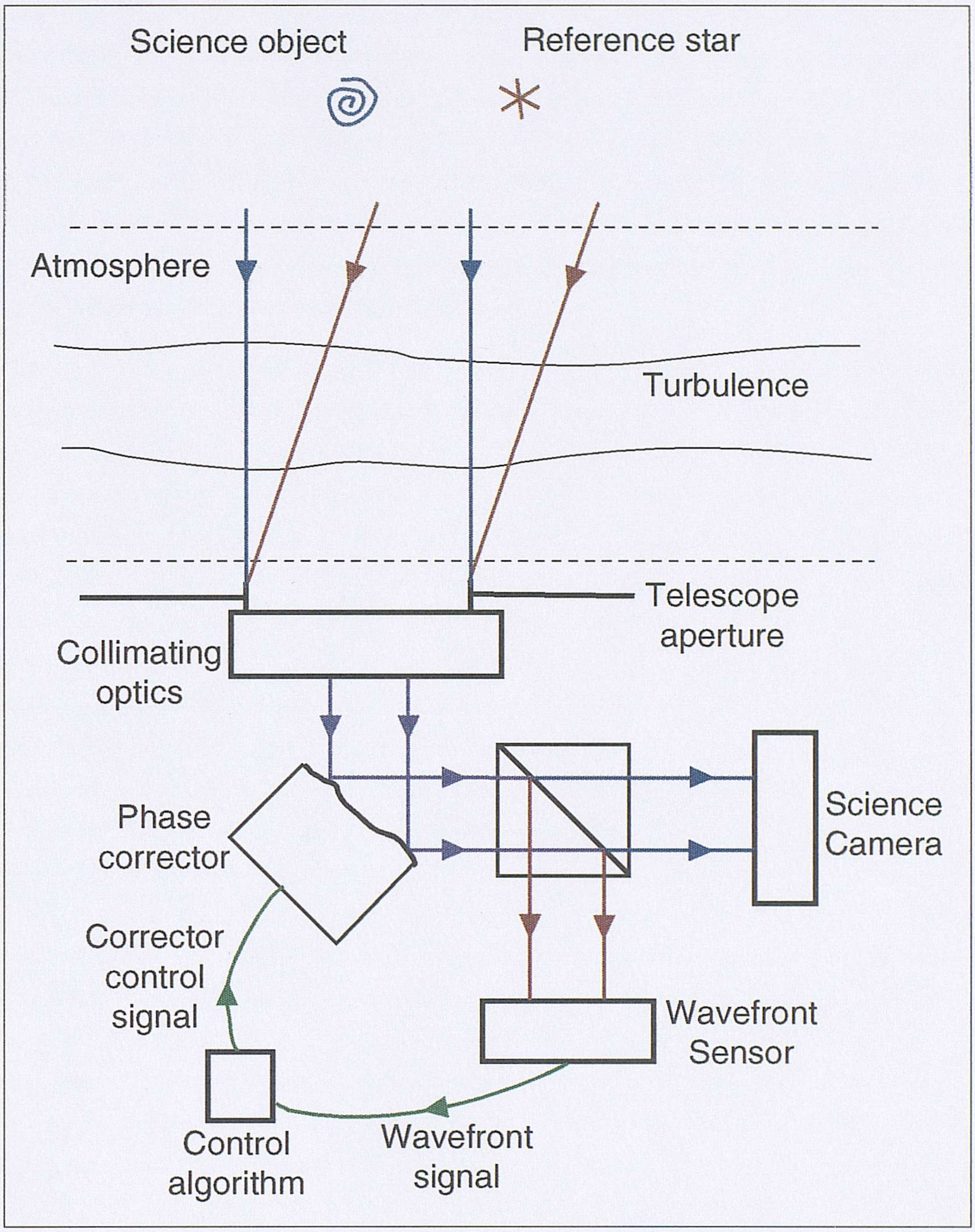



**Figure 1.2**: Schematic diagram showing the principal components of an adaptive optics system.

obtain the same performance.

Any difference between the degraded phase distribution and the (reverse of the) phase distortions introduced in an attempt to correct for it is generally referred to as a *residual*. The Strehl ratio obtainable is directly calculable if the rms residual is known. A residual will occur if the components in the system are unable to function in the ideal way. Clearly, from mechanical considerations, it may not be possible to distort the correcting medium into the exact shape required, and there certainly will be a limiting resolution to which it is possible to sample the incident wavefront. The ability to sense the wavefront depends on the availability of a bright reference source. If too few photons are available, then an inaccurate wavefront will be measured. A lack of photons from the reference source is the factor which will fundamentally limit the performance of any adaptive optics system. Laser beacons are being considered in order to provide bright reference sources anywhere in the sky.

Some aspects of a system which may lead to a residual may be grouped under the general term of *anisoplanatism*, which refers in this instance to the difference between what is measured by the wavefront sensor (*i.e.* the wave incident on the sensor) and the instantaneous state of the wave from the science object. Because correction is made at discrete intervals in time, a residual will be caused by the evolution of the turbulence-degraded phase distribution between successive corrections, a phenomenon known as temporal anisoplanatism. This can also occur if there is a delay, or lag, between sampling the wavefront and providing the correction (due to the finite time taken to convert wavefront sensor signals into corrector control, and any finite response time in the corrector itself). In order to compensate for this, correction might be made more frequently in time. Even if the technologies involved allow this, further problems arise because it will reduce the number of photons available to the wavefront sensor for estimating the phase distortions during each correction cycle. In this case, a brighter star might be chosen for the reference source, if available. The fact that bright stars are less abundant than dim ones means that in all probability, there will be a larger angle between reference star and science object, so that the two sources are now traversing different paths through the atmosphere (as shown in Fig. 1.2). The correlation between the phase distortions introduced into waves propagating along the two paths decreases as the angle between them increases, an effect known as angular anisoplanatism. An 'acceptable' angle (for which the residual remains sufficiently small to provide an acceptable Strehl ratio) is referred to as the isoplanatic angle. The sky coverage is a measure of the proportion of the sky which exists within one isoplanatic angle in distance from any one of the reference stars which are considered bright enough to be useful. A hypothetical science object would need to lie within this region in order for compensation to be available.

## 1.2 Seeing Measurements Relevant to Adaptive Optics

An appropriate mathematical model of the atmosphere would enable all of these effects to be taken into account and the performance of an adaptive optics system to be predicted. Clearly, a complete description of the entire atmosphere within the propagation path in terms of its instantaneous refractive index distribution (for this is what causes the image degradation) is impractical. However, if a statistical description can be given, so that the characterisation of the distortions introduced may be given in terms of a small number of manageable parameters (such as $r_0$), then an estimate of the likely performance of an adaptive optics system may be made. A prerequisite to building any AO system is a programme of *seeing measurements*, where the implications of the night-to-night observing conditions on the design of the system can be assessed. The purpose of such measurements would be to answer certain very general questions: Does the statistical description of the turbulence remain constant during the period of a typical astronomical exposure? If the statistics of seeing do vary in time, then how quickly does this happen and what is the full range and distribution of the conditions which can be expected? Can the fluctuations in the atmosphere be described in terms of a simple statistical model which can be described completely in terms of parameters?

It is important that a knowledge of the statistics of the seeing is gained so that the performance of the AO system can be predicted over the full range of encountered conditions. For example, more specifically, it is vital to know the maximum stroke which must be applied to the different sections of a deformable mirror in order to match the phase distortions introduced by the atmosphere. Also, the characteristic temporal properties of the turbulence must be measured, so that the bandwidth of the system (*i.e.* how regularly the compensation cycle must be applied to prevent unacceptable loss of performance) may be chosen. It must be determined whether there are a suitable number of sufficiently bright natural reference stars to provide the required sky coverage for a particular system design. Failing this, one must consider using a laser beacon.

Seeing measurements are usually made using a wavefront sensor, which may be the same type of sensor as used to measure wavefronts for adaptive compensation. In an adaptive optics system, at the start of a correction cycle, the wavefront sensor measures only the residual wavefront which has evolved since correction was made during the previous cycle. Because of this, such systems are described as being *closed-loop*. For seeing measurements, where the full effects of turbulence are measured, the wavefront sensing is of the open-loop variety. In this case, the measurements may be recorded and processed at a later time. Closing the loop (*i.e.* turning on the correction) should reduce the magnitude and alter the statistics of the distortions being measured. The wavefront information must be processed in real-time in order to facilitate correction.

Almost all models of turbulence are based on the work of Tatarski [9], who took a simple model of refractive index fluctuations proposed by Kolmogorov and predicted how the passage of light would be affected by it. Indeed, the phrase *Kolmogorov turbulence* refers to the type of atmosphere which produces a complex amplitude in the telescope pupil which conforms to the predictions of this model. The label 'Kolmogorov' may also be used to describe the fluctuations in the complex amplitude itself. The major triumph of Kolmogorov theory is that the turbulence can be described in terms of a small number of parameters; indeed the statistical behaviour of the turbulence is well defined and is merely weighted by these parameters. The large volume of information which has been obtained about the subject since this theory was first developed demonstrates that it may not be altogether correct (see section 6.1). However, it may still be sufficiently accurate to remain a useful tool in predicting the outcome of any compensation technique, at least to a first order approximation.

## 1.3 Overview of Thesis

This thesis primarily deals with the effects of turbulence on astronomical imaging and its relevance to the construction of adaptive optics systems. As is to be expected, a large proportion of this work amounts to a critical review of the foundations of the subject of interest. However, a number of its parts can, to the knowledge of this author, be considered original in nature.

- Chapter 1 has provided a general introduction to the concepts and terms involved in adaptive optics.
- Chapter 2 reviews the derivation of most of the major results of Kolmogorov theory from (fairly) basic principles. The statistical behaviour of the turbulence-degraded phase at the observation point is calculated for a model Kolmogorov atmosphere.
- Chapter 3 describes methods of measuring the turbulent wavefront at the telescope pupil using wavefront sensors and how to relate the output of the Shack-Hartmann wavefront sensor to the statistics of the phase itself. These relations are derived for Kolmogorov turbulence and are based on the work of Tatarski [9] and others, although some of the methods involved in the derivations are believed to be original. The procedures which must be observed in order to measure the statistics of seeing with the prescribed accuracy are also presented.
- Chapter 4 discusses the important implications of any variations in the seeing conditions for adaptive optics. Algorithms used for calculating the relevant statistics are given in outline.
- Chapter 5 describes experiments carried out at three observatories to measure the seeing quality. Typical results are shown and the consequences of these experiments with regard to long-term site monitoring campaigns are discussed.

- Chapter 6 describes some original attempts to characterise and measure the effects of non-Kolmogorov turbulence on the statistics of phase and its relevance to predicting the performance of adaptive optics.

## 1.4 Publications

There follows a list of publications, which have been submitted to journals and conferences and to which this author has made a contribution:

1. T.W. Nicholls, C.J.Solomon, N.J.Wooder and J.C.Dainty, "Experiments on Wavefront Sensing at La Palma," in *Adaptive Optics in Astronomy*, Proceedings of the Society of Photo-Optical Instrumentation Engineers. **2021**, 555–561, Kona, 17-18 March (1994).

2. T.W. Nicholls, N.J.Wooder, J-Y.Zhang and J.C.Dainty, "Experiments on Wavefront Sensing for Astronomical Imaging," in *Frontiers in Information Optics.* 97 (conference digest), Kyoto, 4-8 April, (1994).

3. T.W. Nicholls, M.Wells, N.J.Wooder and J.C.Dainty, "Infrared Wavefronts on Mauna Kea", in *Adaptive Optical Systems and Applications*, Proceedings of the Society of Photo-Optical Instrumentation Engineers. **2534**, 257–264, San Diego, 10-11 July (1995).

4. N.J.Wooder, I.Munro, T.W.Nicholls, M.Wells and J.C.Dainty, "Space-time Wavefront Analysis using a Coordinate Remapping Shack-Hartmann Sensor," in *OSA Topical Meeting on Adaptive Optics.* 56–58 (conference digest), Garching, 2-6 October (1995).

5. T.W. Nicholls, G.D.Boreman and J.C.Dainty, "Use of a Shack-Hartmann Wavefront Sensor to Measure Deviations from a Kolmogorov Phase Spectrum," *Optics Letters.* **20**, 2460–2462 (1995).

6. T.W. Nicholls, N.J.Wooder and J.C.Dainty, "Measurement of a Non-Kolmogorov Structure Function," submitted to *OSA Topical Meeting on Adaptive Optics.* Maui, 8-12 July (1996).

7. V.A.Kluckers, N.J.Wooder, T.W.Nicholls and J.C.Dainty, "Results from SCIDAR observations," submitted to *Image Propagation through the Atmosphere.* Denver, 4-8 August (1996).

8. J.C.Dainty, M.J.Adcock, V.A.Kluckers, T.W.Nicholls and N.J.Wooder, "Estimation of Optical Turbulence Profiles using SCIDAR," submitted to *ICO XVII: Optics for Science and New Technology.* Taejon, 19-23 August (1996).

9. N.J.Wooder, I.Munro, T.W.Nicholls, M.Wells and J.C.Dainty, "A Coordinate Remapping Shack-Hartmann Sensor for Use at UKIRT," to be submitted to *Monthly Notices of the Royal Astronomical Society* (1996).

# Chapter 2

# The Kolmogorov Model of Turbulence

When considering the effect of atmospheric turbulence on the imaging properties of a large ground based telescope, a model must be constructed which explains the mechanisms by which the image produced at the telescope focus is affected. Some measure should be available of the extent that the turbulence degrades the image quality. The performance of a telescope can conveniently be characterised in terms of its Optical Transfer Function (OTF), which is a measure of the energy transmitted at each spatial frequency by the telescope when illuminated by a point source at infinity (*i.e.* a non-resolvable quasi-monochromatic star). The task here is to find the mechanisms of turbulence responsible for the changes observed in the transfer function of a large telescope.

Atmospheric effects can be accounted for by considering the complex amplitude distribution ($\Psi(\mathbf{r})$) across the telescope pupil. The OTF can then be formally defined as follows [17]:

$$\mathcal{T}(f) = \frac{\int\int_{-\infty}^{\infty} P(\mathbf{r})\, P^*(\mathbf{r}-\gamma f)\, \Psi(\mathbf{r})\, \Psi^*(\mathbf{r}-\gamma f)\, d^2\mathbf{r}}{\int\int_{-\infty}^{\infty} |P(\mathbf{r})|^2 |\Psi(\mathbf{r})|^2 d^2\mathbf{r}} \tag{2.1}$$

where $P(\mathbf{r})$ is the pupil function (*e.g.* 1 inside the aperture and 0 outside where there are no telescope aberrations), $f$ is the spatial frequency, and $\gamma = \lambda l$ where $l$ is the effective focal length of the telescope and $\lambda$ is the wavelength. In the absence of turbulence, a coherent point source at infinity would produce a uniform intensity and phase across the telescope pupil and the OTF would be governed only by the size and the aberrations of the telescope. The transfer function in this case would be $\mathcal{T}_t(f)$, the OTF of the telescope system. When turbulence is introduced, the transfer function $\mathcal{T}_{atm}(f)$ of the atmosphere must also be considered. For a long exposure (which is normally used for imaging faint astronomical sources), the OTF is given by its average value:

$$\langle \mathcal{T}(f) \rangle = \left\langle \frac{\int\int_{-\infty}^{\infty} P(\mathbf{r}) P^*(\mathbf{r} - \gamma f) \Psi(\mathbf{r}) \Psi^*(\mathbf{r} - \gamma f) d^2\mathbf{r}}{\int\int_{-\infty}^{\infty} |P(\mathbf{r})|^2 |\Psi(\mathbf{r})|^2 d^2\mathbf{r}} \right\rangle . \tag{2.2}$$

This equation cannot be easily simplified further without a complete knowledge of the statistics of the atmosphere and the telescope aberrations. However, in the case of no intensity fluctuations and for *stationary* and *isotropic* phase fluctuations, it can be shown that the OTF can be written as:

$$\begin{aligned} \langle \mathcal{T}(f) \rangle &= \frac{\int\int_{-\infty}^{\infty} P(\mathbf{r}) P^*(\mathbf{r} - \gamma f) \langle \Psi(\mathbf{r}) \Psi^*(\mathbf{r} - \gamma f) d^2\mathbf{r} \rangle}{\int\int_{-\infty}^{\infty} |P(\mathbf{r})|^2 d^2\mathbf{r}} \\ &= \mathcal{T}_t(f) \, \mathcal{T}_{atm}(f) . \end{aligned} \tag{2.3}$$

Thus, in this rather limiting approximation, the effects of the atmosphere and of the telescope system can be considered independently. The validity of the approximation of no scintillation will be discussed in section 2.8; stationarity and isotropy will be considered in section 2.1.

The propagation of an electromagnetic wave from a star can be considered to consist of three distinct phases. The first is a propagation across free space. The second is the journey through the atmosphere, at the start of which the wave can effectively be considered to be plane, and during which the wave-function is somehow altered. The third is the process of collecting and imaging the light which occurs inside the telescope. In fact, the complex amplitude of such a wave can be generalised to any height, h, in the atmosphere (with h=0 at the telescope pupil),

$$\Psi(h, \mathbf{r}) = |\Psi(h, \mathbf{r})| e^{i\varphi(h,\mathbf{r})}. \tag{2.4}$$

It could also be written as:

$$\Psi(h, \mathbf{r}) = e^{\gamma(h,\mathbf{r}) + i\varphi(h,\mathbf{r})}. \tag{2.5}$$

Clearly, it is fluctuations in the log amplitude ($\gamma$) and the phase ($\varphi$), both spatial and temporal, which are of the biggest concern to astronomers when considering the effects of turbulence. It is important to recognise that the phase and the log amplitude have no meaning if the intensity $|\Psi(h, \mathbf{r})|^2$ vanishes to zero. In this case, the intensity would not be a continuous function and therefore would render much of the theoretical analysis invalid. However, this type of strong intensity variation is known not to occur in the application of astronomical imaging through turbulence and it will therefore be possible to proceed within a regime which allows only small perturbations on the complex amplitude. Fluctuations in the log amplitude, both in time and over space, are generally referred to as scintillation.

The only conceivable way in which turbulence could alter the complex amplitude of a wave incident on the atmosphere is via fluctuations in the refractive index from place to place and from time to time within the atmosphere. These refractive index fluctuations produce deviations in the directions of the incoming rays and also changes in the optical path lengths. A well developed theory exists, based on a simple atmospheric model suggested by Kolmogorov, for propagation of a plane electromagnetic wave through a turbulent atmosphere. Ultimately, it is crucial to be able to directly relate the index fluctuations to the fluctuations in the complex amplitude at the telescope pupil. Although it is impossible to predict the exact form of the complex amplitude at a particular time, without a complete knowledge of the refractive index distribution within the whole propagation path, it is, however, possible to understand the nature of its time-averaged statistics. There are any number of ways in which these statistics may be quantified. However, certain statistical quantities are particularly suited to describing the parameters involved in optical propagation through turbulence and some of these are listed here:

- **Correlation Function**

  The spatio-temporal correlation function of a statistical quantity $\chi$ is perhaps its most fundamental description:

  $$B_{\chi}(\Delta\boldsymbol{\rho},\Delta t) = \langle \chi(\boldsymbol{\rho},t)\,\chi(\boldsymbol{\rho}+\Delta\boldsymbol{\rho},t+\Delta t)\rangle\,, \tag{2.6}$$

  where $\boldsymbol{\rho}$ is a vector in three-dimensional space, t is a measure of time elapsed in the system and the angled brackets $\langle\rangle$ denote an integral over all time and all space.

- **Structure Function**

  When considering optical phenomena it is often the difference of an isotropic quantity between two points and/or times that is the most meaningful. For quantities such as the wavefront phase, or the refractive index, an absolute measure has little or no significance, and it is fluctuations that are important. Hence the structure function of the quantity, $\chi$, is defined as follows:

  $$\mathcal{D}_{\chi}(\Delta\boldsymbol{\rho},\Delta t) = \left\langle |\chi(\boldsymbol{\rho},t) - \chi(\boldsymbol{\rho}+\Delta\boldsymbol{\rho},t+\Delta t)|^2 \right\rangle, \tag{2.7}$$

  where $\boldsymbol{\rho}$ is a vector in three-dimensional space.

- **Covariance**

  Related to the correlation function is the covariance:

  $$C_{\chi}(\Delta\boldsymbol{\rho},\Delta t) = \langle [\chi(\boldsymbol{\rho},t) - \overline{\chi}]\,[\chi(\boldsymbol{\rho}+\Delta\boldsymbol{\rho},t+\Delta t) - \overline{\chi}]\rangle\,. \tag{2.8}$$

  For a stationary process (which in this case is where the second-order statistics are time invariant), the covariance is simply the correlation function of the fluctuations of the quantity $\chi$ about its mean:

$$C_{\chi}(\boldsymbol{\rho},\tau) = B_{\chi}(\Delta\boldsymbol{\rho},\Delta t) - \overline{\chi}^2. \tag{2.9}$$

The covariance is related to the structure function as follows:

$$\mathcal{D}_{\chi}(\boldsymbol{\rho},\tau) = 2\left[C_{\chi}(0,0) - C_{\chi}(\boldsymbol{\rho},\tau)\right]. \tag{2.10}$$

Therefore, these three basic statistically descriptive quantities are directly related.

- **Optical Coherence Function**

  A practical example of a correlation function is the optical coherence function, which is described specifically here for a wave with complex amplitude $\Psi$ propagating through space:

$$B_{\Psi}(h,\mathbf{r}) = \langle \Psi(h,\mathbf{r}')\,\Psi^*(h,\mathbf{r}'+\mathbf{r})\rangle, \tag{2.11}$$

  where $\mathbf{r}$ is a vector within the plane perpendicular to the propagation path at position $h$ along the path. This is of interest, due to the fact that the long exposure OTF ($\langle\mathcal{T}(f)\rangle$) is very much dependent on it (this is obvious from Eq. 2.1). A useful property of coherence functions is that they are invariant under a Fresnel propagation, as defined in Eq. 2.42 (one particular choice of model for optical fluctuations will involve the use of the Fresnel approximation to diffraction to determine the wave amplitude at ground level; see section 2.4).

- **Power Spectra**

  The second-order statistics of a quantity $\chi$ can equally well be represented in Fourier space by defining its power spectra in terms of the spatial frequency wave vector $\boldsymbol{\kappa}$ or temporal frequency $\nu$.

$$\Phi_{\chi}(\boldsymbol{\kappa}) = \int d\boldsymbol{\kappa}\, e^{-i\boldsymbol{\kappa}.\boldsymbol{\rho}} C_{\chi}(\boldsymbol{\rho}), \tag{2.12}$$

$$\tilde{\Phi}_{\chi}(\nu) = \int d\nu\, e^{-2\pi i\nu t}\tilde{C}_{\chi}(t), \tag{2.13}$$

  where the integral is over all space or all time accordingly and the quantities $C_{\chi}$ and $\tilde{C}_{\chi}$ represent the spatial and temporal covariance respectively. The spatial power spectrum may be a three-, two-, or one-dimensional quantity, depending on the dimensions of the space being integrated over. It is important at this stage to make a distinction between the spatial frequency ($f$) and the wave vector ($\boldsymbol{\kappa} = 2\pi\mathbf{f}$). One can be directly substituted for the other, as in the case of the power spectrum, for example:

$$\Phi_{\chi}(\mathbf{f}) = \int d\mathbf{f}\, \mathbf{e}^{-2\pi i\mathbf{f}.\boldsymbol{\rho}} \mathbf{C}_{\chi}(\boldsymbol{\rho}). \tag{2.14}$$

The choice between these two quantities depends on convention.

Making use of Eq. 2.10, and working in the spatial domain for the moment, it is possible to relate the structure function to the power spectrum:

$$\mathcal{D}_\chi(\boldsymbol{\rho}) = 2 \int\int\int_{-\infty}^{\infty} \Phi_\chi^{(3)}(\boldsymbol{\kappa}) \left[1 - \cos(\boldsymbol{\kappa}.\boldsymbol{\rho})\right] d\boldsymbol{\kappa}. \tag{2.15}$$

It will be seen later that, in the case of atmospheric turbulence, the statistical quantities of interest are mainly isotropic. In this case, it is possible to simplify the expression for the structure function so that it is in terms of either the three- and two-dimensional power spectra [10]:

$$\mathcal{D}_\chi(\rho) = 8\pi \int_0^{\infty} \Phi_\chi^{(3)}(\kappa) \left[1 - \frac{\sin(\kappa\rho)}{\kappa\rho}\right] \kappa^2 \, d\kappa, \tag{2.16}$$

$$\mathcal{D}_\chi(\rho) = 4\pi \int_0^{\infty} \Phi_\chi^{(2)}(\kappa) \left[1 - J_0(\kappa\rho)\right] \kappa \, d\kappa, \tag{2.17}$$

where the superscripts in brackets refer to the dimensions of the power spectra.

Any of the statistical quantities described may be applied in either the spatial or temporal domains, or a combination of the two.

The remainder of this chapter will be concerned with the derivation of the long-exposure optical transfer function of the atmosphere (and its relation to the quality of an image formed by a telescope) and also the statistics of the *phase* of a wave at the telescope aperture. The models which exist for optical propagation are derived from existing descriptions of the behaviour of turbulence in the atmosphere, particularly its refractive index variations. The first stage then is to give a brief resumé of the most widely accepted model of atmospheric turbulence.

## 2.1 The Statistics Of Refractive Index Fluctuations

The original atmospheric model that is used to predict all the phenomena associated with optical propagation through turbulence was first suggested by Kolmogorov. However, it is important to recognise that Kolmogorov's work describes only *velocity* fluctuations and makes no reference to refractive index. Tatarski [9] was the first to relate Kolmogorov's work to optical propagation. Nevertheless, Kolmogorov's work is still important in its own right because it describes the relative motions of the refractive index inhomogeneities and leads to predictions of the temporal behaviour of the optical quantities of interest.

Kolmogorov's model of velocity perturbations [9] is based on a simple dimensional argument whose major conclusions will be given here. Velocity perturbations in an otherwise stable atmosphere are said to occur over physical scales of up to a size $L_0$, otherwise known as the *outer scale.* Within the largest fluctuations, smaller perturbations arise due to energy being passed down from the larger fluctuations. In fact, the condition for the smaller perturbations to exist is that the energy converted to heat as a result of their formation does not exceed the energy which is contained within them. This is indeed the case for all sizes of perturbation from $L_0$ down to a size $l_0$ at which all the energy contained within the turbulence is converted directly to heat. At scales below this size $l_0$, the *inner scale*, fluctuations are severely damped. In the so-called *inertial range*, where the size of the perturbation $l$ is such that $l_0 \ll l \ll L_0$, the energy conversion rate can be considered to be negligible compared to the energy of formation of the preturbation, so that virtually all the energy is passed down from higher to lower sizes of turbulent eddies until sizes of the order of $l_0$ are reached. The first conclusion which may be reached is that, once the turbulent energy has been passed on from a larger flow to a smaller one several times, ie from a very low order perturbation to a much higher one, there is no longer any preferred direction to the flow (which is clearly not true for the largest laminar flow) and the fluctuations at this scale can be considered to be isotropic and homogeneous. The second conclusion involves the *velocity structure function*, $\mathcal{D}_v$,

$$\mathcal{D}_v(\boldsymbol{\rho}) = \left\langle |v(\boldsymbol{\rho}') - v(\boldsymbol{\rho}' + \boldsymbol{\rho})|^2 \right\rangle . \tag{2.18}$$

The rate at which energy is passed on from larger to smaller scales of turbulence is equal to the rate of dissipation into heat at the smallest scale sizes and is equal for all different sizes of flow. Kolmogorov used this information to derive the form of the structure function; this result will be repeated here, but not derived. By using a dimensional argument, it was demonstrated that the only way in which the structure function of velocity, $\mathcal{D}_v(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1)$ can be written as a function of the variable $|\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1|$ is as follows:

$$\mathcal{D}_v(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1) \propto |\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1|^{\frac{2}{3}} . \tag{2.19}$$

This equation is only valid within the inertial range; if it were valid for infinite separations $|\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1|$, the energy in the system would also be infinite. The inner scale, $l_0$, is thought to occur at sizes of around a few millimetres. The outer scale, $L_0$, has traditionally been thought to range from tens to hundreds of metres in the free atmosphere and near the ground to be of the order of the height above the ground. However, some authors [11] have recently suggested that it could on occasion be much smaller than this (perhaps smaller than the largest telescope apertures) and vigorous debate on this matter has ensued. If the Kolmogorov spectrum were to be extended outside the inertial range, various physical impossibilities (such as infinite energy) would occur. Various attempts have been made to predict the form of the structure function outside these limits but these are of little interest here.

Tatarski's major contribution was to propose that temperature variations within the atmosphere also follow a two-thirds power law [9]. He considered temperature to be a *passive additive*, *i.e.* its variations are caused solely by the mixing effects of the velocity fluctuations already described and that the effects of the temperature directly on the velocity distributions can be considered to be negligible. The temperature inhomogeneity can be considered to be the analogy of the velocity fluctuations described in the earlier discussion. Temperature inhomogeneities occur at the largest eddy size and can similarly sub-divide into smaller and smaller fluctuations until the size at which molecular diffusion into heat becomes important. Following similar dimensional considerations as before, the structure function of temperature, $\mathcal{D}_T(\boldsymbol{\rho}) = \langle |T(\boldsymbol{\rho}') - T(\boldsymbol{\rho}' + \boldsymbol{\rho})|^2 \rangle$, can also be considered to have a two-thirds power relationship with the separation of the reference positions,

$$\mathcal{D}_T(\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1) \propto |\boldsymbol{\rho}_2 - \boldsymbol{\rho}_1|^{\frac{2}{3}}. \tag{2.20}$$

The temperature distribution is important because it can be directly related to the refractive index distribution. Effects of water vapour in the air do not appear to have a significant effect on astronomical imaging through turbulence, so one can write [12]:

$$n - \langle n \rangle = 80.10^{-6} \frac{P}{T^2} \frac{\partial n}{\partial T} (T - \langle T \rangle) \tag{2.21}$$

where $n - \langle n \rangle$ is the refractive index perturbation, $T - \langle T \rangle$ is the temperature perturbation and the multiplicative constant is consistent with the pressure $P$ expressed in millibars and the air temperature $T$ in Kelvin degrees. Hence, the structure function for refractive index can be expressed as

$$\mathcal{D}_n = C_n^2 \rho^{\frac{2}{3}}. \tag{2.22}$$

This two-thirds power dependence on separation forms the basis of all the predictions for the statistics of the wavefront amplitude and phase of a wave incident on an aperture via the atmosphere. It is the starting point from which the optical theory will be derived in full. The constant $C_n^2$, the measure of the relative strength of turbulence, is a vital quantity. From it, many other statistical quantities of interest, such as $r_0$, can be calculated. The measurement of its dependence on height is an important goal [13], particularly in the prediction of the capabilities of laser beacons, which are more sensitive to turbulence in the lower layers of the atmosphere. Such studies [14] [15] have suggested that, in general, most of the atmosphere consists of turbulence which is optically extremely weak. Almost all of the optical turbulence appears to be concentrated into thin layers (with a thickness of the order of $L_0$ as one might expect) with a very high value of $C_n^2$. A number of example plots of the turbulence profile are given in Fig. 2.1.

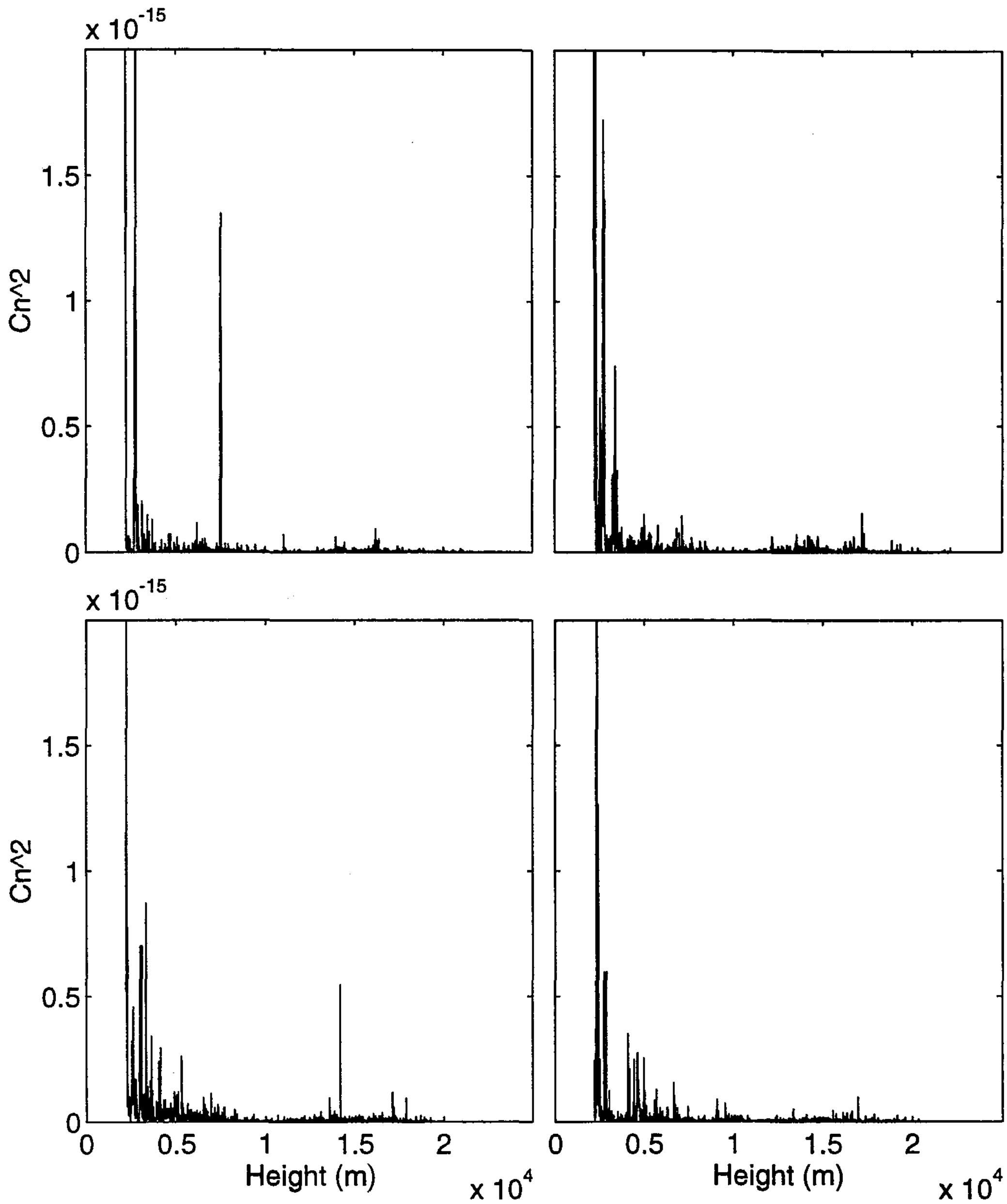



**Figure 2.1**: Graphs showing the turbulence profile $C_n^2$ verses height measured at four different times at the La Palma site. The data are taken from meteorological balloon measurements of temperature differences made at a vertical resolution of 6 metres by J.Vernin (University of Nice). The maximum values measured are $7 \times 10^{-15}$, but the scale has been reduced to enhance the display of lower values.

Since the refractive index perturbations are isotropic, their power spectrum is related to their structure function using equation 2.16,

$$\mathcal{D}_n(r) = 8\pi \int_0^{\infty} \Phi_n^{(3)}(\kappa) \left[1 - \frac{\sin(\kappa r)}{\kappa r}\right] \kappa^2 \, d\kappa. \tag{2.23}$$

The solution to this equation gives the three-dimensional power spectrum of refractive index,

$$\Phi_n^{(3)}(\kappa) = 0.0330 C_n^2 \kappa^{\frac{-11}{3}}. \tag{2.24}$$

## 2.2 Optical Effects: Formulation of Ideas

The obvious starting point for an analysis of the effects of refractive index inhomogeneities on optical propagation is with Maxwell's Equations.

It can be shown that, assuming constant magnetic permeability and ignoring depolarising effects, the scalar wave equation for one of the three components of the electromagnetic field $\Psi$ reduces to:

$$\nabla^2\Psi + k^2n^2\Psi = 0, \tag{2.25}$$

where n (the refractive index) is a function of the coordinate, $\boldsymbol{p}$, and $k = 2\pi/\lambda$, where $\lambda$ is the vacuum wavelength. For a plane wave propagating in the vertical direction z, with horizontal coordinate $\mathbf{r}$ and with n varying with $z$ and $\mathbf{r}$, the solution to this equation at height $z = h$ is given by

$$\Psi(h, \mathbf{r}) = A(h, \mathbf{r})\, e^{ikh} \tag{2.26}$$

(ignoring time dependence since it is irrelevant here; cf Eq. 2.4). $A$ is a complex quantity satisfying [16]

$$\nabla^2 A(h, \mathbf{r}) + 2ik\frac{\partial A(h, \mathbf{r})}{\partial z} + k^2 n(h, \mathbf{r})\, A(h, \mathbf{r}) = 0. \tag{2.27}$$

If the complete description of the refractive index distribution in the atmosphere at a particular instant in time were known, then a solution for $A$ would be calculable. Although only the statistics of the refractive index fluctuations are known, one is still interested in solutions of $A$ in order to describe the statistics of phase fluctuations at height $h$ (in particular the height corresponding to the telescope aperture). An exact solution for $A(h, \mathbf{r})$ is, in any case, not known to exist, therefore several attempts have been made to find approximate solutions which are applicable for the case of astronomical imaging. It is important to appreciate that one or more approximations has been made in order to find an analytic solution for $A(h, \mathbf{r})$ in each of the models about to be described. Clearly the most successful model should be the one in which the approximations are the most applicable to the real conditions encountered for which the solution to the wave equation is sought. Furthermore, it is the closeness of these predictions to available experimental results which has long since led to their widespread acceptance in this field. Tatarski [9] has developed two methods. The first is a very approximate one based on simple ray-tracing. This involves dropping the $\nabla^2 A$ term in Eq. 2.27 and ignoring diffraction effects. His second approach was based on the *Rytov Method* and he provides a complete analysis of the propagation together with the limits of the approximations involved. A summary of this approach will be given in section 2.9. F. Roddier's treatment [12] is a far simpler one which makes use of both the geometric and physical approaches in combination. Although the approximations involved are much more restrictive, Roddier's method does produce the same results as Tatarski's more formal treatment. For this reason, the next six sections will follow closely the reasoning of Roddier. Reference will also be made to Tatarski and, in particular, Goodman's summary [17] of Tatarski's analysis where convenient.

For convenience, the atmosphere will be divided into thin horizontal sections, each of which has a particular strength of turbulence associated with it. The thickness of an individual layer will be assumed to be much larger than the correlation length of the inhomogeneities, which means that the resulting phase shift will be a sum over a large number of statistical realisations but small enough so that diffraction effects can be ignored. In reality, such a layer may not exist, but this will be considered later in section 2.9. Within each layer, a geometrical model will be used to calculate the phase shift induced on an incident wave. Using a small perturbation approximation, the complex amplitude at ground level will be seen to be calculable as a linear function of the phase shifts associated with all the layers in the atmosphere. For this (much longer) propagation path, a physical optics approach is necessary.

## 2.3 Propagation Through a Thin Layer

The first step is to calculate how the refractive index variations within a turbulent layer alter the statistics of the complex amplitude of an incident plane wave. A plane electromagnetic wave propagating in the (vertical) z-direction (although the theory can easily be generalised for a source not at the zenith, or for a spherical wave [18]) with wavelength $\lambda$ and corresponding wave-vector $k = 2\pi/\lambda$ will be incident normally on a single layer, and any perturbations in phase that result must be considered to be small (see Eq. 2.79 for what this means), otherwise a significant spreading of the beam would occur and diffraction effects cannot be ignored [19]. This will allow a ray optics formulation to be used at least within the layers of interest (although between layers and down to ground level of course diffraction effects will have to be accounted for). Hence the atmosphere can be regarded as a series of *phase screens* separated by volumes of free space. The phase shift introduced into a ray with horizontal coordinate $\mathbf{r}$ propagating vertically within a subsection of the layer of thickness $\delta z$ with refractive index n is equal to $n\delta z$, so across an entire layer of thickness $\delta h$, the phase shift introduced is:

$$\varphi(h, \mathbf{r}) = k \int_h^{h+\delta h} dz \cdot n(\mathbf{r}, z) \tag{2.28}$$

Hence the complex amplitude immediately after the layer could be written as:

$$\Psi(h, \mathbf{r}) = e^{i\varphi(h,\mathbf{r})}. \tag{2.29}$$

Since diffraction effects are to be ignored, the amplitude can be taken to be constant across the entire wave immediately after passing through the layer. The coherence function of complex amplitude (see Eq. 2.11 for a definition of this function) can thus be rewritten as:

$$B(h,\mathbf{r}) = \left\langle e^{i[\varphi(h,\mathbf{r}')-\varphi(h,\mathbf{r}'+\mathbf{r})]}\right\rangle. \tag{2.30}$$

Equation 2.28 shows that $\varphi(h,\mathbf{r})$ is actually a sum of a large number of statistically independent variables (under the assumption that the phase screen is much thicker than the refractive index correlation length). The Central Limit Theorem [20] thus allows Gaussian statistics to be invoked for $\varphi(h,\mathbf{r})$ and therefore also for $\varphi(h,\mathbf{r}') - \varphi(h,\mathbf{r}'+\mathbf{r})$:

$$B(h,\mathbf{r}) = e^{-\frac{1}{2}\left\langle|\varphi(h,\mathbf{r}')-\varphi(h,\mathbf{r}'+\mathbf{r})|^2\right\rangle}. \tag{2.31}$$

or in terms of the phase structure function,

$$B(h,\mathbf{r}) = e^{-\frac{1}{2}\mathcal{D}_\varphi(h,\mathbf{r})}. \tag{2.32}$$

Now a relation is sought between the three-dimensional structure function of refractive index and the two-dimensional structure function of the phase shifts introduced by a horizontal turbulent layer. Defining the covariance of phase (shift) as

$$C_\varphi(h,\mathbf{r}) = \langle\varphi(h,\mathbf{r}')\,\varphi(h,\mathbf{r}'+\mathbf{r})\rangle, \tag{2.33}$$

it follows that

$$C_\varphi(h,\mathbf{r}) = k^2\int_h^{h+\delta h} dz\int_h^{h+\delta h} dz'\,\langle n(\mathbf{r}',z)\,n(\mathbf{r}'+\mathbf{r},z')\rangle. \tag{2.34}$$

Substituting $\zeta$ for $z - z'$, it is immediately obvious that $C_\varphi$ is related to the (3-dimensional) covariance of refractive index fluctuations, $B_n(\mathbf{r},\zeta)$.

$$C_\varphi(h,\mathbf{r}) = k^2\int_h^{h+\delta h} dz\int_h^{h+\delta h-z} d\zeta\cdot B_n(\mathbf{r},\zeta). \tag{2.35}$$

Since $\delta h$ is larger than the refractive index correlation size, the covariance of the index fluctuations falls to zero at the limits of the $\zeta$ integration shown in Eq. 2.35, and the limits can be legitimately extended from $-\infty$ to $+\infty$. The z-dependence is removed, hence

$$C_\varphi(h,\mathbf{r}) = k^2\delta h\int_{-\infty}^{\infty} d\zeta\cdot B_n(\mathbf{r},\zeta). \tag{2.36}$$

The stationarity and isotropy of turbulence allows the structure function of phase or refractive index to be written in terms of its covariance:

$$\mathcal{D}_\varphi(h,\mathbf{r}) = 2\left[C_\varphi(h,\mathbf{0}) - C_\varphi(h,\mathbf{r}),\right] \tag{2.37}$$

$$\mathcal{D}_n(z, \mathbf{r}) = 2\left[C_n(0, \mathbf{0}) - C_n(z, \mathbf{r}),\right] \tag{2.38}$$

noting that $\mathcal{D}_\varphi(h, \mathbf{r})$ is a two-dimensional function valid at height h only and $\mathcal{D}_n(z, \mathbf{r})$ is actually three-dimensional (valid in any direction within a layer), but is only relevant in the horizontal direction. Therefore,

$$\mathcal{D}_\varphi(h, \mathbf{r}) = k^2 \delta h \int_{-\infty}^{\infty} d\zeta \left[\mathcal{D}_n(\zeta, \mathbf{r}) - \mathcal{D}_n(\zeta, \mathbf{0})\right]. \tag{2.39}$$

By inserting the Kolmogorov law for refractive index structure function,

$$\mathcal{D}_\varphi(h, \mathbf{r}) = k^2 \delta h C_n^2(h) \int_{-\infty}^{\infty} d\zeta \left[\left(r^2 + \zeta^2\right)^{\frac{1}{3}} - \zeta^{\frac{2}{3}}\right] \tag{2.40}$$

$$= 2.91 k^2 \delta h C_n^2(h)\, r^{\frac{5}{3}}, \tag{2.41}$$

an expression for the structure function of phase perturbations introduced by a single thin layer in terms of the refractive index constant is obtained. It must be recognised that this expression is only valid within the inertial range; if it were valid for $r \to \infty$, perturbations with infinite energy would have to exist.

It is the fact that Gaussian statistics can be applied to the phase that is vital in determining the various power laws involved with the spatio-temporal statistics of the wave at ground level. The model as a whole is known as the *log-normal* model since the quantities of interest, the phase and log amplitude, are related exponentially to the complex amplitude. The weight of scientific evidence [21] favours log-normal statistics, at least for weak fluctuations.

## 2.4 Effect Of A Thin Layer On The Statistics At Ground Level

The easiest calculable statistical property of the complex amplitude of a wave at ground level is the coherence function. Considering a single slice of the atmosphere at height $h$ to be turbulent, with the remainder being completely homogeneous, it will now be shown [12] that the coherence function of a vertically propagating plane electromagnetic wave, which is altered as it passes through the layer, remains invariant in its propagation through the rest of the atmosphere. The Fresnel approximation to diffraction can be used, since the wavelengths involved are much smaller than the smallest scales of refractive index inhomogeneities. Therefore [17],

$$\Psi(0, \mathbf{r}) = \Psi(h, \mathbf{r}) \otimes \frac{1}{i\lambda h} exp\left(i\pi \frac{r^2}{\lambda h}\right) \tag{2.42}$$

($\otimes$ represents a convolution operation). The coherence function was defined in Eq. 2.11; an alternative way of writing this is in terms of a convolution:

$$B(0, \mathbf{r}) = \Psi(0, -\mathbf{r}) \otimes \Psi^*(0, \mathbf{r}). \tag{2.43}$$

Applying Eq. 2.42 to this formulation, and taking advantage of the associative and commutative properties of the convolution operator, the coherence function at ground level can be expressed in terms of the complex amplitude at height $h$:

$$B(0,\mathbf{r}) = \langle \Psi(h,\mathbf{r}')\, \Psi^*(h,\mathbf{r}'+\mathbf{r}) \rangle \otimes \frac{1}{i\lambda h} exp\left(i\pi\frac{r^2}{\lambda h}\right) \otimes -\frac{1}{i\lambda h} exp\left(-i\pi\frac{r^2}{\lambda h}\right). \quad (2.44)$$

Furthermore,

$$\frac{1}{i\lambda h} exp\left(i\pi\frac{r^2}{\lambda h}\right) \otimes -\frac{1}{i\lambda h} exp\left(-i\pi\frac{r^2}{\lambda h}\right) = \delta(\mathbf{r}), \quad (2.45)$$

*i.e.* the coherence function is invariant under a Fresnel propagation:

$$B(0,\mathbf{r}) = B(h,\mathbf{r}) \quad (2.46)$$

and

$$B(0,\mathbf{r}) = e^{-\frac{1}{2}\mathcal{D}_\varphi(h,\mathbf{r})}. \quad (2.47)$$

for a single atmospheric layer at height $h$. So the coherence function at ground level for a wave which has its phase altered by a single layer at height h can be directly related to the phase structure function for that layer. However, a phase perturbation at height h will have an effect on both the phase *and* the log amplitude at ground level due to diffraction effects. The statistics of the phase at ground level are not so easily calculable, and a physical optics approach is necessary.

## 2.5 Coherence Function Through Many Layers

As already mentioned, the coherence function for complex amplitude is a useful function because it can be used to calculate the effects of the turbulence on the resolution of a telescope. Therefore, it will be helpful to calculate the effect of the entire atmosphere on the optical coherence function (see Eq. 2.11 for a definition) at ground level. For this calculation, the atmospheric continuum will be considered to consist of a number of thin layers, each with a different strength of turbulence ($C_n^2$). It is easy to show that, since the coherence function is invariant under a Fresnel diffraction, and that the conditions that have been assumed so far are such that the Fresnel approximation is valid, the coherence function at ground level is merely the product of the coherence functions produced by each layer [12].

Supposing layer $j$ at height $h_j$ with thickness $\delta h$ produces a phase shift $\varphi_j(\mathbf{r})$, then

$$B(h_j,\mathbf{r}) = B(h_j+\delta h,\mathbf{r}) \cdot \left\langle e^{i[\varphi_j(\mathbf{r}')-\varphi_j(\mathbf{r}'+\mathbf{r})]} \right\rangle. \quad (2.48)$$

As each layer is traversed, the coherence function for the propagating wave is simply multiplied by the expression in Eq. 2.30. The coherence function at ground level is simply a linear product of coherence functions of each layer treated individually. Making use of Eqs. 2.32 and 2.41,

$$B(0,\mathbf{r}) = \prod_j e^{-\frac{1}{2}\left(2.91k^2\delta h_j C_n^2(h_j) r^{\frac{5}{3}}\right)}. \tag{2.49}$$

Expressing this product as an integral for a continuous distribution of turbulence,

$$B(0,\mathbf{r}) = e^{\left[-1.46k^2 r^{\frac{5}{3}} \int_0^\infty dh \cdot C_n^2(h)\right]}. \tag{2.50}$$

The statistics of the complex amplitude at ground level have now been related to the distribution of refractive index inhomogeneities in the entire atmosphere.

## 2.6 The Limit Of Atmospheric Resolution

The quality of the image produced by an optical system can be quantified in terms of its Optical Transfer Function (OTF). An expression for the optical coherence function has already been obtained, and its close relation to the OTF allows the resolution of a telescope to be related to the refractive index distribution in the atmosphere. In fact, the OTF of the whole system can be expressed as a product of the individual components of that system [17], so that in the case of a telescope of diameter D combined with a turbulent atmosphere,

$$\mathcal{T}(\mathcal{F}) = \mathcal{T}_D(\mathcal{F}) \cdot \mathcal{T}_{atm}(\mathcal{F}) \tag{2.51}$$

(equivalent to Eq. 2.3, but now written in terms of an *angular* frequency vector $\mathcal{F}$, with dimensions of radians$^{-1}$), since the resolution of an aperture is most easily quantified in terms of angles). Here $\mathcal{T}_D$ is the OTF in the absence of atmosphere and $\mathcal{T}_{atm}$ is the OTF of the atmosphere. It should be made clear that, since the majority of astronomical applications involve integrating over long time periods, the following refers to the long-exposure OTF of the system (although a so-called 'short-exposure' OTF can also be calculated). Fried [22] originally defined the parameter $r_0$ as the maximum size of telescope which is capable of diffraction limited imaging under the prevailing turbulent conditions. Clearly, when the telescope diameter D is smaller than $r_0$, the image is diffraction-limited and when larger it is not. One way of defining the resolution of a system is by equating it to the volume under its OTF. So the parameter $r_0$ can be more precisely defined as the diameter of a telescope for which the resolving power of the atmosphere is equal to the resolving power of the telescope. In other words the reduction in image quality due to the atmosphere is equal to the reduction in quality caused by the telescope having a finite cut-off frequency.

The long exposure OTF of a perfect telescope/atmosphere system can be defined in terms of the coherence function of the complex phase within the aperture by combining Eqs. 2.1 and 2.11 and setting $\Psi$ to zero outside the pupil. If this is normalised so that $\langle \mathcal{T}_0(0) \rangle = 1$ then in terms of the optical coherence function,

$$\langle \mathcal{T}(\boldsymbol{\mathcal{F}}) \rangle = B(0, \boldsymbol{\mathcal{F}}) \cdot \mathcal{T}_D(\boldsymbol{\mathcal{F}}). \tag{2.52}$$

The resolving power of the combined system is

$$\mathcal{R} = \int d\boldsymbol{\mathcal{F}} \cdot B(0, \boldsymbol{\mathcal{F}}) \cdot \mathcal{T}_D(\boldsymbol{\mathcal{F}}). \tag{2.53}$$

The OTF of an aberration-free circular aperture is given by Goodman [17] as

$$\mathcal{T}_D(\boldsymbol{\mathcal{F}}) = \frac{2}{\pi}\left[\cos^{-1}\left(\frac{\mathcal{F}}{\mathcal{F}_0}\right) - \frac{\mathcal{F}}{\mathcal{F}_0}\sqrt{1 - \frac{\mathcal{F}}{\mathcal{F}_0}^2}\right] \tag{2.54}$$

when $\mathcal{F} < \mathcal{F}_0$, with $\mathcal{F}_0 = \frac{D}{\lambda}$, the cutoff frequency of the aperture. For higher frequencies than $\mathcal{F}_0$, the OTF is zero. In the case of a circular telescope with no atmosphere present,

$$\mathcal{R}_D = \int \mathcal{T}_D \cdot d\boldsymbol{\mathcal{F}} = \frac{1}{4}\pi\left(\frac{D}{\lambda}\right)^2. \tag{2.55}$$

A telescope of size $r_0$ should have its resolution limited equally by atmosphere and pupil geometry:

$$\mathcal{R}_{r_0} = \mathcal{R}_{atm}. \tag{2.56}$$

Therefore,

$$\int d\boldsymbol{\mathcal{F}} \cdot B(0, \boldsymbol{\mathcal{F}}) = \int d\boldsymbol{\mathcal{F}} \cdot \mathcal{T}_{r_0}(\boldsymbol{\mathcal{F}}). \tag{2.57}$$

Fried [23] defined the parameter $r_0$, which is the solution to Eq. 2.57, in the following way,

$$r_0 = \left[0.423k^2 \int_0^{+\infty} dh \cdot C_n^2(h)\right]^{-\frac{3}{5}}. \tag{2.58}$$

The dependence of $r_0$ on wavelength should be noted,

$$r_0 \propto \lambda^{\frac{6}{5}}. \tag{2.59}$$

If the line of sight through the atmosphere is at an angle $\vartheta$ to the zenith, then the propagation path through each of the layers is multiplied by a factor $(\cos\vartheta)^{-1}$ and the resolution is correspondingly worse. In this more general case:

$$r_0 = (\cos\vartheta)^{-\frac{3}{5}}\left[0.423k^2 \int_0^{+\infty} dh \cdot C_n^2(h)\right]^{-\frac{3}{5}}. \tag{2.60}$$

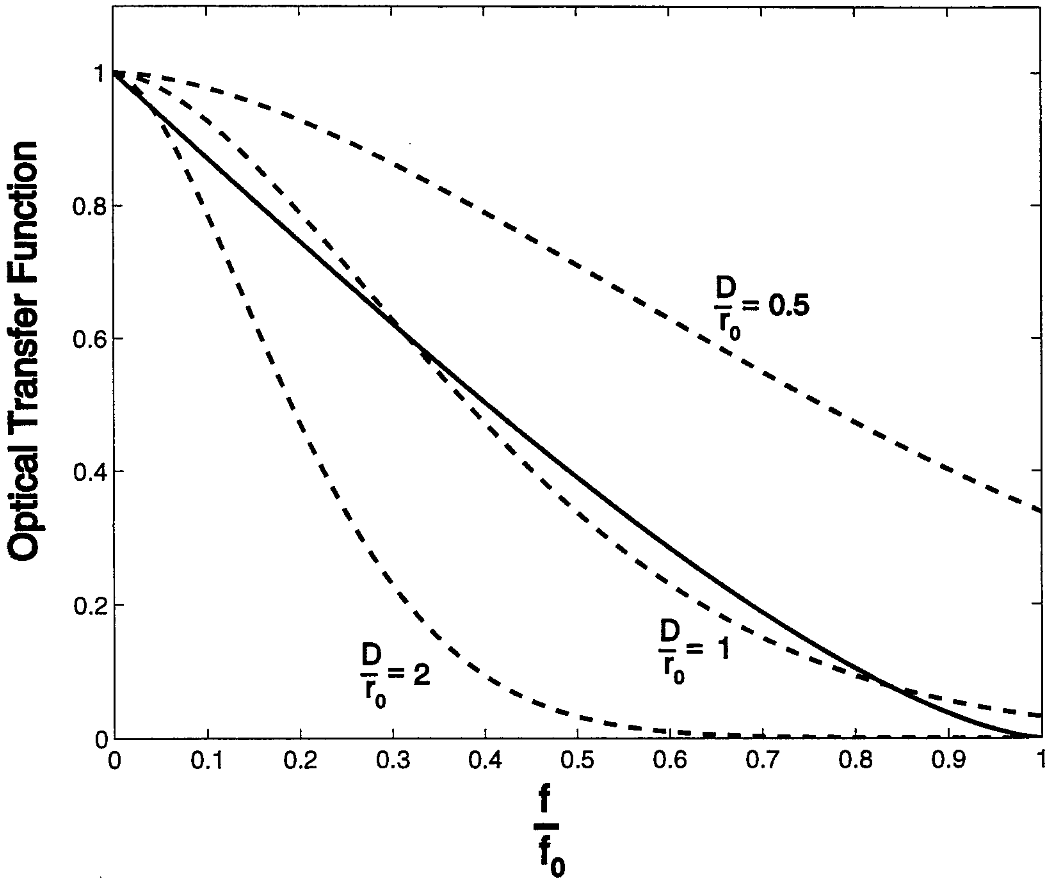



**Figure 2.2**: OTF for a circular aperture of size $D$ (solid line) compared with the atmosphere in various states characterised by different coherence diameters, $r_0 = D/2$, $D$ and $2D$ (dashed lines).

This allows all the important statistical properties discussed so far to be redefined in terms of $r_0$.

$$B(0, \mathbf{r}) = e^{-3.44\left(\frac{r}{r_0}\right)^{\frac{3}{5}}}, \tag{2.61}$$

$$\mathcal{D}_{\varphi}(h, \mathbf{r}) = 6.88\left(\frac{r}{r_0}\right)^{\frac{3}{5}}. \tag{2.62}$$

Perhaps the most important quantity is the long-exposure OTF of the atmosphere, which leads directly from Eq. 2.61.

$$\langle \mathcal{T}_{atm}(\mathcal{F}) \rangle = e^{-3.44\left(\frac{\lambda\mathcal{F}}{r_0}\right)^{\frac{5}{3}}}, \tag{2.63}$$

A cross-section of the OTF of the atmosphere is plotted in Fig. 2.2 together with corresponding OTFs for circular apertures of diameters 0.5, 1 and 2 times $r_0$. The volume under the curve for $D = r_0$ is the same as that under the atmospheric OTF by definition.

Noll [24] has calculated that $r_0$ is the spatial scale across which the mean-squared phase fluctuation (or phase variance) is 1.03 square radians. This was done by calculating the residual wavefront error once the first Zernike mode, *i.e.* the piston, has been corrected for (see the section which defines Zernike polynomials - section 3.2). This is equivalent to finding the fluctuation relative to the mean phase in the aperture. An important point is that $r_0$ is actually calculated as the *long-exposure* resolution of an atmosphere with a particular $C_n^2$ profile. It is still possible to improve on this resolution with a 'lucky' short-exposure image [25].

## 2.7 The Small Perturbation Approximation

Although the form of the coherence function for complex amplitude has been calculated at ground level, it has not yet been possible to predict the separate statistics for phase ($\psi$) and log amplitude ($\gamma$). Since this account is primarily concerned with wavefront sensing, it is the properties of the wavefront phase which are of most significance here. Only phase perturbations were considered at the boundaries of the turbulent layers themselves; however, a phase modulation at a particular height will translate into *both* phase and amplitude fluctuations at ground level when diffraction effects are taken into consideration.

Returning for the moment to the case of a single layer at height h with an induced complex amplitude fluctuation of $\Psi(h,\mathbf{r})$, the complex amplitude at ground level can simply be expressed as a Fresnel convolution of the amplitude at the layer height [12]:

$$\Psi(0,\mathbf{r}) = \Psi(h,\mathbf{r}) \otimes \frac{1}{i\lambda h} e^{i\pi\frac{r^2}{\lambda h}}. \tag{2.64}$$

In order to simplify this expression further, the small phase perturbation assumption that was introduced in section 2.3 will again be used. This restriction allows the (normalised) complex field just after propagation through a single layer to be written as

$$\Psi(h,\mathbf{r}) \approx 1 + i\varphi(h,\mathbf{r}). \tag{2.65}$$

The spatial power spectrum for these phase perturbations $\varphi(h,\mathbf{x})$ within a single layer can be immediately calculated. The covariance of the phase shift for a single layer is related to the three-dimensional covariance of refractive index fluctuations via Eq. 2.36. Therefore the two-dimensional power spectrum for the phase shift can be found in terms of the three-dimensional power spectrum of refractive index fluctuations by using their relation in Fourier space (Eq. 2.12). Making use of the expression for the three-dimensional power spectrum of refractive index fluctuations (Eq. 2.24), and converting from wavenumber to spatial frequency, the two-dimensional power spectrum of the phase perturbations within a single layer of thickness $\delta h$ is given by

$$\begin{aligned}\Phi_n^{(2)}(\mathbf{f}) &= (2\pi)^3\, 0.0330 C_n^2(h)\,(2\pi f)^{-\frac{11}{3}} \\ &= 0.00969 k^2 C_n^2(h)\,\delta h f^{\frac{-11}{3}}.\end{aligned} \tag{2.66}$$

Since the convolution of the Fresnel factor and unity is equal to unity,

$$1 \otimes \frac{1}{i\lambda h} e^{i\pi\frac{r^2}{\lambda h}} = 1 \tag{2.67}$$

(this can be concluded by considering the equivalent relation in Fourier space), the complex amplitude at ground level can also be expressed as a small perturbation on the mean value:

$$\Psi(0,\mathbf{r}) = 1 + \epsilon(0,\mathbf{r}), \tag{2.68}$$

where

$$\epsilon(0,\mathbf{r}) = \varphi(h,\mathbf{r}) \otimes \frac{1}{i\lambda h} e^{i\pi\frac{r^2}{\lambda h}}. \tag{2.69}$$

$\epsilon(0,\mathbf{r})$ is a complex quantity whose real part,

$$\gamma(0,\mathbf{r}) = \varphi(h,\mathbf{r}) \otimes \frac{1}{\lambda h} \cos\left(i\pi\frac{r^2}{\lambda h}\right), \tag{2.70}$$

describes the fluctuations of log-amplitude and whose imaginary part,

$$\varphi(0,\mathbf{r}) = \varphi(h,\mathbf{r}) \otimes \frac{1}{\lambda h} \sin\left(i\pi\frac{r^2}{\lambda h}\right), \tag{2.71}$$

describes the fluctuations of phase at ground level.

Now the significance of the small perturbation approximation will become clear. The complex amplitude perturbation at ground level can be shown to be a linear combination of the effects in each of the layers. Taking two layers at heights $h_1$ and $h_2$ which introduce phase perturbations $\varphi_1(\mathbf{r})$ and $\varphi_2(\mathbf{x})$ respectively, the complex amplitude fluctuation at ground level can be written as:

$$\epsilon(0,\mathbf{r}) = \left[i\varphi_1(\mathbf{r}) \otimes \frac{1}{i\lambda(h_1-h_2)} e^{i\pi\frac{r^2}{\lambda(h_1-h_2)}} + i\varphi_2(\mathbf{r})\right] \otimes \frac{1}{i\lambda h_2} e^{i\pi\frac{r^2}{\lambda h_2}}. \tag{2.72}$$

This reduces to:

$$\epsilon(0,\mathbf{r}) = i\varphi_1(\mathbf{r}) \otimes \frac{1}{i\lambda h_1} e^{i\pi\frac{r^2}{\lambda h_1}} + i\varphi_2(\mathbf{r}) \otimes \frac{1}{i\lambda h_2} e^{i\pi\frac{r^2}{\lambda h_2}}. \tag{2.73}$$

This equation allows the statistics of the complex amplitude at ground level to be quoted as a sum of the effects produced by many layers or, in the case of a continuous distribution of turbulence over height, in terms of an integral. Equations 2.71 and 2.70 yield the statistics of the phase and log-amplitude at the small perturbation approximation. Taking their Fourier Transforms according to Eq. 2.12 gives the following results for the power spectra of fluctuations of complex amplitude, log amplitude and phase respectively [12]:

$$\Phi_{\epsilon}^{(2)}(\mathbf{f}) = \frac{0.383}{\lambda^2} f^{\frac{-11}{3}} \int_h^{\infty} C_n^2(h)\, dh. \tag{2.74}$$

$$\Phi_{\gamma}^{(2)}(\mathbf{f}) = \frac{0.383}{\lambda^2} f^{\frac{-11}{3}} \int_h^{\infty} C_n^2(h) \sin^2\left(\pi\lambda h f^2\right)\, dh. \tag{2.75}$$

$$\Phi_{\varphi}^{(2)}(\mathbf{f}) = \frac{0.383}{\lambda^2} f^{\frac{-11}{3}} \int_h^{\infty} C_n^2(h) \cos^2\left(\pi\lambda h f^2\right)\, dh. \tag{2.76}$$

Thus the turbulent energy contained in the phase fluctuations which occur just after propagation through a thin layer is conserved during the passage to ground level; however, some of this energy is converted from phase fluctuations to log-amplitude fluctuations (scintillation).

## 2.8 The Near-Field Approximation

One further approximation can be made under certain, very limiting, atmospheric conditions. Goodman [17] shows that the structure function of complex amplitude is equal to the sum of the individual structure functions of phase and log-amplitude:

$$\mathcal{D}_{\Psi}(h,\mathbf{r}) = \mathcal{D}_{\gamma}(h,\mathbf{r}) + \mathcal{D}_{\varphi}(h,\mathbf{r}). \tag{2.77}$$

Equations 2.17, 2.75 and 2.76 could be used to find the structure function of either the log amplitude or the phase fluctuations from their respective power spectrum.

It has already been shown that, over small propagation distances within a turbulent layer, amplitude fluctuations can be ignored and all the fluctuations are in the phase:

$$\mathcal{D}_{\Psi}(h,\mathbf{r}) = \mathcal{D}_{\varphi}(h,\mathbf{r}) \quad [\textit{thin layer}]. \tag{2.78}$$

Correspondingly, for spatial phase variations of scale $l$ , the log amplitude variations which evolve after propagation through a distance $z$ are much smaller than the phase variations when $l \ll \sqrt{\pi\lambda z}$. For modulations of any significance, $l$ is always going to be much greater than $r_0$, hence the height $h$ at which the turbulence must occur is given by

$$h \ll r_0^2/\pi\lambda \tag{2.79}$$

in order for log amplitude fluctuations to be neglected. This restriction forms the basis of what is known as the *near-field approximation*, so named because it is equivalent to the condition for a Fresnel approximation to be valid. Under this assumption, the structure function of log amplitude fluctuations is negligible; hence (within the inertial range only)

$$\mathcal{D}_{\varphi}(h,\mathbf{r}) = \mathcal{D}_{\Psi}(h,\mathbf{r}). \tag{2.80}$$

In fact in the limits of the near and far field,

$$\mathcal{D}_{\Psi}(h,\mathbf{r}) = \omega(h,\mathbf{r})\,\mathcal{D}_{\varphi}(h,\mathbf{r}). \tag{2.81}$$

$$\omega(h,\mathbf{r}) = \begin{pmatrix} 0.5 & r^2 \ll \lambda h & far\ field \\ 1 & r^2 \gg \lambda h & near\ field \end{pmatrix} \tag{2.82}$$

The near-field approximation can equally well be applied to the power spectrum of the phase fluctuations given in Eq. 2.76 (thus again neglecting power in the log amplitude perturbations). It is quoted here in terms of the two-dimensional spatial frequency, $f$, rather than the wavenumber, since it is most often used in this formulation:

$$\Phi_{\varphi}^{(2)}(f) = 0.00969k^2 C_n^2(h)\,\delta h f^{\frac{-11}{3}}. \tag{2.83}$$

In fact, for an atmosphere with many layers of turbulence, the power spectrum can be written as an integral over all the turbulence,

$$\Phi_{\varphi}^{(2)}(f) = 0.00969k^2 \int_0^{\infty} C_n^2(h)\,dh f^{\frac{-11}{3}}, \tag{2.84}$$

which can be simplified by using the definition of $r_0$ given in equation 2.58:

$$\Phi_{\varphi}^{(2)}(f) = 0.0229 r_0^{-\frac{5}{3}} f^{\frac{-11}{3}}. \tag{2.85}$$

Young [26] argues that, in general, the near-field approximation appears to be good for the vast majority of the turbulence according to data available to him at the time.

The statistics of the phase at the telescope aperture have now been completely described for Kolmogorov turbulence, albeit under certain very restrictive limiting approximations. Nevertheless, a more general approach leads to the same results; this will be outlined in the next section.

## 2.9 The Restrictions of the Geometric Approach

The analysis of sections 2.3 to 2.8 relies on a number of assumptions, two of which will be discussed here. An assumption in section 2.3 was that a turbulent layer could be defined which was thick enough that a perpendicular path through it would pass through a large number of statistical realisations (*ie* $\delta h \gg L_0$), but also that the path distance was such that diffraction effects could be ignored (*ie* $\delta h \ll l_0^2/\lambda$). Clearly, for outer scales of the order of hundreds of metres and for inner scales of the order of millimetres these two assumptions are not consistent with each other at optical wavelengths. However, it is expected that this derivation may still be valid for two reasons. Firstly, the energy in the fluctuations at low scale sizes is very small and may often be neglected (see Eq. 2.22). Secondly, the structure function of complex amplitude, $\mathcal{D}_\Psi(r)$, is mainly affected by eddies of size of order $r$. For sizes larger than this, the two paths (separated by $r$) generally are affected in the same way and their difference is not affected. For Gaussian statistics to be valid, the path must extend through several independent statistical realisations and this can certainly be true for path lengths much smaller than $L_0$.

The second, more serious restriction, is that of the small perturbation approximation. This was introduced in section 2.7 so that the Fresnel convolutions in Eq. 2.64 could be simplified. The approximation itself may not be valid. Tatarski's more extensive approach lifts these restrictions and, interestingly, the results derived thus far are found to be applicable under a much wider range of conditions. A full description of his approach will not be repeated here; however an outline will be given which includes a description of the approximations made, so at least the significance of the validity of the approach can be understood.

For weak turbulence (the definition of which will be discussed later), it is convenient to treat the complex amplitude given in the wave equation (2.25), repeated here:

$$\nabla^2\Psi + k^2 n^2 \Psi = 0, \tag{2.86}$$

as a perturbation on the unscattered wave:

$$\Psi = \Psi_0 + \Psi_1 + \Psi_2 + \cdots. \tag{2.87}$$

The refractive index distribution can also be considered to be composed of a fluctuation on a mean value, *i.e.* $n = \langle n \rangle + n_1$, and to a good approximation $\langle n \rangle = 1$. Neglecting higher than first order terms (which represent multiple scattering effects), the wave equation transforms to:

$$\nabla^2\Psi_0 + k^2\Psi_0 = 0, \tag{2.88}$$

$$\nabla^2\Psi_1 + k^2\Psi_1 + 2k^2 n_1 \Psi_0 = 0. \tag{2.89}$$

A solution for this equation could be found directly. However, Tatarski introduced the so-called *Rytov Approximation* which involves substituting $\Lambda = \ln(\Psi)$ into the wave equation. The advantage of this will become clear when the validity of further approximations is discussed. Therefore, $\nabla\Psi = \Psi\nabla\Lambda$, and $\nabla^2\Psi = \Psi\nabla^2\Lambda + \Psi\nabla\Lambda.\nabla\Lambda$, and the wave equation transforms to:

$$\nabla^2\Lambda + \nabla\Lambda.\nabla\Lambda + k^2 n^2 = 0. \tag{2.90}$$

Once again, the variable $\Lambda$ can be treated as a perturbation expansion,

$$\Lambda = \Lambda_0 + \Lambda_1 + \Lambda_2 + \cdots. \tag{2.91}$$

As will be discussed later, the solution for $\Lambda_1$ is found to be much more general than the solution for $\Psi_1$ from Eq. 2.89. The physical significance of $\Lambda$ is clear from the fact that

$$\begin{aligned} \Lambda_1 &= \ln\Psi - \ln\Psi_0 \\ &= \ln\frac{\Psi}{\Psi_0} \\ &= \gamma + i\varphi, \end{aligned} \tag{2.92}$$

where $\gamma$ and $\varphi$ are the log-amplitude and phase variation about their respective mean values. Considering first order effects only, where $n_1^2$ is small enough to be neglected and making a first assumption:

$$|\nabla\Lambda_1| \ll |\nabla\Lambda_0|, \tag{2.93}$$

the equation for which a solution must be found is:

$$\nabla^2\Lambda_1 + 2\nabla\Lambda_0.\nabla\Lambda_1 + k^2 n_1 = 0. \tag{2.94}$$

A standard method of solving diffraction problems is to use the *Green's function solution*, which treats every scattering point in the atmosphere (with coordinate $\boldsymbol{\rho}'$) as an individual point source, and calculates the resulting total intensity at the point of interest ($\boldsymbol{\rho}$) from all possible scatterers. This is done, in principle, by integrating the complex amplitude contributions over all vectors $\boldsymbol{\rho} - \boldsymbol{\rho}'$. However, an analytic solution for such problems is normally currently unavailable, so some sort of approximation must be made which is valid under the real conditions which exist in the problem of interest. In the present case, $\boldsymbol{\rho}'$ is a coordinate within a (scattering) turbulent layer at height $h$, so it is reasonable to attempt to expand $\boldsymbol{\rho} - \boldsymbol{\rho}'$ into a component in the direction of propagation, $z - z'$, and a component in a perpendicular plane, $\mathbf{r} - \mathbf{r}'$. A binomial expansion gives:

$$|\boldsymbol{\rho} - \boldsymbol{\rho}'| = (z - z')\left[1 + \frac{(\mathbf{r} - \mathbf{r}')^2}{2(z - z')^2} - \frac{(\mathbf{r} - \mathbf{r}')^4}{8(z - z')^4} + \cdots\right]. \tag{2.95}$$

One must decide how many of the terms in Eq. 2.95 must be retained in order to remain applicable to the problem of interest. The complex amplitude at coordinate $\boldsymbol{p}$ will not be affected by higher order terms for which $\boldsymbol{\rho}-\boldsymbol{\rho}' \ll \lambda$, where $\lambda$ is the wavelength of the light being scattered. The maximum significant scattering angle from the turbulence can be considered to be $\lambda/l_0$ (since $l_0$, the inner scale, is the smallest scale size capable of diffracting light). Therefore, the maximum transverse extent $\mathbf{r}-\mathbf{r}'$ which can scatter to point $\boldsymbol{p}$ is $\lambda h/l_0$. Physically, the coordinate $\mathbf{r}$ is directly 'above' (in the sense that the z-direction is vertical) the point of interest ($\boldsymbol{\rho}$) in the aperture, and $\mathbf{r}'$ is a horizontal vector which connects $\mathbf{r}$ to another scattering point within the turbulent layer. Any such point with $r \geq \lambda h/l_0$ cannot scatter to coordinate $\boldsymbol{\rho}$. The second term in Eq. 2.95 almost certainly cannot be ignored, but the third term, which is of order $\lambda^4 h/8l_0^4$, can be considered to be very small indeed compared with $\lambda$ using very conservative parameters of $\lambda = 10\mu m$ and $l_0 = 1cm$ for propagation paths of hundreds of kilometres. Neglecting this second term, *i.e.* assuming that:

$$h \ll \frac{8l_0^4}{\lambda^3}, \tag{2.96}$$

is equivalent to the condition for Fresnel diffraction to be applicable [27].

Tatarski [21] and, later, Clifford [28] have found the solution to Eq. 2.94, using Green's solution under the conditions for Fresnel diffraction. The power spectra of phase and log amplitude perturbations at a perpendicular distance h away from the turbulence are found to be

$$\Phi_{\gamma}^{(2)}(h,\kappa) = \pi k^2 h \left(1 - \frac{k}{\kappa h}\sin\left[\frac{\kappa^2 h}{k}\right]\right)\Phi_n^{(3)}(h,\kappa), \tag{2.97}$$

$$\Phi_{\varphi}^{(2)}(h,\kappa) = \pi k^2 h \left(1 + \frac{k}{\kappa h}\sin\left[\frac{\kappa^2 h}{k}\right]\right)\Phi_n^{(3)}(\kappa). \tag{2.98}$$

Next, this result will be applied to an extended turbulent region with constant strength of turbulence $C_n^2$. Use of Eqs. 2.23 and 2.17 have led to expressions for the structure functions of phase and log amplitude for a Kolmogorov spectrum:

$$\begin{bmatrix} \mathcal{D}_{\gamma}(h,r) \\ \mathcal{D}_{\varphi}(h,r) \end{bmatrix} = 1.30\, C_n^2 k^2 h \int_0^{\infty}\left(1 \mp \frac{k}{\kappa h}\sin\left[\frac{\kappa^2 h}{k}\right]\right)\left[1 - J_0(\kappa r)\right]\kappa^{\frac{-8}{3}}\, d\kappa. \tag{2.99}$$

The quantities in round brackets are known as the *filter functions* for phase and log amplitude fluctuations. For short path lengths (*ie* $h \ll \pi k/\kappa^2$), the fluctuations in log amplitude can be considered to be negligible in comparison with the phase fluctuations (the near field region). For very long path lengths (*ie* $h \gg \pi k/\kappa^2$), the fluctuations can be considered equal (the far field region). This confirms Eq. 2.81, but for much longer propagation paths. Goodman [17] has evaluated the wave structure function $\mathcal{D}_\Phi$ (which is the sum of $\mathcal{D}_\varphi$ and $\mathcal{D}_\gamma$ as given in Eq. 2.99),

$$\mathcal{D}_\Psi\left(h, \mathbf{r}\right) = 2.91\, k^2 h C_n^2 r^{\frac{5}{3}}. \tag{2.100}$$

This is *exactly the same* expression calculated for a thin turbulent layer by the undeviating path ray optics approach adopted in section 2.3. It is this sum over many paths through different realisations of turbulence which, as before, leads to log-normal statistics.

The atmosphere must, however, be considered to consist of regions of turbulence with varying $C_n^2$. Tatarski has argued that $C_n^2$ cannot vary significantly over distances of order $L_0$. If the atmosphere can be divided into slabs of approximately this size, each one with a constant $C_n^2$, then the fluctuations in different slabs can be assumed to be uncorrelated (using the same argument as given in section 2.3). In this case, the structure functions of all the layers simply add linearly and the total wave structure function at ground level can be expressed as

$$\mathcal{D}_\Psi\left(0, \mathbf{r}\right) = 2.91\, k^2 r^{\frac{5}{3}} \int_0^h C_n\left(h\right)^2 dh. \tag{2.101}$$

The advantages of the Rytov approximation will now become clear. In obtaining Eq. 2.94, it was assumed that $|\nabla\Lambda_1| \ll |\nabla\Lambda_0|$ (Eq. 2.93). This is a much less restrictive assumption than the one which was necessary to derive the results via the geometric approach, which allowed only small perturbations on the phase, rather than the gradient of phase. Therefore, the theory is now valid for *smooth* changes induced in the phase rather than small changes. Fried [29] equates the dropping of the square of the gradient of the phase (the so-called *Eikonal* term) in Eq. 2.94 to the approximation to wave optics which validates ray tracing. So it is not surprising that the two approaches deliver the same results. Furthermore, the restriction of Eq. 2.96 is clearly far less prohibitive than the maximum propagation path within a layer ($\delta h \ll l_0^2/\lambda$) that allowed a geometric approach to be used.

In so-called strong turbulence, where the phase gradients cannot be considered negligible, it might be expected that the Rytov approximation should break down. Such a failure has been blamed for what is known as the *saturation of scintillation*, where, however strong the refractive index perturbations become, the measured value for the normalised log-amplitude variance does not appear to increase above 50% [29]. Fortunately, this little understood phenomenon is not of importance in the case of astronomical imaging through the atmosphere.

## 2.10 The Time Dependence of Seeing

The theoretical predictions for the temporal characteristics of the wave amplitude at ground level are entirely based on the *Taylor Hypothesis of Frozen Turbulence*, which assumes that the turbulent structures that give rise to the wave perturbations are shifted across the aperture fast enough that they do not have time to evolve significantly. At first sight this hypothesis might appear plausible for small sized apertures. The energy in the fluctuations at small scales of turbulence is derived from the energy at larger scales. Therefore, one might expect the larger scale fluctuations to shift the smaller structures across the entire aperture before those smaller structures have had an opportunity to rearrange themselves. It has already been postulated that most of the turbulence does in fact exist in a number of discrete layers, so it is not unreasonable to model the atmosphere as a series of turbulent layers at different heights each moving at its own transport velocity in a different horizontal direction (clearly velocities in a vertical direction would have little effect on the amplitude at ground level).

This 'shifting' of the phase across the aperture can be used to relate the temporal statistics of the phase to the spatial statistics. This will be done initially for a single layer. In this case, the phase distributions at different times at a height h can be related to the transverse wind velocity $v_\perp$ at that height,

$$\varphi\left(h, \mathbf{r}, t+\tau\right) = \varphi\left(h, \mathbf{r} - \mathbf{v}_\perp \tau, t\right). \tag{2.102}$$

A convenient mathematical tool exists to relate the spatial power spectrum of phase fluctuations to the corresponding temporal power spectrum. This procedure can, in principle, be used to calculate the temporal power spectrum of any experimental or theoretical quantity which has a linear relationship with the phase. For this reason, it is the temporal power spectrum which will be used to describe the statistics of the phase in the time domain.

The analysis of Conan *et al* [30] has two distinguishing features. The first is the recognition that the spatial power spectrum of a phase-related quantity $G$ can be related to the power spectrum of the phase as long as $G$ can be expressed as a convolution of the phase with a 'modulation function',

$$G\left(\mathbf{r}, t\right) = M_G\left(\mathbf{r}\right) \otimes \varphi\left(\mathbf{r}, t\right). \tag{2.103}$$

The modulation function $M$ alters the spatial frequency components of the actual phase to describe the new quantity. This leads directly to the spatial power spectrum of G:

$$\Phi_G\left(\boldsymbol{\kappa}\right) = |\tilde{M}_G\left(\boldsymbol{\kappa}\right)|^2 \Phi_\varphi\left(\boldsymbol{\kappa}\right), \tag{2.104}$$

where $M_G(\mathbf{r})$ and $\tilde{M}_G(\boldsymbol{\kappa})$ are a Fourier Transform pair.

The second feature of the work of Conan *et al* is the use of the Taylor Hypothesis of frozen turbulence to relate the spatial and temporal power spectra. Under the Taylor hypothesis, the change in $G$ over time is equivalent to the distribution in $G$ over the wind direction,

$$G(\mathbf{r}, t+\tau) = G(\mathbf{r} - \mathbf{v}_{\perp}\tau, t). \tag{2.105}$$

If $< G >= 0$, the spatio-temporal covariance of G is defined thus:

$$C_G(\mathbf{r}, \tau) = \langle G(\mathbf{r}', \mathbf{t})\, G(\mathbf{r}' + \mathbf{r}, \mathbf{t} + \tau) \rangle \tag{2.106}$$

so that the temporal covariance can be written

$$C_G(0, \tau) = \langle G(\mathbf{r}, \mathbf{t})\, G(\mathbf{r}, \mathbf{t} + \tau) \rangle. \tag{2.107}$$

From the Taylor Hypothesis this is equivalent to

$$C_G(0, \tau) = C_G(\mathbf{v}_{\perp}\tau, 0), \tag{2.108}$$

a spatial covariance.

Using two relations, firstly that the temporal power spectrum is equal to the Fourier transform of the temporal covariance, and, secondly, that the spatial power spectrum is equal to the Fourier transform of the spatial covariance, one can relate the two power spectra to each other:

$$\tilde{\Phi}_G(\nu) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \Phi_G(\mathbf{f}) \exp(-2i\pi\boldsymbol{\kappa}.\mathbf{v}_{\perp}\tau) \exp(2i\pi\nu\tau)\, d\boldsymbol{\kappa}\, d\tau. \tag{2.109}$$

By choosing the x-axis to be the direction of the wind (so that $\mathbf{v}_x = \mathbf{v}_{\perp}$ and considering a vectorial spatial power spectrum, changing the order of integration leads to:

$$\tilde{\Phi}_G(\nu) = \frac{1}{v_x} \int_{-\infty}^{\infty} \Phi_G\left(\frac{\nu}{v_x}, f_y\right) df_y. \tag{2.110}$$

Use of Eq. 2.104 gives:

$$\tilde{\Phi}_G(\nu) = \frac{1}{v_x} \int_{-\infty}^{\infty} |\tilde{M}_G\left(\frac{\nu}{v_x}, f_y\right)|^2 \Phi_\varphi\left(\frac{\nu}{v_x}, f_y\right) df_y. \tag{2.111}$$

The spatial power spectrum of phase has already been calculated (Eq. 2.83), therefore the temporal power spectrum of phase is given by

$$\tilde{\Phi}_\varphi(\nu) = \frac{0.00969 k^2 C_n^2(h)\, \delta h}{v_x} \int_{-\infty}^{\infty} |\tilde{M}_G\left(\frac{\nu}{v_x}, f_y\right)|^2 \left[\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right]^{-\frac{11}{6}} df_y. \tag{2.112}$$

It is useful at this stage to introduce a frequency, $\nu_g$, known as the Greenwood frequency,

$$\nu_g = \left[0.1024 k^2 C_n^2 (h)\, \delta h v_x^{\frac{5}{3}}\right]^{\frac{3}{5}}. \tag{2.113}$$

The significance of this quantity will be discussed later. This allows the expression for the temporal power spectrum to be simplified:

$$\tilde{\Phi}_G(\nu) = 0.0946 \nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} \int_{-\infty}^{\infty} |\tilde{M}_G\left(\frac{\nu}{v_x}, f_y\right)|^2 \left[\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right]^{-\frac{11}{6}} df_y. \tag{2.114}$$

By far the simplest quantity to calculate from this is the temporal power spectrum of the phase itself. In this case $G = \varphi$ and so $M$ is simply a delta function and its Fourier transform is unity. Therefore,

$$\tilde{\Phi}_\varphi(\nu) = 0.00946 \nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} \int_{-\infty}^{\infty} \left[\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right]^{-\frac{11}{6}} df_y. \tag{2.115}$$

After integration, this becomes:

$$\tilde{\Phi}_\varphi(\nu) = 0.00946 \nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} \left(\frac{\nu}{v_x}\right)^{-\frac{8}{3}} \frac{\pi^{\frac{1}{2}} \Gamma\left(\frac{4}{3}\right)}{\Gamma\left(\frac{11}{6}\right)}. \tag{2.116}$$

$$\tilde{\Phi}_\varphi(\nu) = 0.0159 \nu_g^{\frac{5}{3}} \nu^{-\frac{8}{3}} \tag{2.117}$$

This is the same result as originally calculated by Tatarski [21]. Tatarski, however, gives a limiting frequency range where the result is valid,

$$\nu \ll \frac{v_x}{(\lambda h)^{\frac{1}{2}}}. \tag{2.118}$$

The physical explanation of this can be understood by considering the scintillation pattern produced by the layer. Associated with this pattern is a characteristic frequency of scintillation, $\nu_0$, which is related to the time it takes for a typical scintillation patch (*i.e.* a Fresnel zone) to cross a point in the aperture, and which is dependent on the transverse wind speed of the layer. If the layer in question is moving fast enough so that the Greenwood frequency is comparable with the characteristic frequency of scintillation, then part of the power will be converted into intensity fluctuations. In fact, in the opposing regime, $\nu \gg \frac{v_x}{(\lambda h)^{\frac{1}{2}}}$, the phase fluctuations are reduced by a half and a factor of 0.5 would be seen in Eq. 2.117. This is remarkably similar to the condition of Eq. 2.82, which describes the spatial statistics of phase produced by a layer at a particular height $h$, since the distance that the diffraction pattern moves in one characteristic time ($\sim v_x/\nu_0$) should be much smaller than $\sqrt{\lambda h}$ in the first case and much greater in the second case. These can be likened to the near and far field cases respectively. Nevertheless, the slope of the log of the temporal power spectrum remains the same at both ends of the spectrum.

## 2.11 The Significance of the Greenwood Frequency

For an atmosphere consisting of a number of discrete layers of turbulence, the Greenwood frequency can be generalised thus:

$$\nu_g = \left[0.1024k^2 \int_0^\infty C_n^2 (h)\, v_x{}^{\frac{5}{3}} (h)\; dh\right]^{\frac{3}{5}}, \tag{2.119}$$

where $v_x(h)$ is the wind in the transverse direction of each layer.

The Greenwood frequency was originally defined [31] because of its relevance to adaptive optics systems. An important question is this: if an adaptive optics system can be assumed to provide perfect (spatial) correction, but this correction is only performed at time intervals specified by the system bandwidth $\nu_c$, what is the uncorrected power at the end of the cycle just before the next correction is made? The answer can be given in terms of the 'impulse response' of the system, $H(t, \frac{1}{\nu_c})$. The uncorrected phase at a point in the aperture is then given by:

$$\sigma_\varphi^2 = \int |1 - \tilde{H}(\nu, \nu_c)|^2 \tilde{\Phi}_\varphi(\nu)\; d\nu, \tag{2.120}$$

where $\Phi(\nu)$ is the phase temporal power spectrum and $\tilde{H}(\nu, \nu_c)$ is the Fourier transform of the impulse response of the AO system. Greenwood took the frequency response of such a system to be that of a RC binary filter, where $\tilde{H}(\nu, \nu_c)$ is given by

$$\tilde{H}(\nu, \nu_c) = \frac{1}{1 + i\frac{\nu}{\nu_c}}. \tag{2.121}$$

The physical implication of this is that the lowest frequency components of the phase, where $\tilde{H}$ is close to unity, cause very little error in the phase correction. At higher frequencies, $\tilde{H}$ drops below unity and the contribution to the correction error is greater. However, the fast drop off with frequency of the phase power spectrum ($\sim \nu^{-\frac{8}{3}}$) means that there is relatively little power in these high frequencies. The system bandwidth can be chosen so that the uncorrected power is set to an acceptable level. The Greenwood frequency is defined (Eq. 2.119) so that $\sigma_\varphi = 0.2\pi$. The dependence of the Greenwood frequency on wavelength can also be deduced:

$$\nu_g \propto \lambda^{-\frac{6}{5}}, \tag{2.122}$$

therefore, the wavefront correlation times can be expected to be correspondingly larger at longer wavelengths. Equivalent characteristic frequencies for different correction tolerances can easily be calculated from the Greenwood frequency [32]:

$$\sigma_\varphi^2 \propto \left(\frac{\nu_g}{\nu}\right)^{\frac{5}{3}}. \tag{2.123}$$

Using real data from Colavita *et al* [33], Fried [32] has calculated values for $\nu_g$ of between 20 and 50Hz at visible wavelengths.

It is of interest to quote another result of Fried's. If a perfect correction is made at a particular time and the system is left *static* for a time $\Delta t$ (*i.e.* no system impulse response), then the resultant phase error is given by [34]:

$$\sigma_\varphi^2 = 7.85\,(\nu_g \Delta t)^{\frac{5}{3}}\,, \tag{2.124}$$

in units of radians squared.

## 2.12 The Validity of the Kolmogorov Spectrum

The analysis of the effects of atmospheric turbulence on propagation of light from an astronomical source which has been reviewed in the course of this chapter does not take into account three important contributions to the degradation of the wave, which are conventionally grouped together under the term *dome seeing*. The first is artificial turbulence created inside the telescope building, normally due to localised heat sources. The second is caused by the difference in temperature between the air inside the dome and the air outside. The third results from the difference in wind speed across the boundary between the turbulence outside the dome and the still air inside. Although it is, in principle, possible to eliminate the first two contributions (and even the third if the telescope dome can somehow be removed), many existing large observatories still suffer from these problems. If these problems are very severe, they can make a significant difference to the form of the statistics of the wave at the telescope pupil. Since future observatories can be built to minimise these effects, it is important to be able to separate dome and atmospheric effects. This is not a simple exercise, since any non-Kolmogorov behaviour observed may be attributable either to dome seeing or to the fact that the atmosphere is simply not Kolmogorov some or all of the time. Many attempts have been made to verify the predictions of the Kolmogorov theory of propagation through turbulence (see Chapter 6).

## 2.13 Summary of the Major Results of Kolmogorov Theory

There follows a list of the major points covered in this chapter.

- Existing in the atmosphere are volumes of turbulent refractive index inhomogeneities. The variations cause a phase perturbation to be introduced into waves passing through the region.

- The phase variations are of constant magnitude (statistically) throughout the region and have 'small' gradients.

- Phase perturbations from multiple layers cause variations in both the phase and intensity at ground level and add linearly. These layers are assumed to be in the near-field with respect to a telescope aperture at ground level, so that intensity fluctuations can be neglected relative to phase variations in the aperture.

- It is possible to predict the performance of a telescope receiving light which has propagated through such turbulence by considering the spatio-temporal correlation of the phase incident on the entrance pupil.

- The statistics of the phase, when described in terms of a structure function, are found to obey a five-thirds power law. In fact, these statistics can be *entirely* described by the measurable quantity $r_0$. $r_0$ is actually an integral of the turbulence profile $C_n^2(h)$ weighted by height h.

- Kolmogorov theory can be extended to predict the time characteristics of the seeing. If the atmosphere can be assumed to consist of a number of discrete layers of turbulence each with its own wind direction, the temporal statistics of phase are directly calculable from the spatial ones (the Taylor Hypothesis). In fact, a knowledge of $C_n^2$ and the wind vectors is all that is required.

# Chapter 3

# Wavefront Sensors

## 3.1 The Phase as a Wavefront

In Chapter 2 it was argued that intensity fluctuations in a wave that has propagated through atmospheric turbulence could be neglected in comparison with phase perturbations in the near-field approximation. Under this assumption, a perturbed plane wavefront can be represented by a set of phase values at coordinates through the plane, so that the complex amplitude at coordinate $\mathbf{r}$ can be written as $\Psi(\mathbf{r}) = e^{i\varphi(\mathbf{r})}$ in the near-field. It could equally well be represented by a continuous surface joining points of equal phase. This is the concept of the wavefront and it avoids the difficulties of assigning an absolute value to the phase. The two formulations of the wave as a surface and as a phase distribution in a plane are interchangeable.

For a wave with no intensity variations, the formula for the optical transfer function of the atmosphere/telescope system (Eq. 2.1) can be re-written in terms of an autocorrelation of the complex phase inside the pupil ($P(\mathbf{r})$, now normalised to unit volume),

$$\mathcal{T}(\gamma f) = \left[P(\gamma f)\, e^{i\varphi(\gamma f)}\right] \otimes \left[P(\gamma f)\, e^{i\varphi(\gamma f)}\right]. \tag{3.1}$$

Once again, $\gamma$ is a factor equal to the focal length of the telescope multiplied by the mean wavelength. The OTF is therefore directly obtainable from the knowledge of the wavefront phase. The various techniques used in measurement of the latter are collectively referred to as *wavefront sensing.*

The specification of an adaptive optics system will depend on a knowledge of the spatio-temporal correlation of the phase in the telescope aperture. Because the phase is a Gaussian process the fundamental statistical description is the second-order spatio-temporal correlation function (as in Eq. 2.6),

$$B_{\varphi}(\Delta\mathbf{r}, \Delta t) = \langle \phi(\mathbf{r}, t)\, \phi(\mathbf{r} + \Delta\mathbf{r}, t + \Delta t) \rangle\,, \tag{3.2}$$

or its Fourier transform, the cross-spectrum $\tilde{B}_\varphi(\mathbf{f}, \nu)$. However, from a practical point of view, one is usually interested either in parameters that summarise $B_\varphi(\Delta\mathbf{r}, \Delta t)$ or quantities that express its information in a different form. The structure function of phase (see Eq. 2.7), from which the Fried parameter $r_0$ can be found, and the temporal power spectrum of the spatial Zernike modes are examples of the latter.

## 3.2 Modal Wavefront Representation

A convenient method of describing an instantaneous wavefront $W(\mathbf{r})$ over a finite plane is in terms of a set of suitable basis functions $\omega_i$,

$$W(\mathbf{r}) = \sum_{i=0}^{\infty} a_i \omega_i. \tag{3.3}$$

This is a particularly important formulation for mapping the geometry of individual wavefronts. The process of calculating $W(\mathbf{r})$ from its constituent modes is referred to as *reconstruction.* However, if the time-averaged variances, $< a_i{}^2 >$, of the (infinite number of) individual coefficients are known, then the second-order statistics of the wavefront phase, in the form of, say, the structure function of phase, are also known (the expectation values of the time-averages, $< a_i >$, of the coefficients should be zero because of the zero first-order statistics of the wavefront phase fluctuations). The two statistical descriptions are equivalent so there is no particular advantage to using a modal description of the statistics if the structure function of phase can be measured. Modal descriptions are, nevertheless, useful for describing a partially reconstructed wavefront ($W_n(\mathbf{r})$), *i.e.* a reconstruction of a wavefront formed from a finite number of modes (n):

$$W_n(\mathbf{r}) = \sum_{i=0}^{n} a_i \omega_i. \tag{3.4}$$

Ideally, a set of basis functions which are mutually orthogonal would be chosen. In this way, one could obtain the best possible reconstruction with the smallest number of modes. The most efficient set of functions are the Karhunun-Loeve polynomials [35] [36]. The biggest problem with these functions is that they have no analytic expression and must be calculated numerically. They are also a strong function of the shape of the telescope aperture.

Zernike polynomials [37] [38], although not as efficient as the Karhunen-Loeve functions since there is a degree of overlap between the modes, are by far the most popular modes for describing turbulence-degraded wavefronts over a circular or annular region. This is probably because the lowest orders are closely related to the Seidel aberrations [27] (e.g. tip, tilt, defocus, astigmatism and coma) historically used to describe the performance of optical systems. The low orders are actually quite close to their Karhunen-Loeve counterparts (hence there is very little overlap between them), so they are indeed fairly efficient for describing the wavefront. However, the discrepancy has been shown to become larger in the higher orders [39].

If, for a circular aperture of diameter $D$, a radial polynomial of order $(n, m)$ can be defined in terms of the radial coordinate $r$ as follows [24],

$$R_{nm}(r) = \sum_{s=0}^{(n-m)/2} \frac{(-1)^s (n-s)!}{s! \left(\frac{n+m}{2} - s\right)! \left(\frac{n-m}{2} - s\right)!} \left(\frac{2r}{D}\right)^{n-2s} \tag{3.5}$$

then the $i$th Zernike polynomial can be written in terms of products of these radial polynomials with angular functions:

$$\begin{aligned} Z_{even\ i} &= \sqrt{n+1}\, R_n^m(r) \sqrt{2} \cos m\theta \qquad m \neq 0 \\ Z_{odd\ i} &= \sqrt{n+1}\, R_n^m(r) \sqrt{2} \sin m\theta \qquad m \neq 0 \\ Z_i &= \sqrt{n+1}\, R_n^m(r) \qquad m = 0, \end{aligned} \tag{3.6}$$

where $r$ and $\theta$ are the radial and azimuthal coordinates within the pupil. The coefficient $i$ is numbered from 1 to $\infty$ and is a function of n and m. This relation is shown for $i \leq 10$ in table 3.1.

| i | n | m | |
|---|---|---|---|
| 1 | 0 | 0 | piston |
| 2 | 1 | 1 | tilt |
| 3 | 1 | 1 | tilt |
| 4 | 2 | 0 | defocus |
| 5 | 2 | 2 | astigmatism |
| 6 | 2 | 2 | astigmatism |
| 7 | 3 | 1 | coma |
| 8 | 3 | 1 | coma |
| 9 | 3 | 3 | |
| 10 | 3 | 3 | |

**Table 3.1**: Relation between Zernike numbering coefficient i and the radial degree n and azimuthal coefficient m.

Noll [24] has computed the time-averaged variances of the coefficients of the Zernike modes for Kolmogorov statistics. He gives this information in the form of the rms phase error $\Delta_n$ (in radians) contained in the residual wavefront which would remain if the first $n$ Zernike modes were used to partially reconstruct the wavefront:

$$\Delta_n = \left\langle \left(W_\infty(\mathbf{r}) - W_n(\mathbf{r})\right)^2 \right\rangle, \tag{3.7}$$

where the brackets <> denote an average over the entire pupil and over all time. The energies contained in the modes decreases as the order increases, meaning that it is satisfactory to describe the wavefront in terms of a finite number of modes (although the accuracy of such a description depends on the number of modes used). The first polynomial is known as *piston* (n=0, m=0), which describes the dc component of the wavefront; its variance is infinite. If the piston term is removed (and as far as image formation is concerned it is not important) then the remaining variance is given by:

$$\Delta_1 = 1.029\left(\frac{D}{r_0}\right)^{\frac{5}{3}}. \tag{3.8}$$

85% of the remaining energy is contained in the next two terms (known as *tip* and *tilt*). After these first three modes have been accounted for, the rms difference between the wavefront and its Zernike description (for an aperture with diameter $D$) is

$$\Delta_3 = 0.134\left(\frac{D}{r_0}\right)^{\frac{5}{3}}. \tag{3.9}$$

The numerical values of the time-averaged variances of the first 22 polynomials are also explicitly stated by Noll for Kolmogorov turbulence.

Zernike polynomials represent a convenient, but arbitrary, way of describing a wavefront at a particular instant in time. They can be useful in calculations which model the performance of adaptive optics systems, for example in describing the performance of a partially compensating system, which might only provide compensation up to a certain order of Zernike modes. The variances of the modes of higher orders than this can be used to calculate the image degradation caused by the uncorrected phase. If the phase statistics are known (and for Kolmogorov turbulence they are known if $r_0$ is known), then these variances are also known.

## 3.3 Sensing the Wavefront

As has already been discussed, adaptive optics requires wavefront sensing for two different purposes. Firstly, the measurement of single wavefronts, *i.e.* the wavefront geometry sampled in such a short time so that turbulent fluctuations do not significantly alter the wavefront during that time. This has a particular application in adaptive optics, where continuous full compensation of the wavefront would, hypothetically, lead to diffraction-limited imaging over a long-exposure. Secondly, the measurement of the statistics of such wavefronts. The atmospherically-distorted wavefront is a rapidly fluctuating quantity; site monitoring involves a study of how it varies over long and short timescales.

Measurement of the wavefront phase is an important goal, but not a straightforward one. In general, the complex amplitude distribution at the focal plane of a lens is equivalent to a Fourier Transform (in the appropriate coordinate system) of the wave complex amplitude incident on the lens. However, the amplitude distribution is not measurable; only the intensity distribution is, and this does not provide a unique solution to the phase distribution in the incident wavefront. Therefore, it is impossible to recover the phase distribution of the incident wave merely by examining an image formed from it. Instead, many devices have been conceived which somehow modulate the amplitude distribution in the entrance pupil so that it can be retrieved from a measurable intensity distribution somewhere in the system. One of the most popular of these wavefront sensors will be described in detail in the next section.

It should be recognised that the first-order statistics of the turbulent wavefront are zero (this is intuitively obvious), so that a mean wavefront measured over an infinitely long period is flat; therefore second-order statistics must be considered. These are quantities such as the variance, covariance and structure function of phase, both spatial and temporal. A wavefront sensor used in site monitoring must be capable of measuring these wavefront statistics, or at the very least, measuring some sort of statistics which can be related to these statistics using an appropriate model. The output of each sensor will not be the phase itself, but some quantity related to it. The task, then, is to find a simple relationship between the statistics of the phase itself and the output of each particular wavefront sensor.

Since the wavefront is continuously varying, it is desirable to measure the wavefront within time intervals short enough that the wavefront does not evolve significantly (*i.e.* within one correlation time). Clearly, the longer this time, the more the measured wavefront will tend towards its average shape (*i.e.* flat) and the resolution will appear better than in reality [40]. However, since the wavefront sensing involves intensity measurements, and the numbers of photons used to describe the intensity function are always going to be finite, a compromise must be sought between obtaining enough photons to sample the intensity function accurately and the need to *freeze* the wavefront in time [41]. Integral parts of any wavefront sensor are, then, a suitable detector which can handle these requirements and a reference source which is as bright as possible.

## 3.4 Shack-Hartmann Wavefront Sensor

A portion of an aberrated wavefront incident normally to a lens will form an image whose centre of gravity is a measure of the mean wavefront tilt over that portion. The centre of gravity, or centroid, $(c_x, c_y)$ can be estimated from the coordinates $(x_i, y_i)$ of arrival of the photons at the focal plane of the lens:

$$c_x = \frac{\sum_{i=1}^{n_p} x_i}{n_p} \tag{3.10}$$

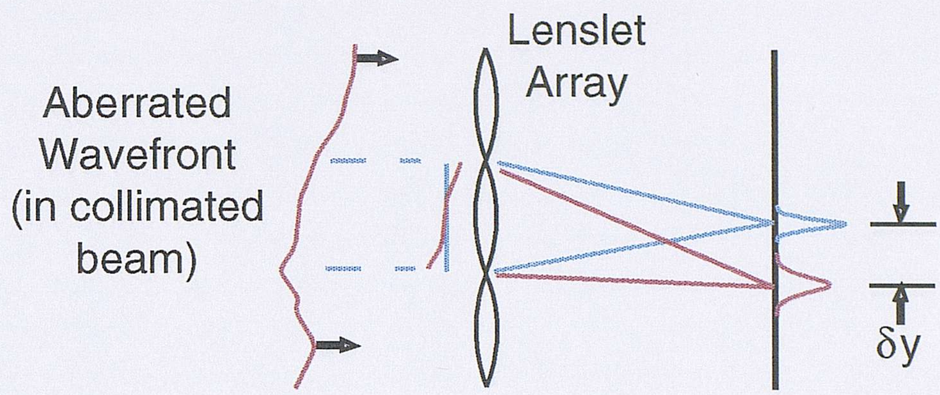



**Figure 3.1**: Principle of the Shack-Hartmann Wavefront Sensor, showing the positions of subaperture images formed by aberrated and unaberrated (reference) wavefronts.

and

$$c_y = \frac{\sum_{i=1}^{n_p} y_i}{n_p}, \tag{3.11}$$

where $n_p$ is the total number of photons passing through the lens (although in practice it is not found in this way; see section 3.7.3).

The Shack-Hartmann wavefront sensor consists of a large number of such lenses placed in the entrance pupil plane of the telescope, so that the full telescope aperture is divided into a number of 'subapertures'. These lenslets form a pattern of images which can be recorded by a suitable camera/detector. The lenslets sample the tilts in various parts of the wavefront so that the full wavefront may, in principle, be reconstructed to some degree of accuracy. If a time series of images is recorded, the evolution of the wavefront can be observed. The optics used is normally a microlens array due to the ease with which such a device can be manufactured. Referring to Fig. 3.1, the mean wavefront tilt (or angle of arrival of the incoming ray) in the y-direction is equal to $\delta y/l$ where $\delta y$ is the shift of the centroid of the sub-image away from its reference position (defined as the centroid of an image formed from a non-aberrated wavefront) in the pupil plane and $l$ is the focal length of the lenslet. This reasoning can easily be extended to find the tilts in two orthogonal directions in the image plane. The popular implementation of the Shack-Hartmann sensor involves placing a microlens array in a beam of light from the telescope which has been collimated (although a novel type of Shack-Hartmann sensor will be discussed in Chapter 5.

The Shack-Hartmann sensor is a popular choice of sensor for use in an adaptive optics system. In this case, the centroid from each lenslet is used to compensate the phase over the pupil area corresponding to that lenslet. The very best compensation which can be achieved is by correcting the two tilt terms across this region; any higher order distortions are not detectable in the centroid motion (although it is suggested that the shape of the subaperture image can give information about higher order distortions [42]). Eq. 3.9 (see earlier) thus gives a lower limit on the rms phase error over the area defined by each lenslet (this is without considering errors caused in the reconstruction process). Clearly, the error is very sensitive to the ratio $D/r_0$; the usual criterion is that the lenslet size $D$ should be less than or equal to $r_0$. A trade-off occurs between the need to sample the wavefront finely enough and the need to keep the photon count per lenslet per sampling interval high enough to obtain an accurate centroid estimation. One of the disadvantages of this technique is that it is impossible to change the system in response to fluctuations in $r_0$. A knowledge of the variation in $r_0$ is therefore essential before an adaptive optics system can be constructed; this emphasizes the importance of site monitoring experiments.

## 3.5 Use of the Shack-Hartmann Sensor in Seeing Measurements

As will be discussed in section 3.9, the Shack-Hartmann Wavefront Sensor is a particularly suitable choice for use in a site monitoring experiment. It is of use, then, to discuss what type of measurements are appropriate in such an experiment. The spatio-temporal correlation of the wavefront would give all the information that is required; however, it is not easy to directly infer this from a series of Shack-Hartmann images. However, in the rather bold assumption that the wavefront obeys Kolmogorov power laws, it is possible to account for these statistics using a very small number of parameters which, it transpires, can be measured by the Shack-Hartmann sensor. In this case, the Fried parameter $r_0$ and the Greenwood frequency $\nu_g$ (the latter under the assumption of the Taylor hypothesis also) completely describe the correlation of the phase. Of course, this assumption a strong one and the effects of non-Kolmogorov statistics will be reviewed and measured in Chapter 6. Nevertheless, it is important to be able to measure $r_0$ (and similar parameters in the temporal domain) directly from the centroid statistics. The Shack-Hartmann wavefront sensor makes an estimate of the first derivative (or *angle-of-arrival*) of the incoming wave (which is in some way averaged over the finite extent of the subaperture). Therefore it is vital to consider the relationship between the statistics of the wavefront and those of its first derivative. Fortunately, under a Kolmogorov assumption, there is a well-defined relationship between the two, and so one is directly calculable from the other. For this reason, the statistics of the angle-of-arrival will be reviewed here.

## 3.6 Angle of Arrival Statistics

The Shack-Hartmann sensor makes estimates of the first derivative of the wavefront, so it is of importance to consider the statistics of the angle-of-arrival for a Kolmogorov atmosphere. In reality, the finite size of the subaperture will distort the frequency spectra of the measured statistics, and these *aperture-averaging* effects must be taken into consideration. The measured quantity from the Shack-Hartmann wavefront sensor is the centroid position from each lenslet. A relationship will first be sought between the statistics of the centroid motions and the wavefront angle-of-arrival statistics. These statistics will be then related to the statistics of the wavefront itself, characterised by means of the Fried parameter and the Greenwood frequency (for Kolmogorov turbulence), in the spatial and temporal domains respectively. The centroid motions can be measured in a variety of ways; the five most important measured quantities which can be related to the phase statistics are the spatial covariance, the spatial variance, the spatial differential variance, the temporal power spectrum and the temporal differential power spectrum of the centroid motions.

### 3.6.1 The Centroid and Angle-of-Arrival

The angle-of-arrival of the wavefront at some point in the measurement plane $(x, y)$ is, by convention, related to the derivative of the phase in the following way:

$$\alpha(x,y) = -\frac{\lambda}{2\pi}\frac{\partial}{\partial x}\varphi(x,y). \tag{3.12}$$

The significance of the negative sign will become clear.

The centroid of an image produced at the focal plane of a lenslet might be expected to be related to the angle-of-arrival of the wavefront at the lenslet (see Fig. 3.1). For an infinitesimally small lens, the centroid position (in angular coordinates) is equal to the angle-of-arrival of the wave. For a finite sized lens, the effects of aperture averaging must be investigated. In order to find the exact relationship, the problem will be considered for a one-dimensional complex amplitude distribution (although the reasoning for the two-dimensional case is similar). The method which will be used is essentially the same as Tatarski's [21], although some well known relationships from basic Fourier theory will be used to simplify the calculation. According to Goodman [43], the angular complex amplitude distribution $A(x)$ in the focal plane of a lens formed from an incident wave with amplitude distribution $\tilde{A}(x')$ can be written as:

$$A(x) = \frac{1}{\lambda}\int\int \tilde{A}(x')\exp\left[\frac{-2\pi i}{\lambda}xx'\right]dx'. \tag{3.13}$$

This equation is valid in the paraxial approximation, where quadratic effects can be ignored (*i.e.* in reality there is a slight reduction in the intensity away from the optical axis). Thus the image plane distribution is equivalent to a Fourier transform of the incident amplitude distribution in reduced coordinates of $x'/\lambda$.

The x-coordinate centroid of the image formed by such a lens defined in Eq. 3.10 can be re-written as:

$$c_x = \frac{\int I(x) \cdot x \, dx}{\int I(x) \, dx} = \frac{\int A^*(x) A(x) \cdot x \, dx}{\int A^*(x) A(x) \, dx} \tag{3.14}$$

For the case of no scintillation, the incident amplitude can be written as

$$\tilde{A}(x') = e^{i\varphi(x')}, \tag{3.15}$$

and thus the denominator can be ignored in Eq. 3.14. Parseval's theorem [44] will be used:

$$\int fg^* \, dx = \int FG^* x' \tag{3.16}$$

where $f(x)$ and $F(x')$ and $g(x)$ and $G(x')$ are Fourier transform pairs. Also using the Fourier transform *derivative rule* [44], where the Fourier transform of $-2\pi i \partial \tilde{A}(x')/\partial x')$ is equal to $xA(x)$ (in appropriate coordinates), the centroid can be written as:

$$c_x = \frac{-2\pi i}{\lambda} \int \tilde{A}(x') \frac{\partial \tilde{A}^*(x')}{\partial x'} \, dx'. \tag{3.17}$$

Putting in Eq. 3.15, for an amplitude distribution with no scintillation,

$$\begin{aligned} c_x &= \frac{-2\pi i}{\lambda} \int \frac{\partial \varphi(x')}{\partial x'} \\ &= \int \alpha(x') \, dx' \\ &= \bar{\alpha}. \end{aligned} \tag{3.18}$$

Thus the centroid position (in angular coordinates) is equal to the aperture-averaged angle-of-arrival (here denoted by $\bar{\alpha}$).

### 3.6.2 Spatial Covariance

The most straightforward measurable quantity to which it is possible to relate $r_0$ is the angle-of-arrival covariance (not aperture-averaged),

$$C_\alpha(x, y) = \langle \alpha(x', y') \, \alpha(x' + x, y' + y) \rangle, \tag{3.19}$$

Following Roddier [12], the power spectrum of the angle-of-arrival can be related to the power spectrum of phase fluctuations by using the relation between a quantity and its derivative in Fourier space (the derivative rule once more; see section 3.6.1):

$$\Phi_\alpha(\mathbf{f}) = \lambda^2 f_x^2 \Phi_\varphi(\mathbf{f}), \tag{3.20}$$

where $\mathbf{f_x}$ is the x-component of the frequency vector $\mathbf{f} = \sqrt{\mathbf{f_x^2} + \mathbf{f_y^2}}$.

Since the covariance of a quantity is simply the Fourier transform of its power spectrum (Eq. 2.14), the covariance of angle-of-arrival along, say, the x-direction can now be related to the covariance of phase,

$$C_\alpha(x,y) = -\frac{\lambda^2}{4\pi^2}\frac{\partial^2}{\partial x^2}C_\varphi(x,y). \tag{3.21}$$

Using equation 2.10 it could equally well be related to the structure function of phase,

$$C_\alpha(x,y) = -\frac{\lambda^2}{8\pi^2}\frac{\partial^2}{\partial x^2}D_\varphi(x,y). \tag{3.22}$$

In the near-field approximation (2.80, 2.62),

$$D_\varphi(x,y) = 6.88 r_0^{-\frac{5}{3}}\left(x^2+y^2\right)^{\frac{5}{6}}. \tag{3.23}$$

The second derivative of $D_\varphi$ is easily calculated and equation 3.22 yields results for the following two measurable quantities,

$$C_\alpha(d,0) = 0.097\lambda^2 r_0{}^{-\frac{5}{3}}d^{-\frac{1}{3}}, \tag{3.24}$$

$$C_\alpha(0,d) = 0.145\lambda^2 r_0{}^{-\frac{5}{3}}d^{-\frac{1}{3}}. \tag{3.25}$$

One would wish to apply this result to measurements made with two apertures of diameter D separated by a distance d in the aperture. In fact, aperture averaging effects can be neglected when the diameter of the apertures is less than about twice their separation (x or y) due to the slow variation of $C_\alpha$ on $d$. The covariance from finite sized apertures can be considered to be an average of the interactions between each point in one aperture and each point in the other. The variation in the separation of these two points is then relatively small relative to the separation of the centres of the apertures, $d$, so that the $d^{\frac{1}{3}}$ factor (see Eqs. 3.24 and 3.25) can be considered the same for all such pairs of points. Therefore, in this regime, the covariance of the centroid motions (scaled from detector to angular coordinates) can be considered the same as that of the angle-of-arrival fluctuations, *i.e.*

$$C_\alpha(x,y) = C_{\bar{\alpha}}(x,y). \tag{3.26}$$

### 3.6.3 Single Aperture Variance

An easily measurable quantity is the angle-of-arrival variance as seen through a finite-sized aperture. Most frequently, apertures are either square or circular in shape, so the variance will be derived for both of these. The variance of a quantity is simply its power spectrum integrated over all frequencies:

$$\sigma_G^2 = \int_{-\infty}^{\infty} \Phi_G(f)\, df. \tag{3.27}$$

Using Conan's formulation once more, the aperture-averaged angle-of-arrival can be expressed as a convolution of the phase with a modulation function (as in Eq. 2.103):

$$\bar{\alpha}(\mathbf{r},t) = M_{\bar{\alpha}}(\mathbf{r}) \otimes \varphi(\mathbf{r},t). \tag{3.28}$$

with

$$M_{\bar{\alpha}}(\mathbf{r}) = -\frac{\lambda}{2\pi} P(\mathbf{r}) \otimes \frac{\partial}{\partial x}. \tag{3.29}$$

Here $P$ is the *normalised* pupil function which allows for the effects of aperture-averaging (the same as the definition used in Eq. 2.1). It is the Fourier transform of $M$, *i.e.* $\tilde{M}(\mathbf{f})$, which will be used to relate the spatial power spectrum of phase to that of angle-of-arrival (as in Eq. 2.104):

$$\Phi_{\bar{\alpha}}(\boldsymbol{\kappa}) = \left|\tilde{M}_{\bar{\alpha}}(\boldsymbol{\kappa})\right|^2 \Phi_{\varphi}(\boldsymbol{\kappa}). \tag{3.30}$$

For a circular aperture of diameter $D$,

$$\tilde{M}_{\bar{\alpha}}(\mathbf{f}) = i\lambda f_x \frac{2J_1(\pi D f)}{\pi D f}, \tag{3.31}$$

(since the Fourier transform of the differential operator, $\partial/\partial x$, is $-2i\pi f_x$). Using the result for the near-field Kolmogorov spatial power spectrum given in terms of $r_0$ in Eq. 2.85, the centroid variance (in angular coordinates) can then be written as:

$$\sigma_{\bar{\alpha}}^2 = 0.0229\lambda^2 r_0^{-\frac{5}{3}} \frac{4}{\pi^2 D^2} \cdot I_1(D), \tag{3.32}$$

where $I_1(D)$ is the integral:

$$I_1(D) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_x^2 \left(J_1\left[\pi D \left(f_x^2+f_y^2\right)^{\frac{1}{2}}\right]\right)^2 \left(f_x^2+f_y^2\right)^{-\frac{17}{6}} df_x\, df_y. \tag{3.33}$$

This integral can be performed in polar coordinates by setting $x = Df_x$, $y = Df_y$, and $r = \sqrt{x^2+y^2}$. Then,

$$I_1(D) = D^{\frac{5}{3}} \int_0^{2\pi} \cos^2\theta\, d\theta \int_0^{\infty} \left(J_1[\pi r]\right)^2 r^{-\frac{8}{3}}\, dr \tag{3.34}$$

$$= \pi D^{\frac{5}{3}} \int_0^{\infty} \left(J_1[\pi r]\right)^2 r^{-\frac{8}{3}}\, dr \tag{3.35}$$

$$= \pi D^{\frac{5}{3}} \frac{\pi^{\frac{5}{3}}\Gamma\left(\frac{8}{3}\right)\Gamma\left(\frac{1}{6}\right)}{2^{\frac{8}{3}}\Gamma\left(\frac{11}{6}\right)\Gamma\left(\frac{17}{6}\right)\Gamma\left(\frac{11}{6}\right)} \tag{3.36}$$

$$= 18.30 D^{\frac{5}{3}}. \tag{3.37}$$

Therefore

$$\sigma_{\tilde{\alpha}}^2 = 0.170 \frac{\lambda^2}{D^{\frac{1}{3}} r_0^{\frac{5}{3}}}. \tag{3.38}$$

This is exactly the same as the result originally derived by Tatarski [21]. The variance of the centroid motion in two dimensions (*i.e.* the x-y plane) is simply twice that of $\sigma_{\tilde{\alpha}}^2$.

The calculation can be repeated for a square aperture with an area of $D^2$. In this case, the frequency modulation function is given by:

$$\tilde{M}_{\tilde{\alpha}}(\mathbf{f}) = i\lambda f_x \frac{2 \sin(\pi D f)}{\pi D f}. \tag{3.39}$$

Following the same procedure as for the circular aperture:

$$\sigma_{\tilde{\alpha}}^2 = 0.0229 \lambda^2 r_0^{-\frac{5}{3}} \frac{4}{\pi^2 D^2} \cdot I_2(D), \tag{3.40}$$

with $I_2(D)$ as

$$I_2(D) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_x^2 \left( sin \left[ \pi D \left( f_x^2 + f_y^2 \right)^{\frac{1}{2}} \right] \right)^2 \left( f_x^2 + f_y^2 \right)^{-\frac{17}{6}} df_x \, df_y. \tag{3.41}$$

As before,

$$I_2(D) = \pi D^{\frac{5}{3}} \int_0^{\infty} (sin[\pi r])^2 r^{-\frac{8}{3}} dr \tag{3.42}$$

$$= \pi D^{\frac{5}{3}} \cos\left(-\frac{5\pi}{6}\right) \Gamma\left(-\frac{5}{3}\right) 2^{\frac{2}{3}} \left(\frac{\pi}{2}\right)^{\frac{5}{3}}. \tag{3.43}$$

Therefore

$$\sigma_{\tilde{\alpha}}^2 = 0.161 \frac{\lambda^2}{D^{\frac{1}{3}} r_0^{\frac{5}{3}}}. \tag{3.44}$$

A comparison can be drawn between the results for circular and square apertures by considering the two to have the same area. A square aperture with an area of $\pi D^2/4$ would have a variance of:

$$\sigma_{\tilde{\alpha}}^2 = 0.174 \frac{\lambda^2}{D^{\frac{1}{3}} r_0^{\frac{5}{3}}}. \tag{3.45}$$

This is $\approx$ 2% larger than the variance calculated for a circular aperture. It is often customary to use relations derived for a circular aperture when using a square aperture (whose area defines the radius of the equivalent circle).

### 3.6.4 Two-Aperture Differential Variance

The two methods described so far for measuring $r_0$ have one limitation in common; they are both sensitive to what are known as 'tracking errors', where the reference positions of the images move during an observation. This is a common feature of any telescope system, whose task it is to track a star as it moves across the sky. Non-uniformities in the tracking motion and other vibrations, such as wind buffeting, may cause the telescope to move unpredictably with respect to the actual position of the star in the sky. The basis of what has now become known as the *seeing monitor* is to choose two apertures within the pupil and measure the differential variance between their centroid motions. Any movement in their reference positions is removed by taking the difference of the two centroid positions, so this technique is insensitive to tracking errors.

Following the reasoning of Sarazin and Roddier [45], the differential variance of image motions $\mathcal{D}_{\bar{\alpha}}$, so labelled because of its equivalence to a structure function, can be related to the covariance $C_{\bar{\alpha}}$ via Eq. 2.10 which can be rewritten as:

$$\mathcal{D}_{\bar{\alpha}}(d) = 2\left[C_{\bar{\alpha}}(0) - C_{\bar{\alpha}}(d)\right], \tag{3.46}$$

The covariance of the image motions is given in Eqs. 3.24 and 3.25 for the longitudinal and transverse cases respectively. These equations are, however, not valid for values of $d$ less than about $2D$, so the expression for the centroid variance will be used instead at $d = 0$, *i.e.*

$$C_{\bar{\alpha}}(0) = \sigma_{\bar{\alpha}}^2. \tag{3.47}$$

The differential variance can, once again, be divided into its two components; longitudinal ($\mathcal{D}_{\bar{\alpha}_{\|}}$), which refers to the variance of the difference of the two image positions as measured along the axis connecting the two reference position, and transverse ($\mathcal{D}_{\bar{\alpha}_{\perp}}$), measured perpendicular to this axis:

$$\mathcal{D}_{\bar{\alpha}_{\|}}(d) = 2\left[\sigma_{\bar{\alpha}}^2 - C_{\bar{\alpha}}(d,0)\right], \tag{3.48}$$

$$\mathcal{D}_{\bar{\alpha}_{\perp}}(d) = 2\left[\sigma_{\bar{\alpha}}^2 - C_{\bar{\alpha}}(0,d)\right], \tag{3.49}$$

which leads to the result:

$$\mathcal{D}_{\bar{\alpha}_{\|}}(d) = \left[1 - 0.571S^{-\frac{1}{3}}\right]\sigma_{\alpha}^2, \tag{3.50}$$

$$\mathcal{D}_{\bar{\alpha}_{\perp}}(d) = \left[1 - 0.852S^{-\frac{1}{3}}\right]\sigma_{\alpha}^2. \tag{3.51}$$

Sarazin and Roddier [45] used a coefficient of 0.179 instead of 0.170 for the single aperture differential variance (see Eq. 3.38); no explanation is given for this (although it is possible that the expression for variance of the Zernike tilt term, which is similar but not exactly the same, was used incorrectly). Therefore, they have obtained coefficients of $S^{-\frac{1}{3}}$ of 0.541 and 0.811 in Eqs. 3.50 and 3.51 respectively.

This result can be considered no more valid than the expression for the covariance which was used to derive it (*i.e.* it becomes more accurate as $d$ increases, and is useful for $d \geq 2D$). Fried [46] has calculated a general expression for the differential variance of centroid motions for two circular apertures at any orientation which can be applied to any expression for the structure function of phase fluctuations (and therefore any model of turbulence).

$$\begin{aligned}\left\langle(\alpha_1-\alpha_2)^2\right\rangle = &\left(\frac{4}{\pi}\right)^4\left(\frac{\lambda}{D}\right)^2\int_0^{2\pi}d\theta\int_0^1 u\,du\\ &\left[\frac{1}{8}\cos^{-1}u+\left(1-u^2\right)^{\frac{1}{2}}\left[\left(\frac{u^3}{12}-\frac{5u}{24}\right)+\left(\frac{u^3}{3}-\frac{u}{3}\right)\cos^2\theta\right]\right]\\ &\cdot\left[\frac{1}{2}\mathcal{D}_\varphi\left(D\left[S^2+2Su\cos(\theta+\psi)+u^2\right]^{\frac{1}{2}}\right)\right.\\ &\left.+\frac{1}{2}\mathcal{D}_\varphi\left(D\left[S^2-2Su\cos(\theta+\psi)+u^2\right]^{\frac{1}{2}}\right)-\mathcal{D}_\varphi(Du)\right].\qquad(3.52)\end{aligned}$$

Here, $D$ is the diameter of a single aperture, $S$ is the ratio of the sub-aperture separation to $D$, and $\psi$ is the angle between the x-axis and the line joining the centres of the two apertures.

Figure 3.2 shows a comparison between values calculated numerically from Eq. 3.52 using the Kolmogorov structure function (solid line) and the ratio of values calculated from Eqs. 3.50 and 3.51 (dotted line). Values calculated from Sarazin and Roddier's slightly different formulae are also shown (dashed line).

### 3.6.5 Single Aperture Temporal Power Spectrum

#### Single Layer Spectrum

The calculation of the aperture-averaged angle-of-arrival temporal power spectra is a straightforward matter. Conan's formulation will be used to obtain the temporal power spectrum from the frequency modulation function $\tilde{M}_{\vec{\alpha}}$ (defined in section 3.6.3) and the spatial power spectrum of phase fluctuations. Equation 2.114 will be repeated here:

$$\tilde{\Phi}_G(\nu) = 0.0946\nu_g^{\frac{5}{3}}v_x^{-\frac{8}{3}}\int_{-\infty}^{\infty}\left|\tilde{M}_G\left(\frac{\nu}{v_x},f_y\right)\right|^2\left[\left(\frac{\nu}{v_x}\right)^2+f_y^2\right]^{-\frac{11}{6}}df_y. \qquad (3.53)$$

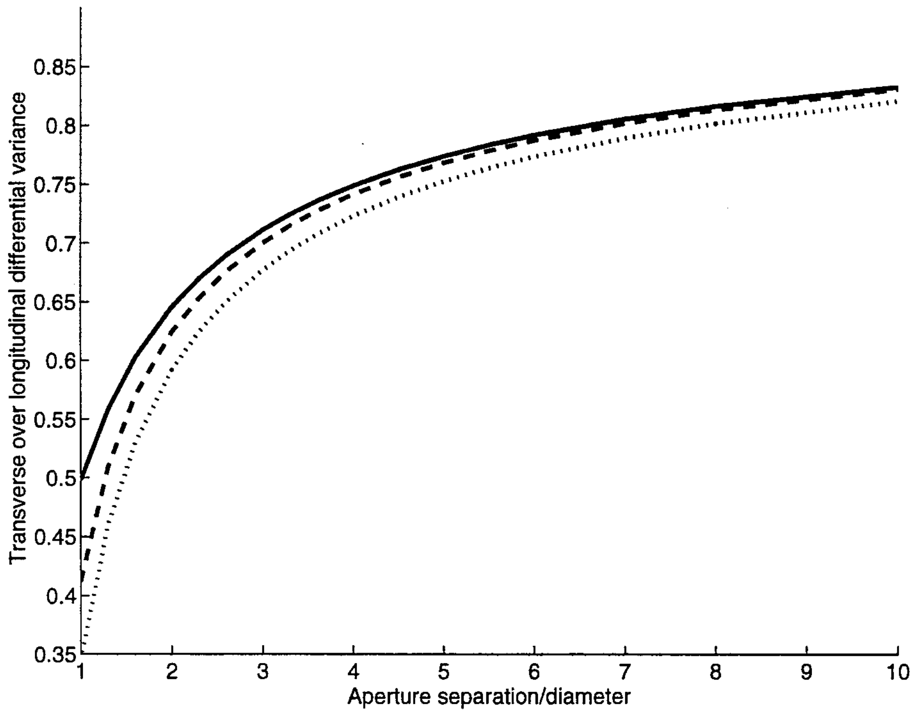



**Figure 3.2**: A comparison between different methods of calculating the two-aperture differential variance ratio for different aperture separations; solid line - values calculated using a numerical integration of Eq. 3.52; dotted line - values calculated from the semi-empirical formulae of Eqs. 3.50 and 3.51; dashed line - values calculated from the (slightly different) formulae of Sarazin and Roddier.

For a circular aperture of diameter D, the function $\tilde{M}$ is given by Eq. 3.31. Hence,

$$\tilde{\Phi}_{\bar{\alpha}}(\nu) = 0.0946\frac{4}{\pi^2}\nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} D^{\frac{2}{3}} \cdot I_3\left(\left(\frac{\nu}{v_x}\right), D\right), \tag{3.54}$$

where the integral $I_3$ is equal to

$$I_3\left(\left(\frac{\nu}{v_x}\right), D\right) = D^{-\frac{8}{3}} \int_{-\infty}^{\infty} \left(\frac{\nu}{v_x}\right)^2 \left(J_1\left[\pi D\left(\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right)^{\frac{1}{2}}\right]\right)^2 \cdot \left[\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right]^{-\frac{17}{6}} df_y. \tag{3.55}$$

A $D^{\frac{8}{3}}$ factor has been taken outside the integral for convenience later. By setting $a = \nu/v_x$ and $y = Df_y$,

$$I_3(a, D) = \int_{-\infty}^{\infty} (aD)^2 \left(J_1\left[\pi\left((Da)^2 + y^2\right)^{\frac{1}{2}}\right]\right)^2 \left[(Da)^2 + y^2\right]^{-\frac{17}{6}} dy. \tag{3.56}$$

the integration can be performed numerically.

A graph of $I_3$ against $(aD)$ is shown in a log-log form in Fig. 3.3. $I_3$ shows the exact shape of the temporal power spectrum, the actual power being a function of $\nu_g$, $v_x$ and $D$. There is a sharp cut-off at a frequency of around $\nu_c$, shown by the dotted line, where

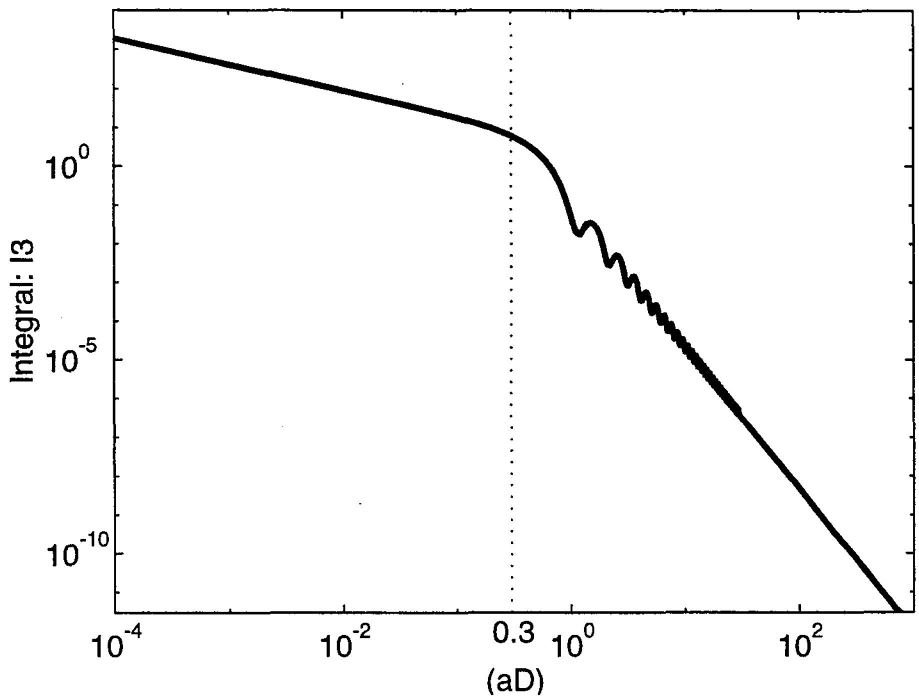



**Figure 3.3**: Graph of the integral $I_3$ against $(aD)$, used to calculate the temporal power spectrum of the centroid motions of a single circular aperture for Kolmogorov turbulence. This graph has been altered since submission.

$\nu_c = 0.3v_x/D$. Below $\nu_c$, the slope of the graph is $-2/3$. Above this frequency, the slope is $-11/3$ and one can write $I_3 = \chi_0(aD)^{-11/3}$ where $\chi_0$ is found from the graph to be approximately $-0.94$. Both of these slopes agree with those found by Conan [30].

Conan has also investigated the behaviour of the temporal power spectrum for square apertures. The form of the curve is found to be the same as for the circular aperture, except in the high frequency domain where the slope depends on the orientation of the aperture with respect to the wind direction. In this case the 'corner' regions of the square have a significant effect on the behaviour. Assuming a single turbulent layer, it is possible to obtain slopes of between $\nu^{-14/3}$ and $\nu^{-11/3}$.

A method is now proposed for determining the Greenwood frequency from an experimentally found graph of the power spectrum against frequency. From equation 3.54 the power spectrum in the region above $\nu_c$ can be written as:

$$\tilde{\Phi}_{\tilde{\alpha}}(\nu) = 0.0946\frac{4}{\pi^2}\nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} D^{\frac{2}{3}} \chi_0 \left(\frac{v_x}{D}\right)^{\frac{11}{3}} \nu^{-\frac{11}{3}}. \tag{3.57}$$

The coefficient $\chi$ of a log-log graph of measured power spectrum against frequency in this $-11/3$ slope region is estimated so that:

$$\tilde{\Phi}_{\tilde{\alpha}}(\nu) = \chi\nu^{-\frac{11}{3}}. \tag{3.58}$$

Combining Eqs. 3.57 and 3.58, and using the relation $\nu_c = 0.3v_x/D$ gives:

$$\frac{\chi}{\chi_0} = 0.038\frac{\nu_g^{\frac{5}{3}}}{D^2}\frac{\nu_c}{0.3}. \tag{3.59}$$

By measuring $\chi$ and estimating the cutoff frequency $\nu_c$, the Greenwood frequency is estimated thus,

$$\nu_g^{\frac{5}{3}} = 0.8\frac{D^2}{\nu_c}\frac{\chi}{\chi_0}. \tag{3.60}$$

It is important to recognise that the measured frequency depends on the direction of the wind relative to the direction of the measurement of the centroid position (*i.e.* the x-direction here). By measuring the cut-off frequencies in orthogonal directions, one can deduce the true wind direction and perform the calculation on the centroid power spectrum in this direction.

### Multi-Layer Spectrum

Of course, the analysis just described assumes not only that the Taylor hypothesis holds, but also that there is one single layer. In general, the temporal power spectra resulting from the *transverse* wind vectors from many layers will simply add, assuming that the phase distributions in each are independent. It is useful here to repeat the definition of the Greenwood frequency (Eq. 2.119):

$$\nu_g = \left[0.1024k^2 \int_0^\infty C_n^2(h)\, v_x^{\frac{5}{3}}(h)\, dh\right]^{\frac{3}{5}}. \tag{3.61}$$

A number of authors [47] [33] have suggested that this could be written as

$$\nu_g = \left[0.1024k^2\overline{v_x}^{\frac{5}{3}} \int_0^\infty C_n^2(h)\, dh\right]^{\frac{3}{5}}, \tag{3.62}$$

where $\overline{v_x}$ is an effective wind speed resulting from many different layers. Therefore,

$$\nu_g = \left[0.242{r_0}^{-1}\overline{v_x}^{\frac{5}{3}}\right]^{\frac{3}{5}}. \tag{3.63}$$

The resulting shape of the multi-layer spectra would be as follows: Instead of the sharp cutoff frequency observed in Fig. 3.3, a transition region would occur between frequencies of $0.3v_{min}/D$ and $0.3v_{max}/D$, where $v_{min}$ and $v_{max}$ are the transverse velocities (in the direction orthogonal to centroid measurement) of the slowest and fastest layers.

Above and below this transition region, the slopes would be the same as for the single layer spectrum. Therefore, one could still make an approximate estimate of the Greenwood frequency by observing where $\nu_c$ occurs, as long as $r_0$ could be measured using, say, the differential image motion method of section 3.6.4.

### 3.6.6 Differential Centroid Motion Power Spectrum

Once again, the centroid motion from a single aperture is subject to tracking errors. Its temporal power spectrum thus contains a component which is the power spectrum of the tracking error. This may appear in the form of anomalous peaks in the measured power spectrum. If the tracking errors are bad, it may be impossible to measure the slopes involved or even to see where the cutoff frequency lies. It would be much more satisfactory to use the power spectrum of the differential image motions.

It is immediately obvious from Eqs. 3.50 and 3.51 that the area under this power spectrum will depend on the separation of the apertures, d. The modulation function for the two apertures can be written:

$$\tilde{M}_{D_{\parallel}}(\mathbf{f}) = i\lambda f_x \sin(\pi \mathbf{d}.\mathbf{f}) \frac{2J_1(\pi D f)}{\pi D f}, \tag{3.64}$$

if the differential measurements are chosen to be made in the longitudinal direction. If the apertures are oriented along the x-direction, then,

$$\tilde{\Phi}_{D_{\parallel}}(\nu) = 0.0946 \frac{4}{\pi^2} \nu_g^{\frac{5}{3}} v_x^{-\frac{8}{3}} D^{\frac{2}{3}} \cdot I_4\left(\left(\frac{\nu}{v_x}\right), D, S\right), \tag{3.65}$$

where the integral $I_4$ is equal to

$$I_4\left(\left(\frac{\nu}{v_x}\right), D, S\right) = D^{-\frac{8}{3}} \int_{-\infty}^{\infty} \left(\frac{\nu}{v_x}\right)^2 \sin^2\left(\pi d\left(\frac{\nu}{v_x}\right)\right) \cdot \left(J_1\left[\pi D\left(\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right)^{\frac{1}{2}}\right]\right)^2 \left[\left(\frac{\nu}{v_x}\right)^2 + f_y^2\right]^{-\frac{17}{6}} df_y. \tag{3.66}$$

Again, by setting $a = \nu/v_x$ and $y = Df_y$ and with $S = d/D$,

$$I_4(a, D) = \int_{-\infty}^{\infty} (aD)^2 \sin^2(\pi S D a) \cdot \left(J_1\left[\pi\left((Da)^2 + y^2\right)^{\frac{1}{2}}\right]\right)^2 \left[(Da)^2 + y^2\right]^{-\frac{17}{6}} dy. \tag{3.67}$$

This integral can also be performed numerically for different values of $S$, and the results are displayed in Fig. 3.4. At very low frequencies, $I_4$ approximates to the value of $I_3$ (see earlier) multiplied by a factor of $(aD)^2$ (since $\sin(\pi SDa) \approx \pi SDa$). Hence a slope of $+4/3$ is observed; the magnitude depends on $S$. Once again, there is a cut-off frequency of $\nu_c \approx 0.3 v_x/D$. Above this frequency, the behaviour is more complex. Discontinuities occur at values of $SDa = m$, where $m$ is any integer. In fact, the high frequency end of the spectrum is enveloped by the cyclic $\sin^2$ term in Eq. 3.67. However, if this oscillating behaviour is ignored, the average value of the power spectrum over each sinusoidal cycle can be calculated; $< \sin^2(\pi SaD) >= \pi$, therefore in this region, $I_4 \approx \pi I_3$ and the differential power spectrum is simply a factor of $\pi$ larger than the single centroid power spectrum (in the direction of the axis joining the differential pair) in the high frequency regime (*i.e.* with an $-11/3$ slope).

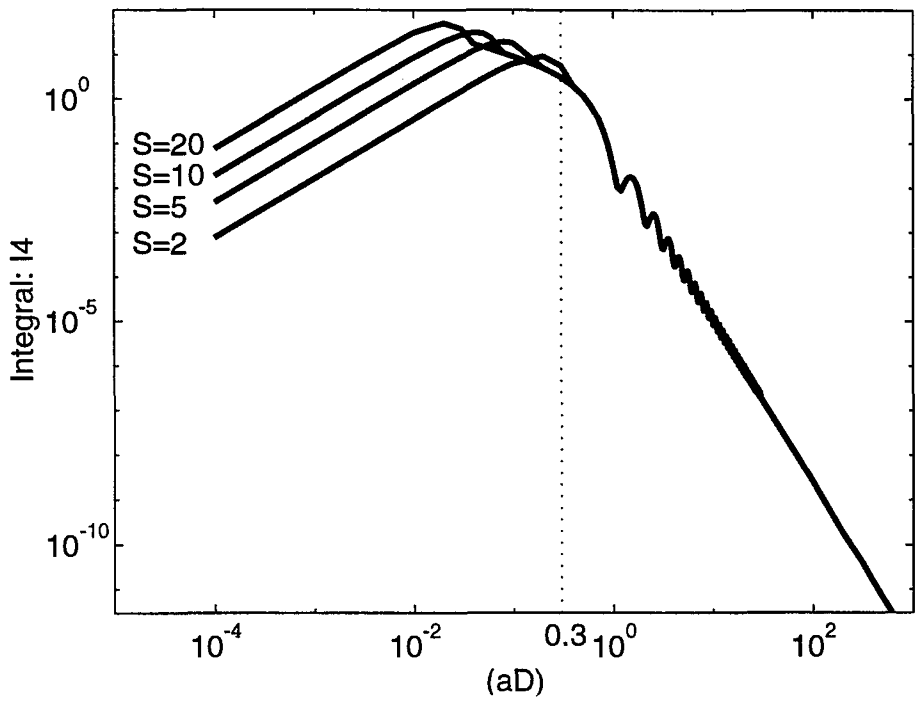



**Figure 3.4**: Graph of the integral $I_4$ against $(aD)$, used to calculate the temporal power spectrum of the differential centroid motions of pairs of circular apertures with different separations (S) for Kolmogorov turbulence. This graph has been altered since submission.

## 3.7 Performance of the Shack-Hartmann Sensor

### 3.7.1 Performance Limitations

Like any other wavefront sensor, the performance of the Shack-Hartmann sensor will depend on its application. It is normally used in two situations where the considerations are quite different, *closed-loop*, as part of an adaptive compensation system, where only the change since the previous measurement will be observed with the sensor, and *open-loop*, where the effects of the turbulence itself are measured. In the specific case of Adaptive Optics, the Strehl Ratio of the closed-loop image (assuming a perfect system elsewhere) is the most popular method of measuring performance and, if the effects of all other components in the AO system are neglected, this is directly related to the rms phase difference between the aberrated wavefront and the estimated wavefront used to provide correction. This estimate will only be accurate to some limited degree, due to both the fundamental limitations of the measurement technique and also to noise propagation in the system. The specific case for this thesis is open-loop turbulence statistics measurement, so only this aspect will be considered in detail here. The problem is that it is necessary to estimate the statistics which themselves are time-varying; this must be done using a finite number of photons and therefore a finite sampling rate and a finite subaperture size.

The error on the centroid statistics is caused by two other main sources of measurement error:

- Errors on the centroid determination, and
- Error due to finite measurement sample size,

and are propagated through the various algorithms to the end result.

The three major causes of centroid errors are photon noise, detector read noise and speckle noise. Photon noise is unavoidable in any system, where there is inevitably a finite number of frames collected, and is due to the fundamental quantum nature of light. A centroid measurement must be made using the light collected within a characteristic correlation time period of the turbulence being measured. During this period, a finite number of photons will pass through a lenslet and the positions where as many as possible of these photons arrive at the focal plane of the lenslet must be measured in such a way as to give an accurate centroid measurement. The finite number of photons imposes a limit on the accuracy of the measurement.

Speckle noise refers to the result of an undesirable shape of image. If the lenslet size is set so as to be less than or equal to $r_0$, then the image is, by definition, diffraction limited, and a good quality lenslet should produce close to a circularly symmetric image in the shape of the characteristic Airy disk of the lens. This, most concentrated, distribution of the light is optimum for determining the centroid. If $r_0$ becomes less than the lens size, the image will begin to break up into speckles. The light will be spread non-uniformly across a larger area and the centroid will be more susceptible to photon noise and readout noise. Thresholding techniques also become more difficult due to the asymmetric shape of the image and the larger image area. This problem can be eliminated if the lenslet size is always less than the Fried parameter.

Detector technologies available for the centroid measurements vary in quantum efficiency (*i.e.* what proportion of the incident photons contribute to the detected signal) and noise characteristics. Each detector element (or pixel) occupies a particular area of the detecting plane and a certain proportion of the photons incident on this area (depending on the quantum efficiency) within an exposure time result in a corresponding signal. In addition to this signal, a certain extra random value (whose standard deviation is referred to as the *readout noise*) will be added by the electronics which reads the signal from the detector. In general, this noise increases as the readout rate increases. The preferred detector should have a high quantum efficiency and a low readout noise.

If the subimage formed is circularly symmetric, it is possible to measure the position of the centroid with a 2x2 array of square detectors (*i.e.* a *quad cell*). This arrangement requires each sub-image to be centred between the four pixels. Ideally, this would consist of four Avalanche-Photo-Diodes (APDs), whose readout noise is practically zero, so that the wavefront can be sensed at very low light levels. However, APDs are extremely expensive. There is also a *registration* problem, in that the centre of each quad cell has to be aligned perfectly, on the axis of its associated lenslet. An AO system would certainly require use of faint reference stars but for turbulence measurements there are normally a number of bright stars available (usually the brightest star in the part of the sky being monitored is chosen) so it is most cost efficient to use an array, such as a CCD array, to detect the spot pattern. In any case, in a closed-loop system, the wavefront being measured is one that has been partially corrected by the system and there is relatively little motion of the image; however, for open-loop turbulence measurements, the image wander is comparable in magnitude to the Airy disc radius of the lenslet (see Eq. 3.44). Therefore, a larger array of elements (than 2x2) is required to locate a centroid position in this case.

Séchaud *et al* [48] and Rousset [49] have provided an in-depth analysis of the error propagation in a Shack-Hartmann sensor.

### 3.7.2 Design of a Shack-Hartmann Sensor with a CCD Array

As already discussed, in noiseless conditions with a closed-loop system, a detector area of 2 by 2 pixels, with an image diameter (that of the Airy disc of the lenslets under normal seeing conditions) considerably smaller than this would be sufficient to obtain a perfect measurement of the image centroid. When noise is added to the system, an increase in the size of the spot relative to the pixel size (and therefore also a larger number of pixels used in the centroid determination) is necessary to obtain the optimum centroid accuracy. There is a trade-off, however, since a larger spot size reduces the signal-to-noise ratio in each pixel (and consequently increases the number of pixels which must be read per unit time). Some sort of thresholding is essential to reduce the number of noisy pixels. Winick [50] has shown that when photon noise is considered independently, the most efficient value for the pixel to image-spot-size ratio (where the size is defined as the full-width-half-maximum) is between 2 and 2.5 (a result that depends little on the actual magnitude of the noise). Once detector noise has been taken into account, this ratio must be increased to between 3 and 5 [49]. An even bigger area than this must be designated for each subaperture image to cope with the image wander produced by the turbulence (although thresholding will allow 'empty' pixels to be ignored in the centroid calculation).

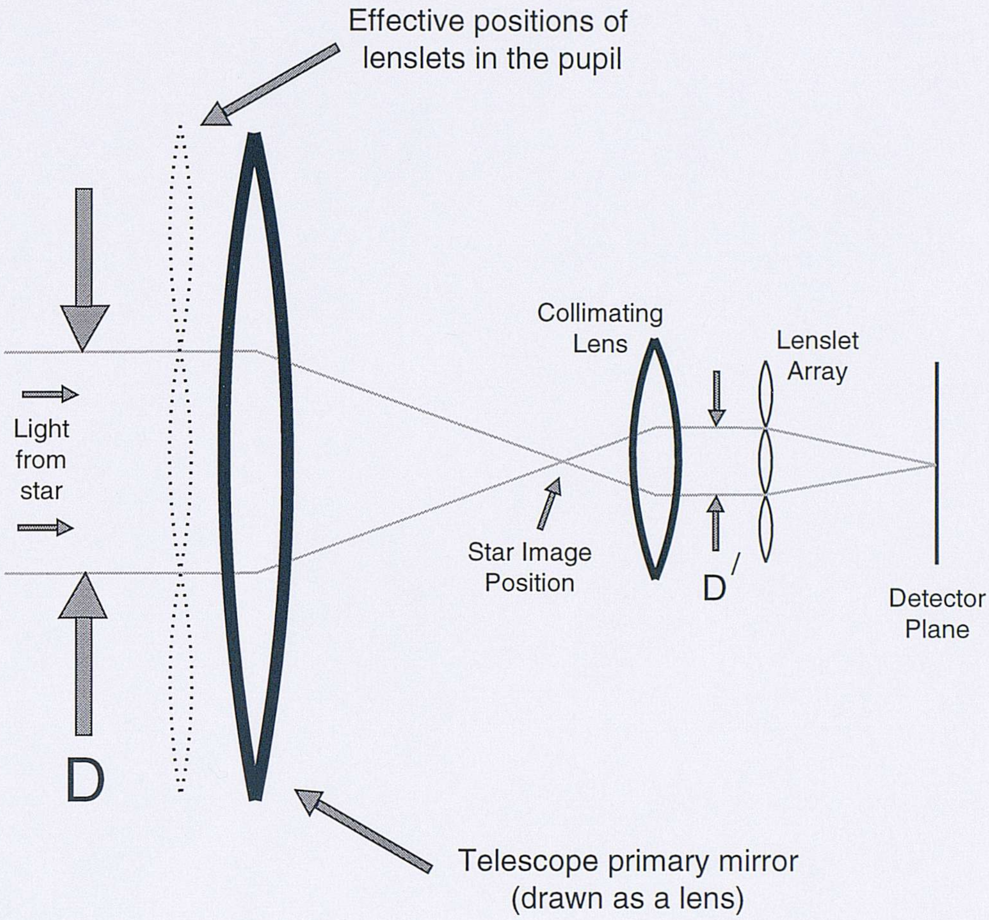



**Figure 3.5**: Diagram showing how the telescope pupil is reimaged to a size which is appropriate for the pixel size of the detector. The rays drawn describe the formation of a single subaperture image.

Once a suitable detector has been chosen, the procedure for designing the geometry of the rest of the system is fairly straightforward. Initially, certain parameters must be agreed before this can be done. These are the mean wavelength of the light to be detected; the expected minimum correlation time of the turbulence; the expected minimum correlation length ($r_0$) of the turbulence and the faintest magnitude of reference stars to be used. The lenslet array will, in general, be many times smaller than the actual telescope pupil, due to the small sizes of the detectors used. A re-collimation of the light beam, as in Fig. 3.5, is normally employed to deal with this. The ratio between the size of the sub-aperture at the telescope pupil, $D$, and the physical lenslet size, $D'$, is denoted by $S_0$, and the focal length of the lenslets is defined as $F$. Appropriate values for $D$, $D'$ and $F$ can be chosen if the typical $r_0$ and the size of the detector pixels are known.

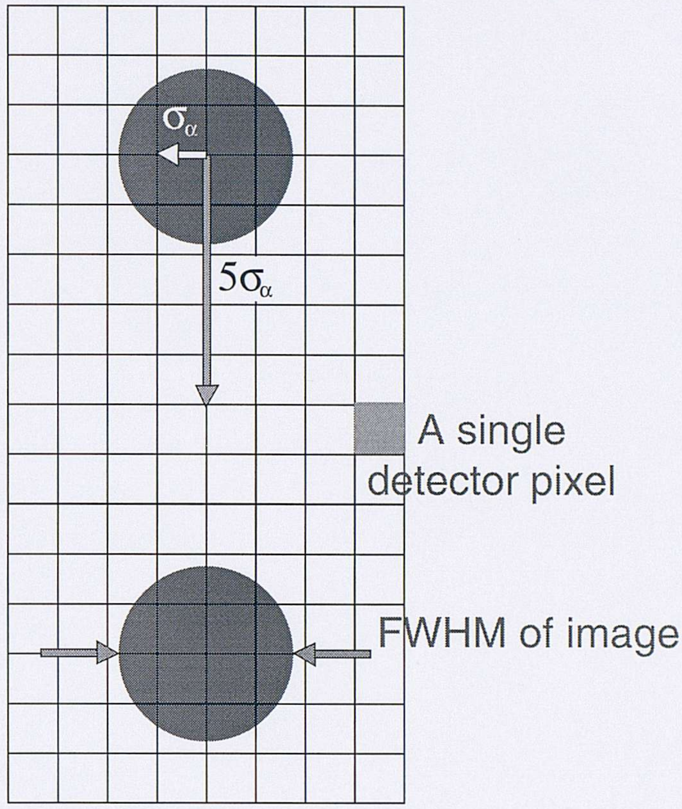



**Figure 3.6**: Diagram showing the array of pixels which make up a CCD detector and suitable separations for neighbouring Shack-Hartmann images.

Firstly, the image diameter on the detector is approximately $1.22\lambda F S_0/D$, and secondly, from Eq. 3.44, the standard deviation, $\sigma_{c_x}$, of the centroid motions is $0.40\lambda F S_0/D$ for square lenslets (ray angles scale up by a factor of $S_0$ between the telescope pupil and the re-imaged pupil at the lenslet plane). The moving image from each lenslet must occupy its own region on the detector (see Fig. 3.6). This region must be as small as possible so that the minimum number of pixels is read in each frame, but large enough so that the probability that neighbouring images collide is suitably remote. If the radius of the region is chosen to be, say, $5\sigma_{c_x}$, then the separation between the images should be around 10 pixels [49] (or a little larger, if the images are deemed to be larger than their full-width-half-maxima; the radius of the image can be defined as that of the region within which the 'useful' signal for centroiding occurs; this in turn depends on the ratio of the detector read noise to the peak signal in the image). The radius of this region defines the physical radius of the lenslets, $D'$, and thus the scaling factor $S_0$ is also deducible if $D$ is required to be equal to the expected value of $r_0$.

The image diameter, $1.22\lambda F S_0/D$, must be set equal to between 3 and 5 pixel widths; for the sake of an example this will be chosen to be 3.5 pixels here. The lenslet focal length $F$ can thus also be inferred. Typical values are pixel size $= 0.027\mu m$, $F = 45mm$ and $D' = 0.5mm$, for $r_0 = 100mm$.

### 3.7.3 Determination of Centroid Accuracy

The number of photons detected in each subimage can be estimated by considering the geometry of the system, the exposure time, the magnitude of the reference star and the quantum efficiency of the detector. From this it is possible to estimate the accuracy of the centroid measurement from the photon noise and detector noise considerations. The centroid errors can then be converted to errors on the statistical quantities of interest using the appropriate relations.

The centroid coordinates can be estimated from the readout of a CCD array by considering the intensity signal values $I_i$ at pixels numbered from $i = 1$ to $n$ over the region of interest. If the $i$th pixel has coordinates $(x_i, y_i)$, then the centre of gravity, or centroid, $(c_x, c_y)$ is estimated via the following two equations:

$$c_x = \frac{\sum_{i=1}^{n} x_i I_i}{\sum_{i=1}^{n} I_i} \tag{3.68}$$

and

$$c_y = \frac{\sum_{i=1}^{n} y_i I_i}{\sum_{i=1}^{n} I_i} \tag{3.69}$$

As already mentioned, there are many different forms of noise. However, each source of noise can be characterised in terms of its standard deviation $\sigma_i$ at each pixel ($i$). As long as the noise levels from pixel to pixel are uncorrelated, and if the noise levels $\sigma_i$ are known or can be estimated, the variance on the centroid is given as:

$$\sigma_{c_x}^2 = \frac{1}{n_{ph}^2} \sum \sigma_i^2 {x_i}^2, \tag{3.70}$$

where $x_i$ is now the distance from the centre of the $i$th pixel to the estimated centroid position.

Photon noise is a signal dependent noise which is inherent in the quantum nature of the light being collected. The variance of the photon events at each pixel can be considered to be equal to the intensity at that pixel, *i.e.* $\sigma_i^2 = \langle I_i \rangle$, using Poisson statistics. Then

$${\sigma_{c_x}}^2 = \frac{1}{n_{ph}^2} \sum I_i {x_i}^2, \tag{3.71}$$

The standard deviation of the distribution of photons in the image (in angular coordinates) can be considered to be equal to the half-width of its Airy disc (in the diffraction limited case), so that $\sum I_i {x_i}^2 = n_{ph} \left(\frac{\lambda}{2D'}\right)^2$. With $n_{ph}$ detections across the entire region of n pixels, the centroid error (in angular coordinates) is [49]

$$\sigma_{c_x}^2 \approx \frac{1}{n_{ph}} \left(\frac{\lambda}{2D'}\right)^2. \tag{3.72}$$

In the case of detector readout, the noise can be considered to affect each pixel equally, no matter how high the signal level is. Assuming the noise level to be the same for each pixel so that ${\sigma_i}^2 = {\sigma_e}^2$ per pixel per exposure, then for read noise

$$\sigma_{c_x}^2 = \frac{1}{n_{ph}^2} \sum_{i=1}^{n} {\sigma_e}^2 {x_i}^2. \tag{3.73}$$

This can be calculated directly if the region over which the centroiding is performed is known. For example, if this region is $N$ pixels square in size with the pixel in column $i$ and row $j$ being a distance $x_{i,j}$ from the centroid estimate, then in pixel coordinates:

$$\sigma_{c_x}^2 = \frac{1}{n_{ph}^2} \sum_{i=-N/2}^{N/2} \sum_{j=-N/2}^{N/2} \sigma_e^2 {x_{i,j}}^2. \tag{3.74}$$

Therefore,

$$\sigma_{c_x}^2 \approx \frac{1}{n_{ph}^2} \frac{N^4}{12} \sigma_e^2. \tag{3.75}$$

The error is critically dependent on the number of pixels used for centroiding; therefore some sort of thresholding to reduce the number of pixels involved is essential. This can be done by ignoring any pixels where the noise is larger or comparable to the signal level (if one assumes the shape of the image is known or can be predicted). For example, one may choose to use only those pixels within a circular region which contains 80% of the starlight falling on the detector. An algorithm which attempts to optimise the thresholding procedure in a centroiding process is described in section 4.2.3.

### 3.7.4 Error Calculation: A Case Study

The effect of photon and readout noise on the centroid accuracy has been studied for a real detector and wavefront sensor configuration. The detector considered has a performance typical of the Princeton Instruments CCD camera used in the seeing measurements which will be described in section 5.2.2. The readout noise $\sigma_e$ of the camera is taken to be 9 electrons (equivalent to 9 detected photons). A numerical calculation of the effects of noise on the accuracy of the centroid measurement using a typical simulated image has been performed using the mathematical language of *MATLAB*. A two-dimensional Gaussian distribution is first created over a 32 by 32 grid of points. The grid is then divided into smaller squares consisting of 4 by 4 grid points. The values within each of these squares are summed together to create a new 8 by 8 array which represents the signals found on an array of square pixels when a Gaussian image is formed from a Shack-Hartmann lenslet. The total signal across the region is normalised to unity. The final array is such that the full-width-half-maximum (FWHM) of the image is 3.5 pixels. This process is visualised in Fig. 3.7.

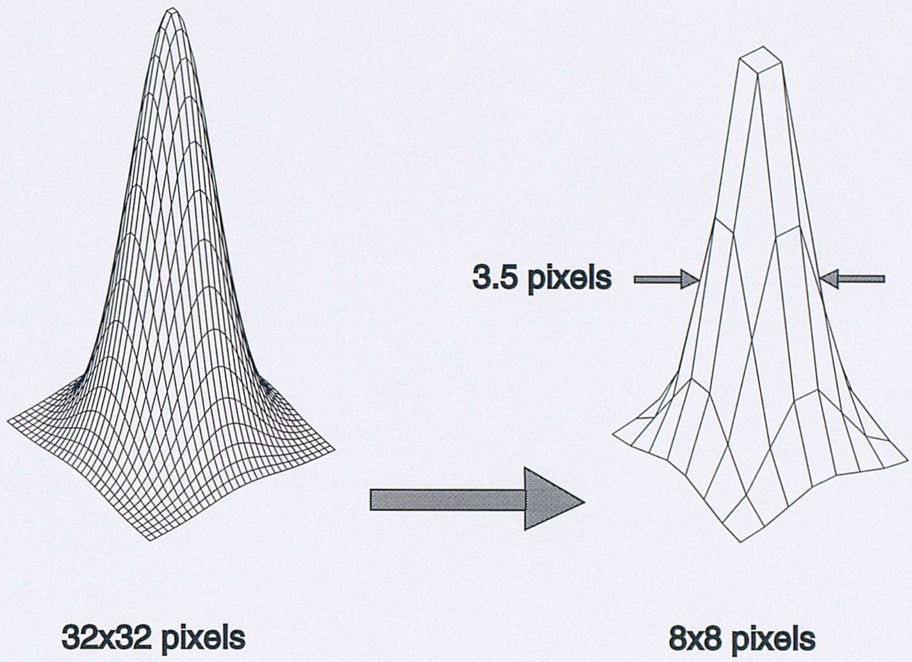



**Figure 3.7**: Diagram showing the process involved in simulating a Gaussian image formed on a discrete array of detector elements.

Equations 3.71 and 3.73 can now be applied numerically to the array of signal values now created. Different light levels are simulated by multiplying the entire array by $n_p$, the total number of photons in the image. The effects of the detector quantum efficiency are not included, but these can be assessed by scaling the value of $n_p$ appropriately. A graph showing the centroid accuracy (relative to the image FWHM) against total detection level for a range of light levels for both the photon limited and read noise limited cases is displayed in Fig. 3.8. The dashed-dotted line is for a read noise calculation using all 64 pixels in the centroid estimation. In order to improve on this accuracy, a very crude thresholding process is carried out, where only those pixels whose centres lie within three pixel widths of the centroid are used in the calculation (this is equivalent to rejecting pixels with signal levels below 20% of the maximum). This calculation is represented by the dashed line. A more rigorous thresholding algorithm might be expected to increase the accuracy further, so that the magnitudes of the centroid errors caused by photon noise and by read noise would be comparable in size for this detector specification.

### 3.7.5 The Effect Of A Finite Number Of Measurement Samples

Values of $r_0$ are calculated using either a single aperture or pair of apertures (depending on the method chosen) from an angle-of-arrival variance (either single-aperture or differential) estimated using a finite number of sampled centroid measurements (equivalent to the number of Shack-Hartmann images, or frames, obtained). When using

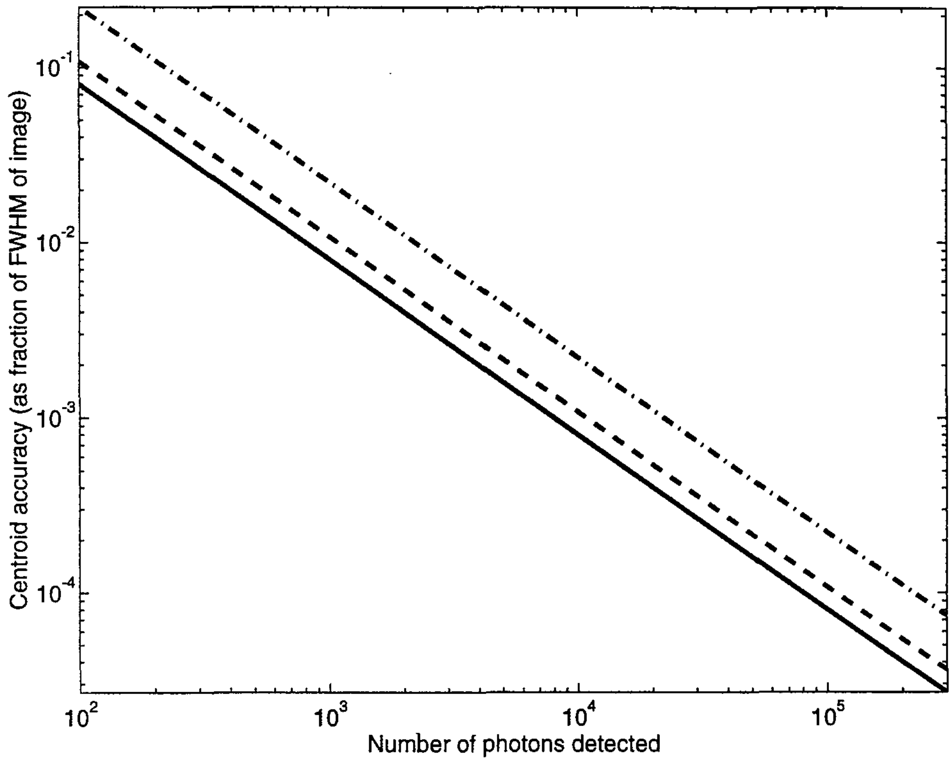



**Figure 3.8**: Graph showing the relative centroid accuracy as a function of the number of photons detected per subaperture image by the Princeton CCD camera. The solid line is for photon noise only. The dashed-dotted line is for read noise only. The dashed line is for read noise using a simple thresholding algorithm.

a Shack-Hartmann sensor to obtain time series of $r_0$ values, the number of images which can be obtained is limited either by the exposure rate of the detector or by the magnitude of the reference star. It is desirable to obtain $r_0$ values as regularly as possible, but this will be at the expense of a correspondingly small measurement sample size. The smaller the size of the sample, the greater is the expected inaccuracy of the estimated variance, from simple statistical considerations. Each variance value estimated (from a certain number of measurements) will have an associated error due to the finite sample size. This error can be calculated from the *probability distribution* of the variance estimate. The larger the number of measurements acquired, the narrower is this probability distribution, hence one is more likely to obtain an answer close to the 'true' value of variance. Since the time-averaged angle-of-arrival, $< \alpha >$, is zero, its variance can be calculated simply as the mean of the squares, $< \alpha^2 >$. The probability distribution of the mean of several independent measurements of $\alpha^2$ (angle-of-arrival squared) can be found by convolving the probability distributions of each of the individual measurements of $\alpha^2$ together [20] (a discussion of what is meant by 'independent' is given in section 4.3.2). This procedure is shown for two such mea-

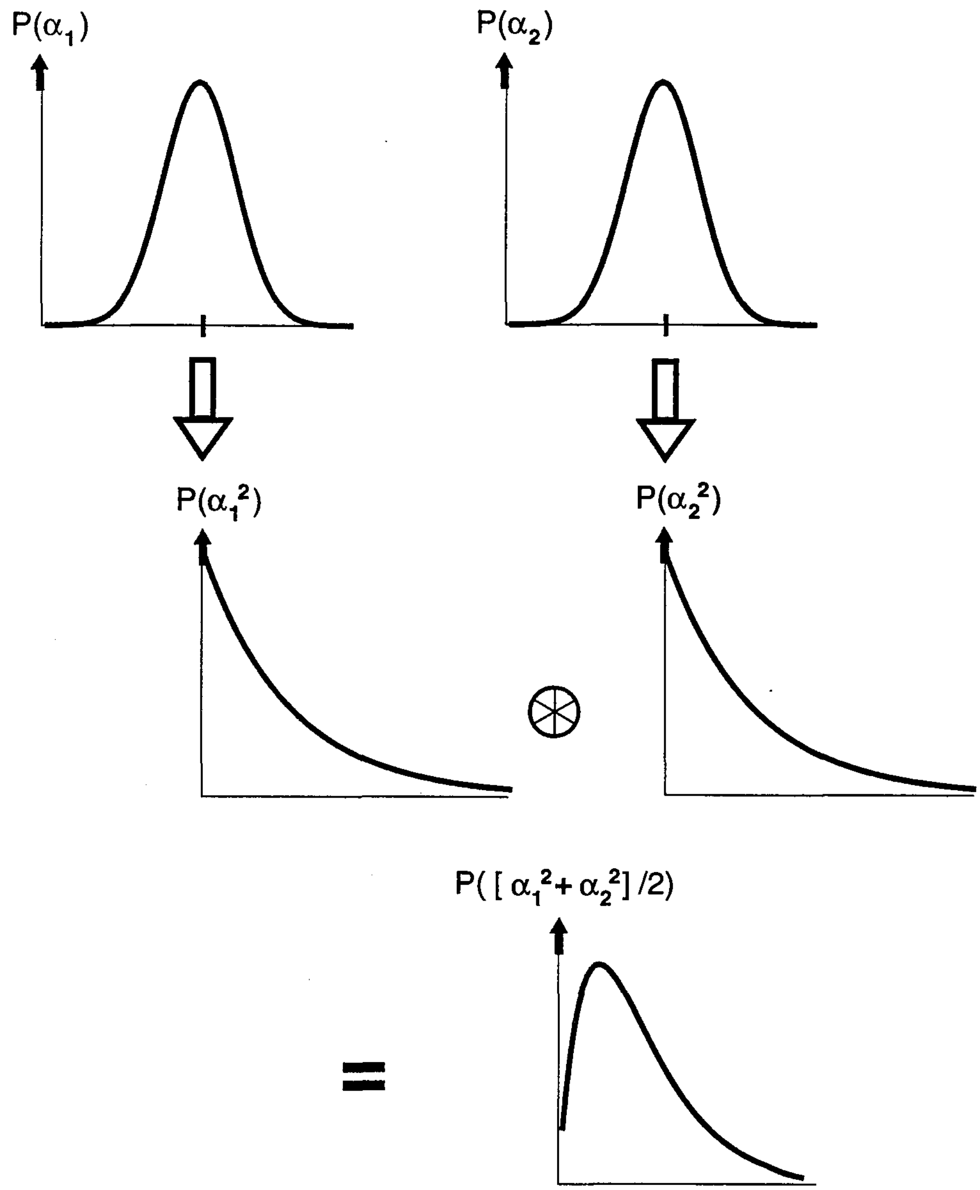



**Figure 3.9**: The procedure for estimating the probability distribution for the variance of a Gaussian variable calculated using a finite number of independent measurements. Top: probability distributions for two angle-of-arrival measurements ($\alpha_1$ and $\alpha_2$). Centre: Probability distributions of their squares. Bottom: Probability distribution of the mean of their squares, found by convolving the two together.

surements in Fig. 3.9. For moderately large numbers of measurements (*e.g.* larger than 30), the probability distribution of the variance estimate becomes close to Gaussian in shape (a consequence of the central limit theorem). If the (expected) angle-of-arrival variance is ${\sigma_\alpha}^2$, then an error on an estimate of ${\sigma_\alpha}^2$ is found to be ${\sigma_\alpha}^2\sqrt{2/n}$, where n is the number of measurements made (see Eq. A.4). So, for example, if 50 independent measurements of angle-of-arrival are used to estimate the variance, then its accuracy is 20%. However, if the $r_0$ values from a large number of subapertures or subaperture pairs can be averaged together, then this error can be reduced accordingly.

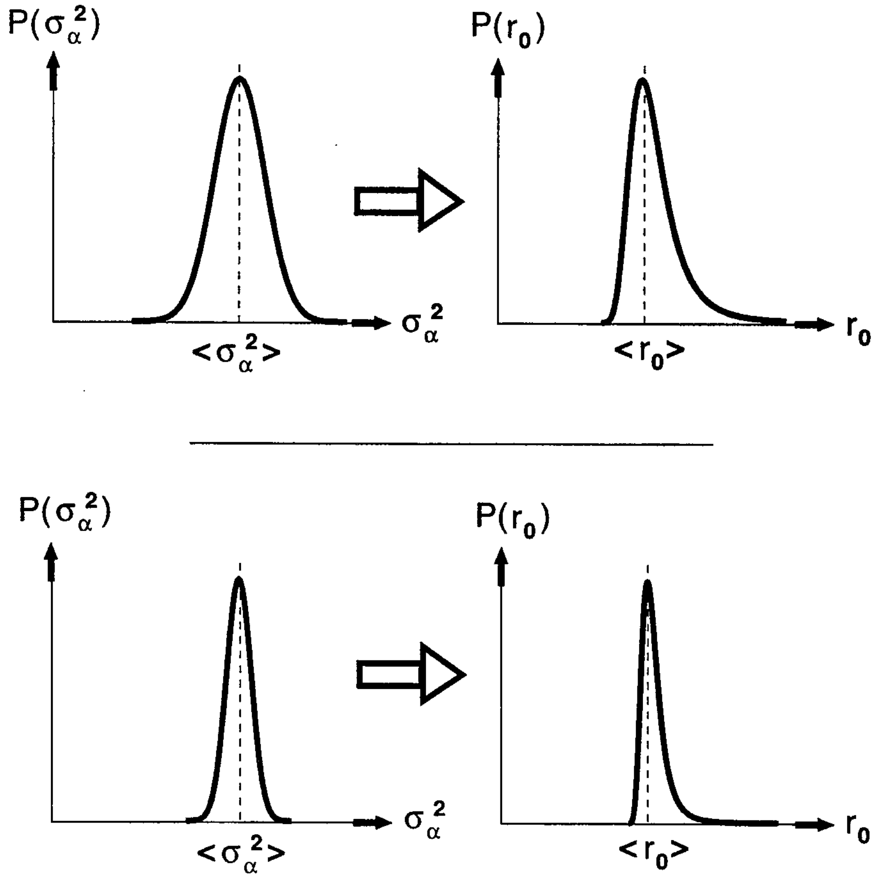



**Figure 3.10**: Diagram showing the reasons why a measurement of a centroid variance leads to a bias in the value of $r_0$. Top: probability distribution of an estimate of centroid variance taken from $N_1$ independent measurements and the resulting distribution of $r_0$ measurements. Bottom: the same distributions for estimates taken from $N_2$ independent measurements, where $N_2 > N_1$.

Of greater importance, however, is any biasing effect that might occur when too small a sample size is used. Although, as already mentioned, the probability distribution of the estimated variance is close to Gaussian in shape and therefore symmetric, the corresponding $r_0$ value is calculated by raising the variance to the power of $-3/5$. As shown in Fig. 3.10, The probability distribution of an $r_0$ value estimated in this way is likely to be asymmetric, in which case $r_0$ estimates will be biased high. However, the width of this distribution is decreased if a larger sample of measurements is used. This also reduces any biasing effect.

An attempt has been made to estimate this bias by using a simple algorithm written in the mathematical computer language of *MATLAB*. This has been done for the $r_0$ calculation which uses the single aperture variance (although the principle is the same for the differential variance). Assuming that the angle-of-arrival values obey Gaussian statistics (this is not strictly true due to aperture averaging effects), the probability distribution of the angle-of-arrival squared of a single measurement can be simulated

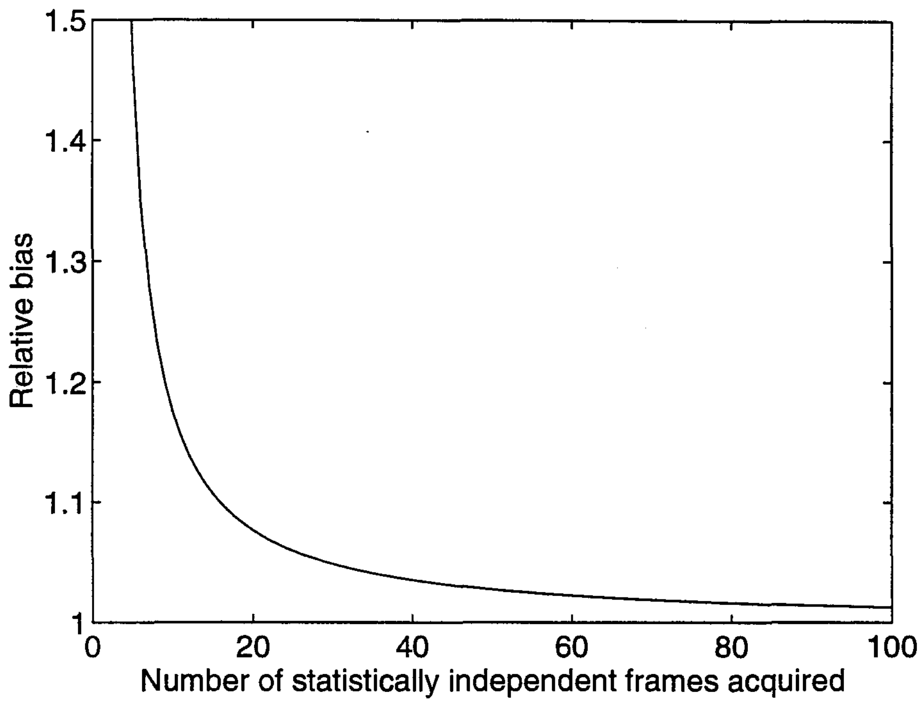



**Figure 3.11**: Expected bias on $r_0$ through measurement using a finite number of frames.

in the form of a discretised normal distribution. A convolution is then performed to calculate the probability distribution of an estimate of angle-of-arrival variance found from a certain number of angle-of-arrival measurements. Finally, the probability distribution arising from taking the variance raised to the power of $-3/5$ is calculated. This distribution shows the spread of $r_0$ values that would be obtained from measurement samples of the appropriate size. Its mean is the most likely value of $r_0$ which would be obtained when the $r_0$ estimates from several subaperture pairs are averaged together. Its median (*i.e.* the 'peak' of the distribution) is where the 'true' value of $r_0$ lies. In general, these two values will not coincide due to the fact that the distribution is skewed (see Fig. 3.10). The bias in $r_0$ can be estimated since the true value is known from the initial probability distribution. This has been done for different sample sizes up to 100 and the result is displayed in Fig. 3.11. The error is approximately 3% at a sample size of 50 frames. When calculating time series of $r_0$ values there will, therefore, always be a compromise between the need to measure $r_0$ as regularly as possible and the need to obtain an unbiased estimate.

## 3.8 Wavefront Curvature Sensor

Beckers [51] shows how it it possible to relate the second derivative of the wavefront phase to the intensity distribution in an out-of-focus image. In very simple terms, a curvature introduced in the wavefront at the entrance pupil will result in an increase

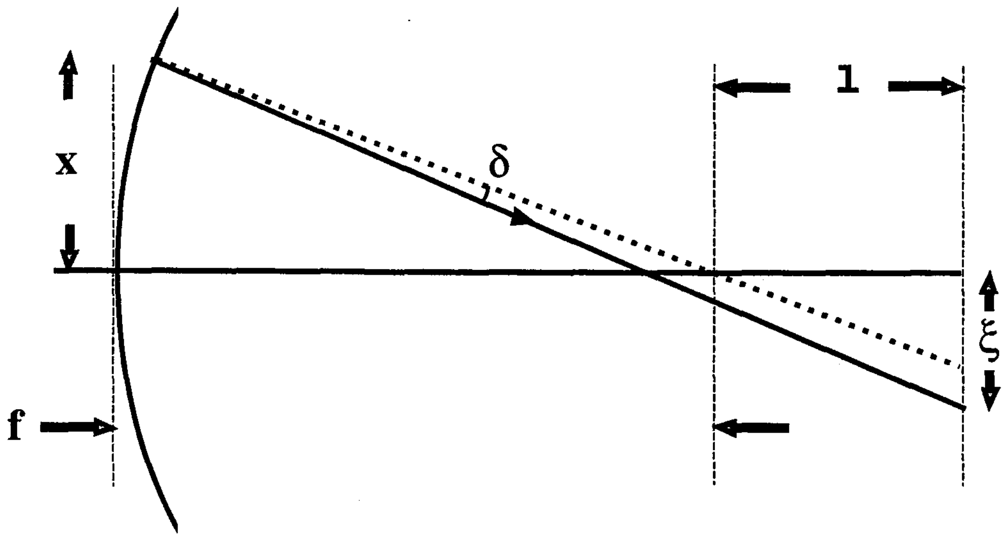



**Figure 3.12**: Ray optics representation of a single defocused image.

in intensity somewhere in a plane in front of focus and a decrease in intensity in a plane behind focus (or vice-versa, depending on the direction of the curvature). Using a ray optics approximation, if the wavefront first derivative is distorted by a value $\delta(x)$ then the intersection of the ray with a plane at a distance $l$ from focus will occur at coordinate $\xi = xl/f + \delta(x)(F+l)$ (see Fig. 3.12). The intensity in the image is proportional to $d\xi/dx = |l/f + (F+l)d\delta(x)/dx|$; therefore intensity variations in the plane are proportional to the wavefront curvature.

Clearly, this simplified model does not take into account scintillation or diffraction effects (particularly at the edge of the pupil). Roddier has evaluated the effects of diffraction at the pupil edge. This is done in so-called *object space*, so that any imaging optics is removed, and the telescope mirror can be assumed to have been replaced by a circular aperture. Again in the paraxial approximation the *irradiance transport equation* [52] [53] can be applied to the intensity distribution $I$ and the phase distribution $\varphi$ of a wave propagating along the z-axis,

$$\frac{\partial I}{\partial z} = -\left(\nabla I \cdot \nabla\varphi + I\nabla^2\varphi\right). \tag{3.76}$$

In the case of no scintillation, $\nabla I$ is zero everywhere except at the very edge of the pupil where

$$\nabla I = -I_0 \mathbf{n} \delta_c, \tag{3.77}$$

$\delta_c$ being a dirac delta function along the pupil edge and $\mathbf{n}$ a unit vector perpendicular to and in the plane of the pupil edge. Equations 3.76 and 3.77 lead to

$$\frac{\partial I}{\partial z} = I_0 \frac{\partial \varphi}{\partial n} \delta_c - I_0 P \nabla^2 \varphi. \tag{3.78}$$

Here P is the pupil function (*ie* 1 inside and 0 outside). The intensity distributions $I_+$ and $I_-$ in two planes at a distance $\Delta z$ either side of the pupil can then be calculated (as long as $\Delta z$ is small enough that phase gradients do not become important):

$$I_\pm = I_0 \pm \frac{\partial I}{\partial z} \Delta z. \tag{3.79}$$

A signal $S$ (this is the 'sensor signal', to be measured by the curvature sensor) is defined as

$$S = \frac{I_+ - I_-}{I_+ + I_-}. \tag{3.80}$$

Therefore,

$$S = \left( \frac{\partial \varphi}{\partial n} \delta_c - P \nabla^2 \varphi \right) \Delta z. \tag{3.81}$$

When the focussing effects of the telescope imaging optics are taken into consideration, it is found that the two planes either side of the pupil in object space are equivalent to two planes at a distance $l$ either side of the telescope focal plane in *image space.* The effect of the imaging optics is to 'shrink' the space following the mirror so that the plane which was at infinity in object space is now at the focal plane (in image space). A coordinate transformation takes place so that $\Delta z = f(f - l)/l$ and the coordinates in the two planes are multiplied by a factor $l/f$, where $f$ is the focal length of the telescope (the coordinates in the second plane are also reversed). Hence,

$$S = \frac{f\,(f - l)}{l} \left( \frac{\partial \varphi}{\partial n} \delta_c - P \nabla^2 \varphi \right). \tag{3.82}$$

Roddier's more complete analysis [54] also takes into account intensity variations within the pupil and diffraction effects and gives the same result as Eq. 3.82. This result has led Roddier to the idea of the Curvature Sensor [55] [56] [52]. The sensor requires two images to be recorded simultaneously at an equal distance either side of the focal plane. The normalised difference between their intensities is a measure of the second derivative (or curvature) of the wavefront phase. With reference to the geometry represented in Fig. 3.13, the two images are positioned in planes $P_1$ and $P_2$ respectively. The distorted wavefront is focussed by lens or mirror $L_1$ of focal length $f$. $L_2$ is a lens which re-images $L_1$ at a distance $f$ behind focus. This is used to make the image space symmetrical either side of focus, and thereby retain the equivalence with the two planes in object space; it may or may not be necessary in practice. A complete analysis

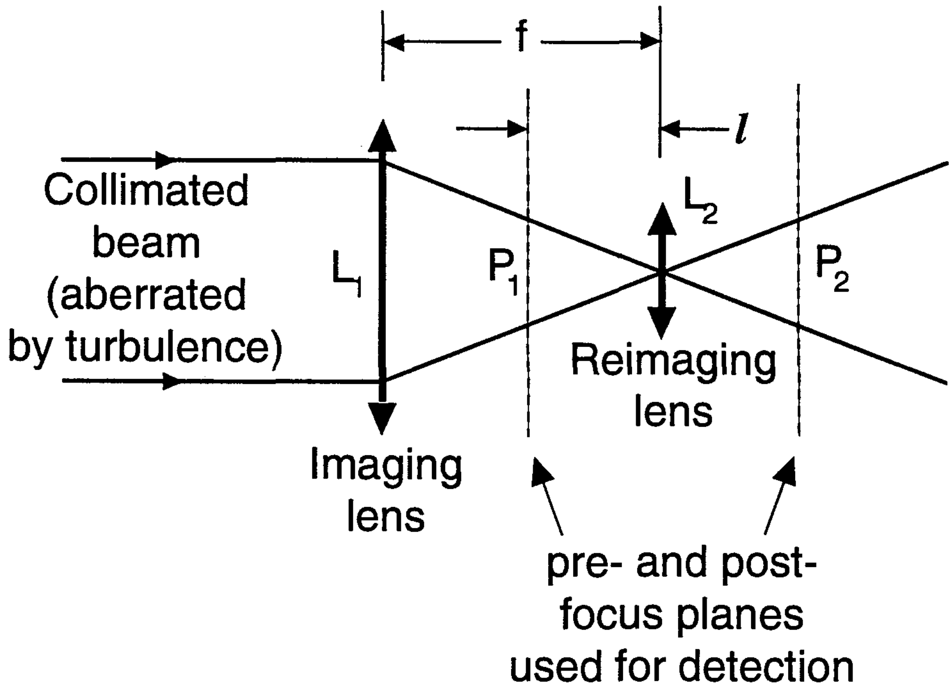



**Figure 3.13**: Principle of the curvature sensor.

[54] gives an expression for the normalised difference in illuminations $I_1$ and $I_2$ in planes $P_1$ and $P_2$, under the assumption that $r_0$ is sufficiently large for geometric optics approximations to apply (*i.e.* the phase gradients in the pupil are not sufficiently large for there to be any significant spreading of the beam). The complete wavefront surface can be reconstructed by solving Poisson's equation with the appropriate Laplacian measurements and boundary conditions provided by the sensor.

In fact, in the case of little scintillation, Hickson and Burley [57] have shown that the wavefront can be solved satisfactorily from a single defocused image, thus further reducing the number of measurements required by a factor of two.

## 3.9 Curvature Sensing vs Shack-Hartmann: A Comparison

The Shack-Hartmann sensor estimates the wavefront first derivative over an area of the pupil known as a subaperture; in order to make a comparison an equivalent area must be defined for the curvature sensor. A subaperture in the curvature sensor is thus defined as the area in the pupil over which a single estimate of the wavefront curvature is made. It is immediately obvious that this requires only two detector elements (one before and one behind focus), whereas the Shack-Hartmann requires at least four elements (or pixels) to make a single measurement. Therefore, the photon throughput required for wavefront sensing in a curvature sensor is at most half of that required in a Shack-Hartmann system.

A second advantage of the curvature sensor is that potentially any geometry of detector elements may be used to measure the signal. Thus, in an adaptive optics system, for example, the arrangement of the subapertures, with respect to the telescope pupil, can be designed to match the geometry of the correcting elements in the most efficient way.

Perhaps the biggest advantage, however, is that the sensitivity of the curvature sensor (*i.e.* the ratio between $(I_+ - I_-)/(I_+ + I_-)$, the sensor signal, and the noise level) can be altered by changing the positions of the defocus planes relative to focus (*i.e.* by increasing or decreasing $l$). The maximum sensitivity is achieved when the size of the blur (caused by the finite size of the image in the focal plane) is equal to the subaperture size scaled down to the size of the beam cross-section by a factor of $(l/f)^2$ [55] [56]. If the blur size is larger than this, it will be spread across several subapertures; conversely, if $l$ is increased too far, then the sensor signal is reduced in accordance with Eq. 3.82. In open loop operation, the blur size is equal to the seeing disk size, so this alteration to the system geometry should be made whenever $r_0$ changes. In a Shack-Hartmann system, closing the loop does not affect the size of the image (assuming $d \leq r_0$). However, in the closed-loop curvature sensor, the image blur is reduced to close to an Airy disk in size, and the sensitivity can be increased by a factor of $(D/r_0)$, by increasing $l$ accordingly. Roddier [58] has shown that due to the error propagation in the reconstruction procedure, this advantage holds only for small numbers of subapertures. This is because the error due to reconstruction from $n$ subapertures grows proportionally to $n$ for a curvature sensor, but proportionally to $A + B \ln n$ [59] (where $A$ and $B$ depend on the subaperture geometry) in a Shack-Hartmann system. For large arrays, the Shack-Hartmann sensor is to be preferred.

For the purpose of open-loop turbulence measurements, the Shack-Hartmann remains an effective sensor, especially since the abundance of bright stars imposes no strict requirement on the sensitivity of the sensor. Its advantage is in the ease of its construction. The curvature sensor requires two separate images (before and after focus) to be detected. This can be done either simultaneously, or with the use of a dynamic device, such as a membrane mirror [60], which produces pre- and post-focus images alternately on the same detector. In either case the system geometry is more complicated than that of a Shack-Hartmann sensor, which can be constructed from common optical components.

Roddier has, however, described a seeing monitor based on a curvature sensor [61]. It is claimed that a single curvature sensor measurement is equivalent to a differential image measurement with a Shack-Hartmann system; therefore a full circular pupil can be used instead of two holes. The system employs a Photomultiplier tube instead of a CCD array, making the whole system less expensive to construct. Two focus positions are created in the system, the two focussed beams being in mutually orthogonal planes of polarisation, using a birefringent lens. The two defocused images are detected on the photomultiplier alternately using a polarisation-sensitive waveplate, which can be switched at rates of up to 1000Hz.

# Chapter 4

# The Shack-Hartmann Wavefront Sensor: Algorithms for Measuring Turbulence

## 4.1 Seeing Measurements Relevant to Adaptive Optics

The purpose of this chapter is to describe how Shack-Hartmann images can be analysed to obtain the statistics of the turbulence-degraded wavefront. It is assumed that a time series of images has been obtained, with each frame containing a pattern of subaperture images, each of which corresponds to one subaperture in the pupil. The precise experimental details with which this has been done are not important for now (a number of experimental configurations will be discussed in Chapter 5). It is assumed that these frames were acquired at regular intervals using a suitable reference star.

A common method of performing simple seeing measurements is to use a differential image motion monitor (or DIMM; see section 3.6.4) [45] which is independent of the telescope and its artificial surroundings. This would normally provide a continuous series of $r_0$ measurements, perhaps even simultaneously with astronomical observations at a site, and would give an indication of the best possible observing conditions at that site. A DIMM could also be used to provide real-time seeing information which could be used to optimise an adaptive optics system. The Shack-Hartmann wavefront sensor automatically provides a large number of pairs of subapertures, all of which act simultaneously as DIMMs, and thus gives a greater measurement capability. However, this must be used in conjunction with existing telescope optics, consuming valuable observation time. Such experiments are also susceptible to dome seeing effects, which often cannot be distinguished from the effects of turbulence.

Using this sensor, the important measurements with regard to the construction of adaptive optics systems are:

- The testing of the major results of the Kolmogorov theory and the Taylor hypothesis of frozen turbulence.
- Measurement of the Fried parameter, $r_0$.
- Measurement of the Greenwood frequency, $\nu_g$.

It is important to note the variation in the statistical quantities between different times of the year, over different times of the night and the variation during the period of a typical long-exposure. Measurement of these two parameters is essential because, in the Kolmogorov approximation, they lead to a complete description of the statistical behaviour of the wavefront. It allows a complete modelling of a particular proposed adaptive optics system to be made and its performance to be predicted. This can be done, for example, in terms of Zernike polynomials, whose statistical behaviour can be entirely inferred from parameters such as these. However, a Zernike description of the wavefront, or its statistics, implies a knowledge of the behaviour of a large number of coefficients; verification of either the Kolmogorov model or the Taylor approximation would allow a drastic reduction in the number of parameters involved. It is also desirable to measure the turbulence profile ($C_n^2(h)$; relevant to adaptive compensation with a laser beacon), angular isoplanatism (which is a function of the turbulence profile), and the wind vectors of layers of turbulence. However, these measurements are not possible with a simple Shack-Hartmann wavefront sensor.

Of course, some of the parameters involved (such as $r_0$) have no meaning if the statistics of seeing are not Kolmogorov. Chapter 6 discusses an approach to characterising the statistics of non-Kolmogorov turbulence, so that these parameters can be used in a more generalised way. The particular aspects of the theory which can be tested are:

- Stationarity, *i.e.* the invariance of the correlation function of the wavefront from one part of the aperture to another; this may be observed by measuring the variation of $r_0$ from one subaperture pair to another, for example.
- Outer scale effects: if the outer scale is small enough to affect the structure function at separations smaller than the telescope diameter, the turbulent energy at these scales will be less than expected, and adaptive compensation need not be so great; again, this can be observed in an apparent increase in $r_0$ as the subaperture separation is increased.
- The accuracy of the (spatial) Kolmogorov power laws Eq. 2.62, tested by observing the differential variance of the centroid motions (see Chapter 5). This

information can also be inferred from the statistics of a modal decomposition of wavefronts obtained from individual Shack-Hartmann patterns; further discussion is given elsewhere [62].

- The power spectral slopes for single centroid and differential centroid time series predicted from the Taylor hypothesis.

## 4.2 Determination of Centroid Positions

### 4.2.1 Short Exposures as a Method of Sampling

The first task in any seeing experiment is to obtain a time series of centroid locations relative to their respective reference positions for all the subapertures formed by the Shack-Hartmann lenslets. For this purpose, two types of Shack-Hartmann image (*i.e.* the intensity signal profile measured over the surface of the detector in the focal plane of the lenslets) are required. Firstly, a series of short-exposure images, or *frames*, which together constitute a single *run*, whose exposure time is considerably less than the wavefront correlation time, in order to take statistical snapshots of the atmosphere. Secondly, a long-exposure image, from which the zero-point or *reference* positions of the centroids are found; this is often taken as a sum of all the short-exposures in one run.

The precise method of light detection varies depending on the type of detector, two of which are used in the seeing measurements described in this thesis. The first type is a frame transfer system, where an image is formed on the detector over a short period, then the entire frame is transferred elsewhere ready for reading out at the same time. An example of this type of system is the DALSA CCD camera used in section 5.4. The second is a line transfer system, where each line is transferred in turn for reading out. The disadvantage of this choice of system is that each line will be exposed to the star during a slightly different time period, resulting in some 'smearing' of the image (*i.e.* each subaperture image is effectively captured at a different instant in time). The IRCAM3/ALICE system used at UKIRT is an example of this. The Princeton Camera used at the JKT is also a line transfer system. However, it is also mechanically shuttered, so that the entire frame is read while the chip is unexposed. This avoids the effects of smearing.

The analysis of such an image can be divided into two parts; firstly, the frame-by-frame centroiding of each of the subaperture images, and, secondly, the statistical analysis of the evolution of each of these centroids.

### 4.2.2 Cleaning Images

The raw Shack-Hartmann images must undergo some processing before the centroiding can take place. Firstly, a background image must be subtracted. This background is normally an image of a dark part of the sky, so that any biasing effects from both the telescope and the camera can be taken into account. Secondly, any effects from badly behaved or burnt out pixels must be removed. These pixels can normally be identified by simply looking at a background image. This is done on each frame by averaging bad pixels with their nine nearest neighbours. In all cases described in this thesis, a linear relationship between pixel value and incident intensity is assumed, disregarding any need for a calibration of the camera to be made.

### 4.2.3 Centroiding Algorithm

The centroiding algorithm used was written by N.J. Wooder of the Applied Optics Group at Imperial College. Since it is of vital interest to the Shack-Hartmann data reduction it will be described in detail here. Its purpose is to locate a centroid coordinate corresponding to each subaperture in every frame. A modular diagram showing the major parts of this algorithm can be seen in Fig. 4.1. A number of parameters must be set which depend fairly critically on the design of the system, the size of the spots on the camera and the level of noise. These parameters will be shown in italics in the following description.

The centroiding process is essentially a three-stage process. The first two stages are operated on a long-exposure image, normally an average of a large number of frames. These stages identify the reference, or non-aberrated, positions.

The first stage is to identify groups of connected pixels which have larger signal values (relative to the maximum value found in the image) than the *first stage threshold.* The output of the first stage detection is the set of coordinates of the centres of each of these groups of pixels. Coordinates which are separated by a distance smaller than the *elimination width* are eliminated, since they are assumed to be too close to be from neighbouring subapertures.

The second stage begins by finding the brightest of the thresholded groups of pixels. A quasi-circular region is defined (where pixels are given a weighting between 0 and 1) around each of the coordinates. The edge of this circle is calculated to be where the average signal level in all directions drops to a fraction of the maximum defined as the *precision threshold.* Only the pixels inside this circle are used to find the centre-of-gravity. The result is used as the centre to define another circle and the entire process is repeated with the same radius of circle until the coordinate shift between one iteration

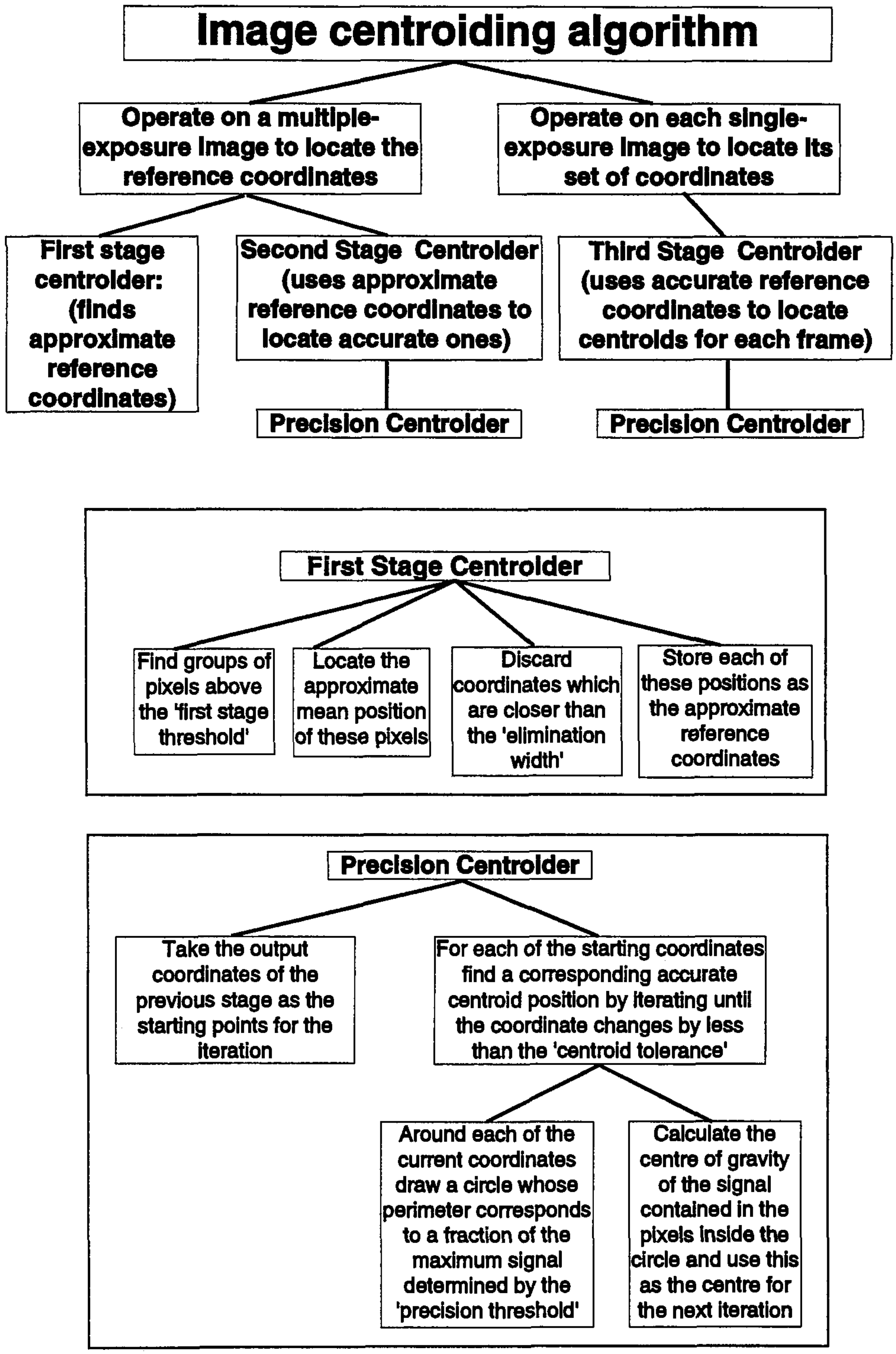



**Figure 4.1**: Modular diagram showing the major parts of the centroid algorithm (written by N.J. Wooder).

and the next is smaller than the *centroid tolerance.*

The third stage is performed on each of the single images. It is essentially identical to the second stage of the reference calculation, except with a different set of parameters to reflect the higher noise levels present in a short exposure image (*e.g.* higher thresholds and larger acceptable tolerances on the centroid coordinates). The reference coordinates (*i.e.* the output of stage 2) are used to define the circular regions used for the first iteration.

## 4.3 Statistical Analysis of Centroids

A set (referred to here as a frame) of centroid coordinates corresponding to each Shack-Hartmann image is thus calculated; the series of frames belonging to a run are stored on permanent media. This is a large amount of information, therefore these frames are only retrieved one-by-one for processing.

The calculation of the spatial statistics is made as follows. In order to study the evolution of $r_0$, the run will be divided into time intervals, for each of which a single estimate of $r_0$ is desired. The duration of this sampling interval must be chosen so as to encompass enough frames to avoid the biasing effect described in section 3.7.5, bearing in mind that the shorter this interval, the more regularly spaced (in time) will be the $r_0$ estimates. (for further discussion on obtaining sufficient numbers of *independent* measurements, see section 4.3.2). For each frame, possible pairs of subapertures are identified in turn. The coordinate system is rotated and aligned with the direction of the axis joining the two reference positions. Transverse and longitudinal centroid deviations can be calculated for the two subapertures in this coordinate system. In this way, sums and sums-of-squares of differential centroid positions are built up for every pair so that variances can be found for the entire run and for each sampling interval. An error on the variance can be estimated by summing the squares of the squares. An $r_0$ value in both directions is thus obtained for each sampling interval by taking the mean of $r_0$ calculated for every pair using the formulae of Sarazin and Roddier (Eqs. 3.50 and 3.51). The standard error on this value is found from the distribution of $r_0$ values from all pairs. Over the entire run, $r_0$ is also found for each subaperture pair. An error on this value is found from the standard error of the corresponding variance.

To calculate the temporal statistics, a temporal block size must first be defined. This is the length of the sequence of coordinates whose Fourier transform will be found in order to estimate the power spectrum. A large number of such estimates will be averaged together for each centroid over the whole run in order to accurately calculate the power spectrum. It is normally chosen to be a power of 2 (512 is the length chosen for the data described in Chapter 5) to make the most efficient use of the fast Fourier-transform algorithm; it is long enough to sample the spectrum finely, but short enough so as to fit a large number of blocks within a single run. An average power spectrum is calculated for all centroids in the two directions parallel and perpendicular to the direction of tracking of the star across the sky. The significance of this will become clear from the measurements in Chapter 5; the geometry of the telescopes used in those measurements results in these two directions corresponding to known directions on the detector arrays. A number of pairs of subapertures whose separations are equal are identified for calculation of a differential power spectrum. For every centroid, a time series of coordinates is built up using a sequence of frames equal to the block size. The squared modulus of the Fourier transform of each series is found; an estimate of the power spectrum is made by averaging these spectra from all centroids together. Several frame sequences make up a run; the mean of the power spectrum estimates from each of these is found to calculate an accurate power spectrum. This is also done for the differential case.

The entire processing algorithm is summarised in a modular form in Fig. 4.2.

### 4.3.1 Implementation of Algorithms

Computer programs have been written which incorporate algorithms to retrieve the data from the appropriate medium, to clean the images, subtract the bias and mask bad pixels, calculate a long-exposure image and calculate centroid positions for the long-exposure (reference) image and each single frame. The implementation is slightly different for each of the Shack-Hartmann/detector configurations described in Chapter 5. The centroid processing algorithm described in section 4.3 has also been implemented as a single callable function which is the same for each configuration. All algorithms are coded in the programming language 'C' and run on SUN SPARC-stations.

There follows a list of parameters which must be measured accurately in order to avoid any systematic error in the calculations:

- The focal length of the lenslets.
- The diameter and shape of the lenslets.
- The dimensions of the pixels in the detector array.
- The mean wavelength of the light detected.
- The exposure time and interval between frames.
- The ratio between the size of the telescope pupil and that of the re-imaged pupil in the plane of the lenslets.

These and other parameters, such as the $r_0$ sampling interval and the temporal block size, are chosen by the user.

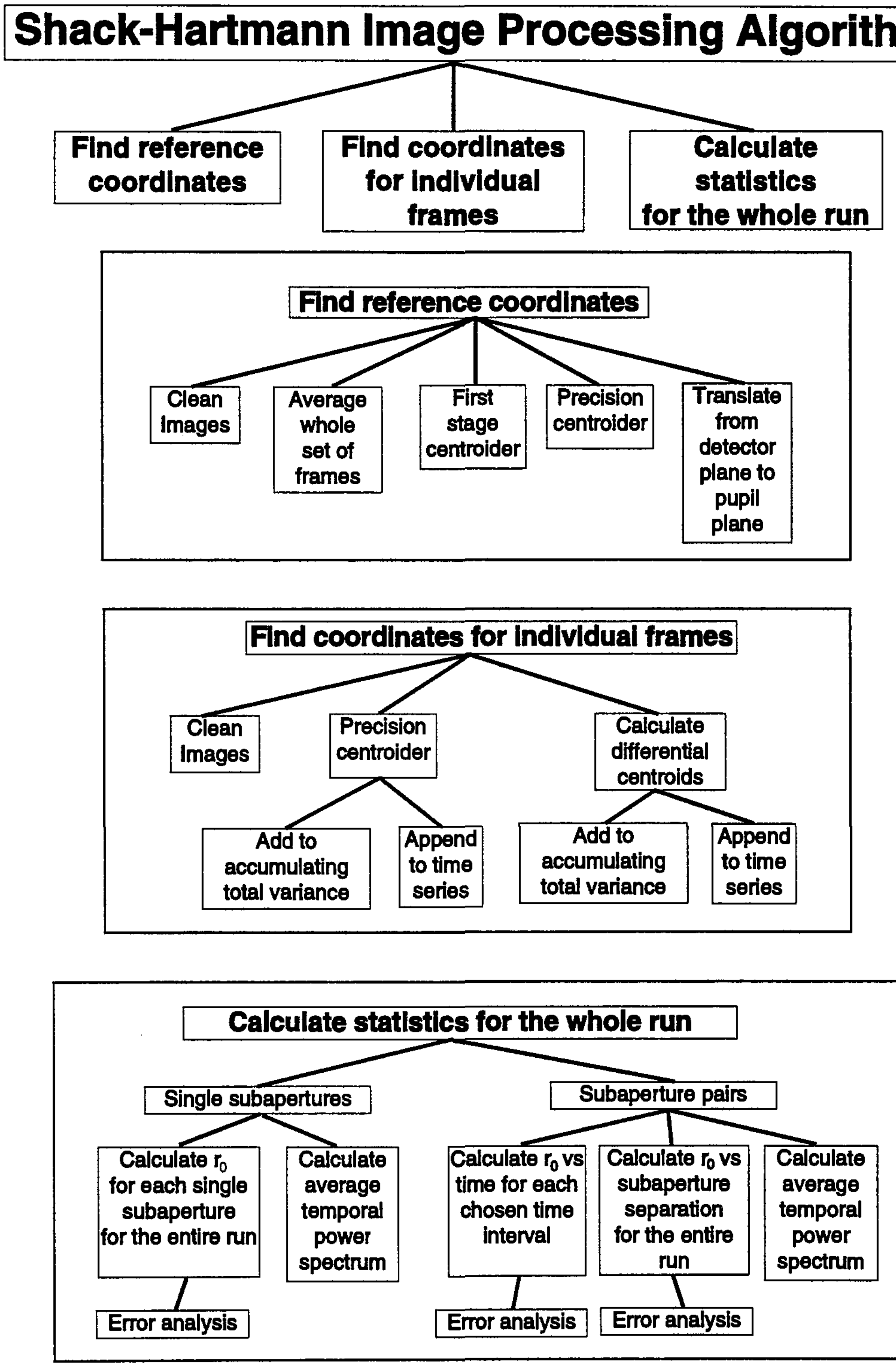



**Figure 4.2**: Modular diagram showing the centroid statistics processing algorithm.

### 4.3.2 Statistical Error Analysis

A discussion of likely sources of error involved in Shack-Hartmann measurements has already been made (section 3.7). The random errors involved could, in principle, be estimated if the noise characteristics of the system were entirely known. These errors are more readily calculable from a study of the spread of measurements made used to estimate a particular quantity; formulae for such calculations are presented in Appendix A.

In order to make some sort of estimate of the likely accuracy of a particular calculation, a knowledge of the correlation between measurement sample points is required. For example, if two estimators of a quantity are completely correlated, then there is no advantage to using both above one; if uncorrelated, the estimates can be considered independent. In the more general case where there are $n$ partially correlated estimators of the quantity (separated in space or time, or both), then the effective number of independent measurements can be considered to be somewhere between 1 and $n$. In the application of Shack-Hartmann image processing to calculate $r_0$ values there are at least two cases where the correlation between $r_0$ estimators must be considered in order to be able to place an error bound on the result obtained. Firstly, where a time series of centroids from a single image or pair of images is used to make an estimate of centroid motion variance (and consequently $r_0$), unless the time interval between frames is much larger than the wavefront correlation time, one cannot consider successive centroid positions to be independent. In this case, the temporal correlation (or covariance) of the centroid position or centroid differential must be studied in order to make an estimate of the 'redundancy factor' $F$ defined in Eq. A.11. Specifically, the time series of measurements from a single centroid or pair must be examined for correlation as a function of time interval between measurements (this imples correlating every frame with every other frame). The temporal characteristics of different pairs will inevitably be different. However, calculation of temporal covariance is impractical for all combinations of pairs, simply because of the large processing time required. Therefore, when examining a particular run, only a small number of pairs will be examined in this way. The second case of interest is where a large number of combinations of pairs of Shack-Hartmann images are all used to estimate a value $r_0$, using a calculation of the differential centroid variance for each pair over a large number of frames. In this case, the number of pairs is likely to be larger than the number of images; each image may combine with a number of other images to form several pairs, and its centroid position in any particular frame may consequently be used in the calculation of many of the differential variances. Therefore, one would expect a fairly high degree of correlation between these $r_0$ estimators. The covariance between centroid differentials must therefore be studied to estimate the redundancy. There may also be some correlation between the angles-of-arrival caused by turbulence at closely separated locations in the aperture. The effects of this will also be taken care of in the covariance calculation.

# Chapter 5

# Seeing Measurements

## 5.1 Overview of Experiments

As discussed in section 4.1, much of the seeing information required from a site monitoring campaign can be obtained from $r_0$ measurements and centroid power spectra. Before describing the specific measurement techniques and the results obtained, it is necessary to discuss some of the practical considerations involved. Ideally, one would wish to have a system which would be kept running at all times and which could observe a star close to the zenith continuously. However, if measurements are to be made using a telescope, as here, only a small amount of time will inevitably be available for such use. In the present case, short periods of observation have been allocated at irregular times around the year for performing seeing measurements.

Observations have been necessarily divided into individual runs using a suitable reference star for each run; the most appropriate star would normally be chosen as the star closest to the zenith which is bright enough to give a satisfactory signal-to-noise ratio (see the discussion in section 3.7). However, if there is a requirement to measure the dependence of seeing on declination angle or telescope orientation with respect to the sky, stars in the appropriate regions in the sky need to be selected.

A telescope is normally housed in a dome, with neighbouring buildings containing various instrument rooms and the control room. If the effects of the atmosphere are to properly accounted for (indeed, if any future adaptive system is to run at optimum correction), then it is vital that dome seeing be reduced to a minimum. This can be done by reducing any heat sources in the building to a bare minimum and by trying to equalise the temperature between inside and outside the dome. Sensible practice involves opening the dome some time before observing to try to flush out warm air from inside.

Three different experimental configurations of wavefront sensor have been selected for their own reasons; firstly, a high precision system for measuring spatial phase statistics in the visible spectrum only; secondly, a fast readout system for measuring temporally correlated wavefront information in the visible; and thirdly, a novel configuration designed to obtain temporal information in the near infrared. The locations and dates of use of these systems are summarised in table 5.1.

| Date | Location | Mean wavelength | Typical Exposure Time |
|---|---|---|---|
| Dec '93 | La Palma (JKT) | $650nm$ | $15ms$ |
| July '95 | Calar Alto | $655nm$ | $2.5 - 5ms$ |
| August '94 | Mauna Kea (UKIRT) | $2200nm$ | $16 - 33ms$ |

**Table 5.1**: Table of observations made and their respective sites.

Brief descriptions of the three wavefront sensing systems will now be given. Following this, some typical results will be described. The purpose will be to test some of the aspects of the Kolmogorov model and the Taylor hypothesis, and also to compare the merits of the three sites which were visited.

## 5.2 High-Precision Measurement Of Spatial Wavefront Information

### 5.2.1 Objectives of the Visible Shack-Hartmann Observations

The purpose of these experiments was to make highly accurate angle-of-arrival measurements using light from the visible part of the spectrum. A detector with 16-bit storage capabilities was used (to obtain high precision) whose readout noise could be kept low by using a very slow readout rate, the drawback being that no temporal information could be obtained. Such high precision is particularly useful when calculating the modal expansion of individual wavefronts (see elsewhere [62]). The wavefront sensor was designed to make it possible to take speckle images of the source simultaneously with the wavefront measurements with the purpose of performing some wavefront deconvolution. Some earlier measurements using this system have been described by J-Y. Zhang in his PhD Thesis [63].

The measurements were made at the Jacobus Kapteyn Telescope (JKT), which is an equatorially-mounted telescope with a 1-metre diameter parabolic primary mirror, situated at the Isaac Newton Observatory site on the island of La Palma in the Canary Islands at an altitude of 2400 metres. The hyperbolic secondary mirror, which gives a conventional Cassegrain optical system with a focal ratio of $f/15$, was used. This creates a central obscuration of approximately 0.5m diameter. The site also contains two other much larger telescopes. The JKT is particularly suitable for seeing measurements for a number of reasons. It is not in such high demand for astronomers who require the larger light collecting capabilities and the range of instrumentation available at the larger telescopes. A smaller dome size is generally accepted to be less susceptible to local effects and the local environment can be more easily controlled by visiting observers.

### 5.2.2 Instrumentation

A diagram of the system used is given in Fig. 5.1. The $f/15$ beam produced by the telescope optics is split into two using a beamsplitter cube. One beam is collimated and then passes through a lenslet array positioned in the pupil plane before being directed onto the CCD camera for detection. The other beam is used to form a direct image of the source adjacent to the Shack-Hartmann image on the CCD, using a microscope objective positioned away from its proper conjugate position. The lenslet array was manufactured in photoresist by Observatoire de Paris and consists of square apertures with dimensions of $0.5 \times 0.5$mm and a focal length of 45mm. A collimating doublet with focal length of 80mm was chosen to give a subaperture size of just under 100mm at the telescope pupil.

High-precision mechanical components are incorporated to allow small manual motions to be made, thus making easier a careful alignment of the optical system. In the wavefront sensing arm of the system, the collimating lens is placed untilted, on axis and so that its focal point coincides with the telescope focus position. This is done with the assistance of an artificial alignment tool, which consists essentially of a spatially filtered and focussed laser beam. The beam of starlight can later be brought to the same position using the telescope guiding controls and a pinhole which defines the location of the telescope focus. The lenslet array must be placed as close to the re-imaged pupil plane as possible and untilted, to avoid unnecessary aberration of the subaperture images (coma in this case). This can be done by observing the point-spread functions of the individual images while tilting the lenslet array. This is best done using bright starlight during a period of good seeing. The annular pattern of illumination is made symmetric with respect to the grid of lenslets (*i.e.* apertures at one edge of the pupil are illuminated with the same pattern as those on the opposite side) by small adjustments of the position of the lenslet array relative to the axis of the beam.

The detector used is a Princeton Instruments CCD camera containing a Tektronics TK512 chip, a front-illuminated CCD with a 512 square array of pixels, each with dimension $27\mu m \times 27\mu m$. The controller can read the chip at between 25 and 200 kHz pixel readout rate with up to 16-bit precision per pixel. The detector was calibrated to

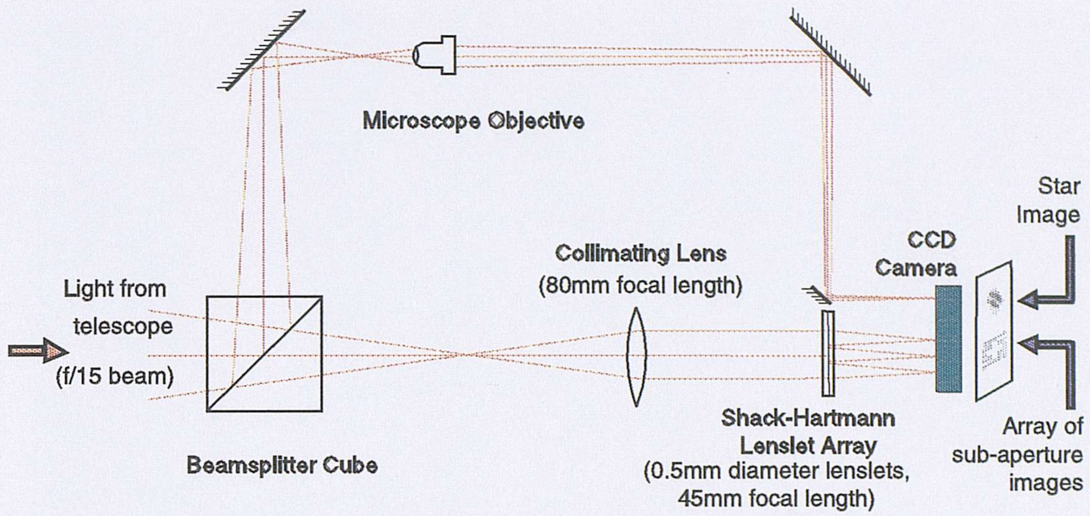



**Figure 5.1**: Diagram of the simultaneous Shack-Hartmann wavefront sensor and speckle imager used at the JKT in La Palma.

register one count for 8 photons detected. It is a silicon detector, and therefore can be considered to have a peak wavelength detection in the red part of the visible spectrum. The readout noise was measured by the manufacturer to be 1.36 counts rms at 50kHz pixel readout rate and 2.64 counts at 200kHz. The chip is Peltier cooled in air to $-5$ Celsius, giving a negligible dark count at typical exposures of 15ms. The Princeton camera is controlled by a SUN SPARC IPC via a GPIB interface. Images are recorded directly on DAT using a system controlled by a graphical-user-interface running on the SUN.

The adjacent Shack-Hartmann and speckle images occupy a 256 by 512 pixel region of the CCD array (see figure 5.2). At the fastest readout speed, images can be obtained at the rate of approximately two frames per second using this configuration. The camera makes use of a mechanical shutter with a minimum exposure time of approximately $15ms$ due to the large size of the chip. This is almost certainly not fast enough to 'freeze' the turbulence fluctuations, or to obtain consecutive wavefronts which are temporally correlated. The measured resolution may thus appear better than in reality.

Although this system is clearly not capable of giving time-resolved wavefront information, it is a high-performance system due to the high dynamic range and the low dark count. An example of a short-exposure image acquired during the December 1993 run is shown in Fig. 5.2.

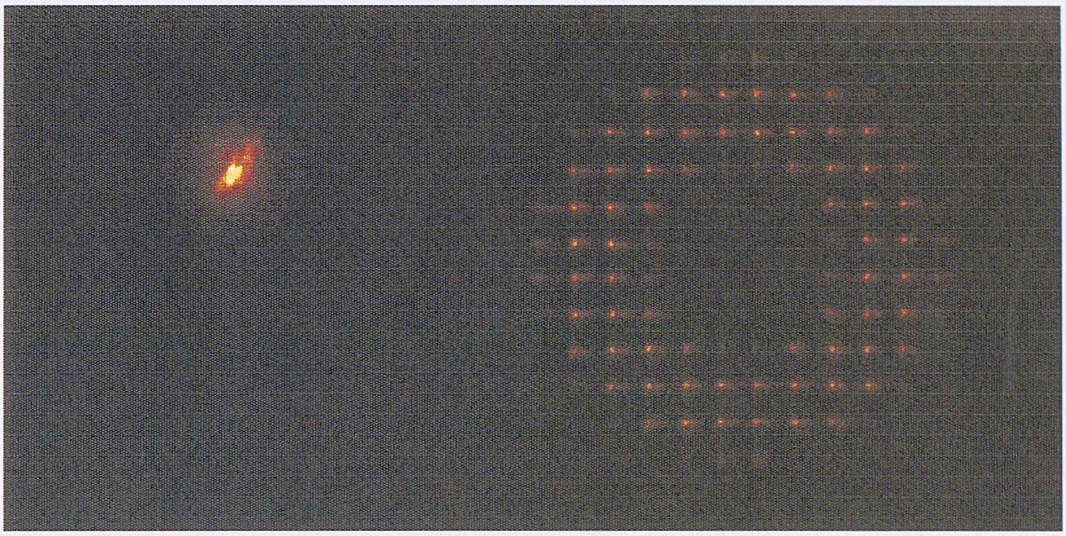


**Figure 5.2**: A short-exposure image acquired with the Shack-Hartmann/Speckle imaging system at the JKT.

## 5.3 Temporally Resolved Infrared Wavefront Measurements

### 5.3.1 The Difficulties of Infrared Seeing Measurements

Diffraction-limited images with a 4-metre telescope at 0.5 $\mu$m would represent a 30 times increase in resolution over images taken in 1-arcsecond seeing conditions. However, at 2.2 $\mu$m, the potential increase in resolution is only about 7 times, assuming the Kolmogorov scaling of seeing with wavelength. Nevertheless, improving the resolution of imaging in the infrared is considered to be of prime importance, despite the relatively poor angular resolution, due to the large number of potentially resolvable objects in the sky at these wavelengths.

Conversely, the relatively poor gain in resolution at these wavelengths is offset by the expected larger correlation scale and time of the turbulence (Eqs. 2.59 and 2.122). This means that seeing measurements are made easier by the need to sample the wavefront with lower spatial and temporal resolution. Although the physical geometry of the wavefront can be considered the same at all wavelengths (the refractive index fluctuations result in a wavelength-independent path change), the magnitude of the corresponding phase fluctuations is inversely proportional to the wavelength. Therefore, a smaller aperture is required at shorter wavelengths in order to keep the phase fluctuations below the tolerance needed for diffraction-limited imaging. Also, the correlation time of the full-aperture tilt (which is essentially what is measured through a Shack-Hartmann subaperture) is shorter for the smaller apertures which are, therefore, necessary for shorter wavelengths. Indeed, Mariotti [64] has measured correlation times

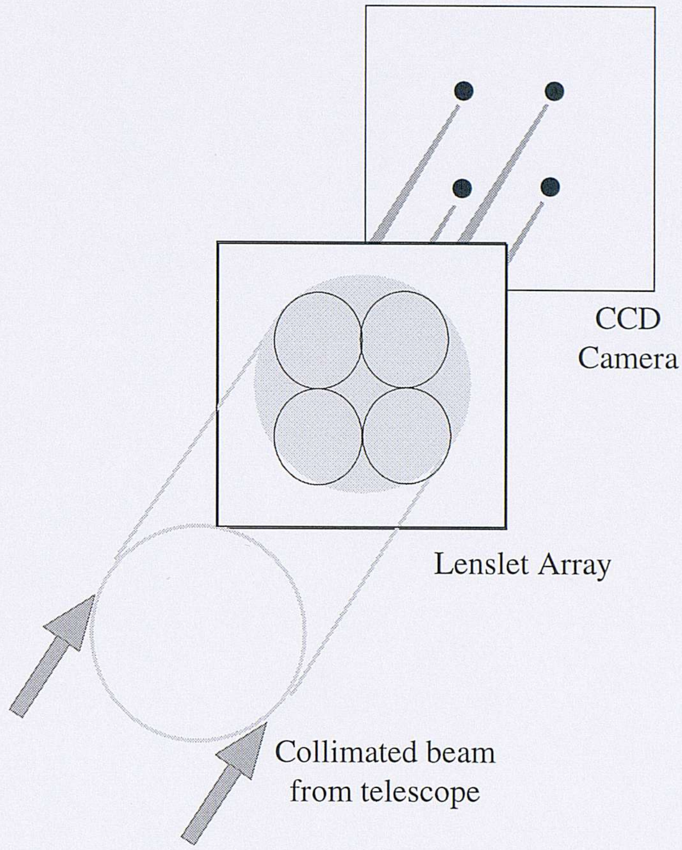



**Figure 5.3**: Principle of the standard Shack-Hartmann wavefront sensor, showing the unused space between subaperture spots.

of between 50 and 100 milliseconds at 2.2 $\mu$m, and between 100 and 300 milliseconds at 4.6 $\mu$m. This was calculated as the $1/e$ correlation time of intensity at a specific location in a speckle image. This would correspond to around 10ms at $0.5\mu m$, although times considerably smaller than these would be required if the fluctuations are to be properly 'frozen' for wavefront sensing (*i.e.* $1/e$ is a rather large decay in intensity). A Shack-Hartmann array with 12 by 12 lenslets running at an exposure rate of 50Hz ought, therefore, to be capable of measuring the time evolution of the perturbed wavefront under a range of conditions at a 4-metre telescope (assuming $r_0$ to be approximately 100mm in the visible). Alternatively, the seeing statistics could be measured using far fewer lenslets (each with a diameter of the same order as $r_0$), without the need to sample the entire aperture. In the past, seeing measurements of the type described in section 5.2 have been virtually impossible at infrared wavelengths. This has been due to the lack of availability of fast detector arrays; in the past, techniques have relied on single detector elements [65].

A small amount of observing time was allocated at the UK Infrared Telescope (UKIRT) in Hawaii for infrared Shack-Hartmann measurements. UKIRT is a 3.8-metre equatorially mounted telescope situated on Mauna Kea, a mountain on the 'big island' in

Hawaii at an altitude of approximately 3500 metres. The intention was to use a novel Shack-Hartmann sensor together with the latest fast-readout infrared detector technology to obtain information about the spatial and temporal properties of the wavefront in the infrared.

### 5.3.2 A Modified Shack-Hartmann Sensor

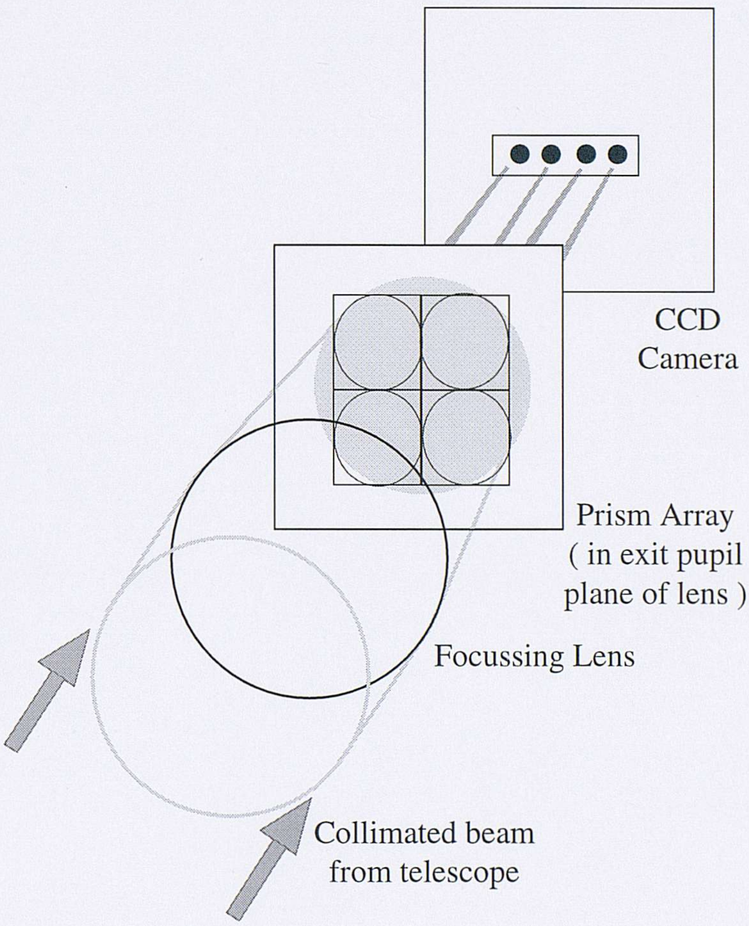



**Figure 5.4**: Principle of the coordinate-remapping Shack-Hartmann wavefront sensor.

A new infrared detector has recently been installed at UKIRT as part of the upgraded IRCAM3 system. A transputer network, known as ALICE, has been designed to allow data output through 16 parallel channels. By August 1994 the output was only configured to run through a single channel, limiting the pixel readout rate to 1MHz. At this speed, the readout noise was measured to be approximately 57 electrons [66]. With a standard Shack-Hartmann sensor configuration, the very fastest frame rates desired would not be available. If the positions of the subaperture images were to be moved with respect to each other on the CCD, however, a significant gain in readout speed could be achieved (see Fig. 5.4) for three reasons:

- To reduce the area on the detector array to be read out.

  Figure 5.3 shows that the readout time is unnecessarily prolonged by some dead space between the subaperture images. This is inherent in the design of a standard Shack-Hartmann wavefront sensor. A coordinate re-mapping system would, in principle, be able to utilise some of this dead space.

- To make the active area on the CCD the most efficient to read out.

  The transputer network which delivers the data to the outside world can be programmed so that only part of the array is read out with little overhead time. The fastest readout speed would be achieved by positioning the spots on the CCD so that this overhead is reduced to a minimum.

- So that the aperture can be deliberately undersampled to gain in readout speed.

  With a standard Shack-Hartmann system, the area on the CCD which must be read out always matches the active area on the lenslet array. If space is left between the individual lenslets, an equivalent area on the CCD would still have to be read. The re-mapping could remove this area so that the area which must be read is proportional only to the number of subapertures.

Such a sensor has been constructed [67], using an array of prisms to replace the lenslets, and installed in the filter wheel in the dewar at UKIRT. Each prism redirects its beam to a new location on the detector without altering the relative motion of the centroid with respect to its reference position. There are 24 prisms in total, each aligned with a circular hole in the pupil mask of effective diameter 31cm at the telescope pupil. These are positioned in two concentric rings in the pupil. A 25th subaperture image is produced from a hole in the mask. The actual diameter of the array is 17mm, each prism being 2.5mm square, with the holes in the mask having a diameter of 1.4mm. The prisms have been optically contacted onto the glass substrate and are positioned immediately behind a lens of focal length 80mm. The array was manufactured by ICOS and the entire assembly was placed into the filter wheel and cooled some months before the observations took place. The subaperture images are redirected so as to form two lines of images on the CCD. Figure 5.5 shows a scale diagram of the prism layout in the telescope pupil together with the corresponding positions of the images formed on the detector. A major advantage of this coordinate-remapping configuration is its versatility. The readout electronics can, in principle, be programmed so that different combinations of subapertures can be chosen to be read out. For example, spots from opposite sides of the pupil could be positioned side-by-side and read together.

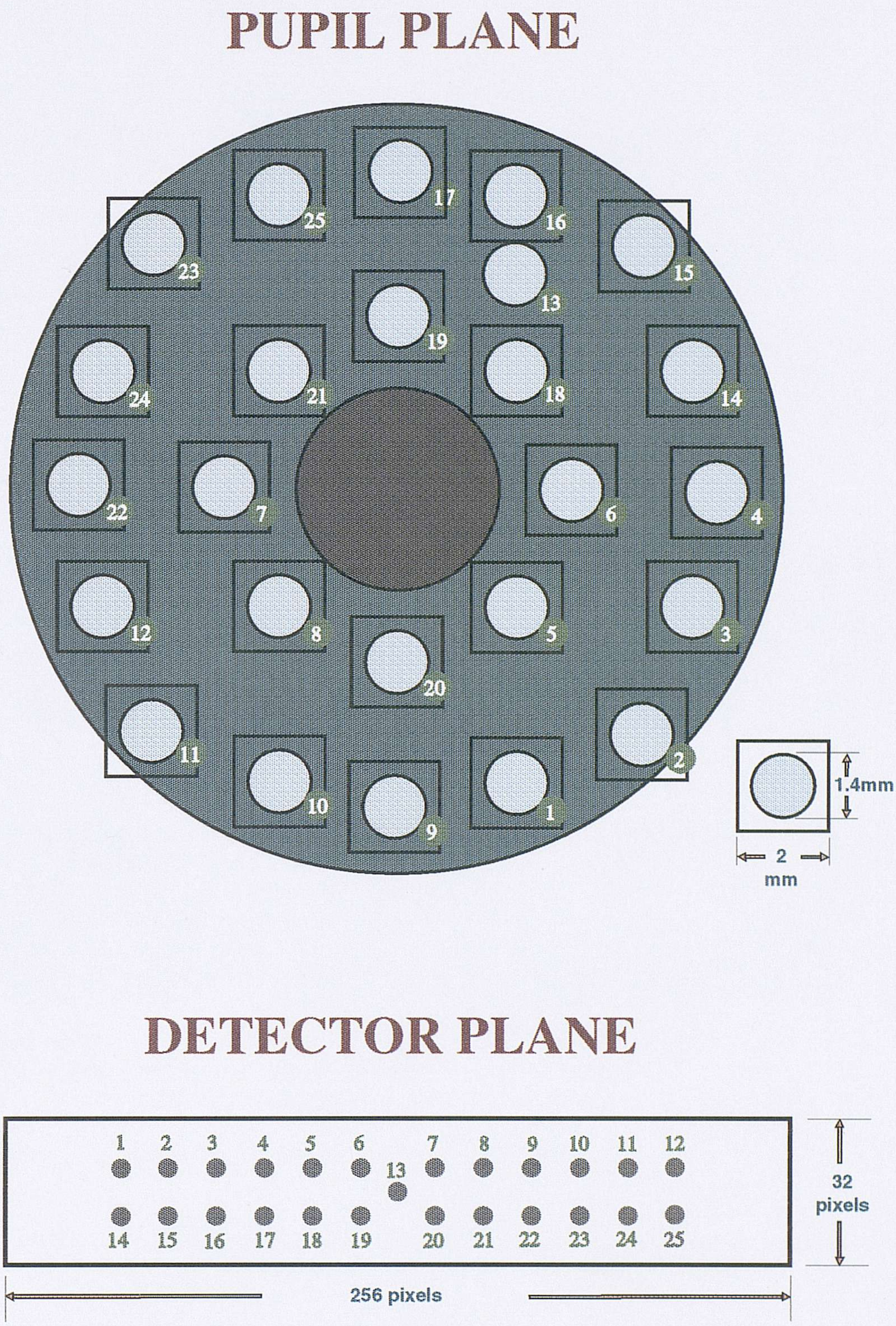



**Figure 5.5**: Scale diagram of the prism geometry in the pupil plane of the 3.8-metre telescope shown with the corresponding subaperture image locations on the detector array. The 25th subaperture, which is a hole in the mask, is also shown (spot 13), together with the central obscuration.

### 5.3.3 Data Acquisition System

At the time of the observing run, no on-line system existed which would enable data to be stored at the highest rate allowed by the readout electronics. A complete data acquisition system has therefore been constructed for this purpose and is designed to link up to the transputer system. The ALICE transputer system was programmed for the Shack-Hartmann configuration by P. Puxley at ROE. Optical disk was chosen as the medium for permanent storage of data, for reasons of speed, and also since it was appropriate for infrared speckle measurements being performed on the same observing run. However, at the storage speed required for the seeing measurements even optical disk is not fast enough. Therefore, the data is first dumped onto magnetic disk and later backed up onto optical disk, resulting in a large overhead time between observations. This also places a limit on the maximum length of each seeing run.

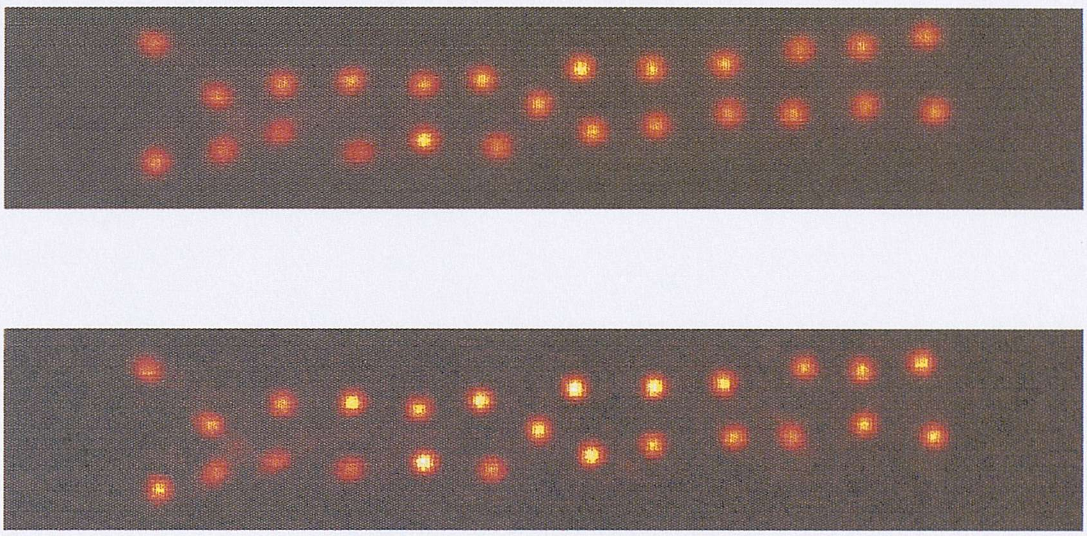


**Figure 5.6**: Multiple-exposure image (average of 3600 frames – top) and a single image (bottom) acquired at UKIRT using the 256x48 pixel readout mode.

Two readout configurations have been programmed. The first allows the reading out of an area on the CCD which contains the entire array of subaperture images. Unfortunately, the prism array was placed in the system slightly rotated with respect to the optical axis, causing the lines of subaperture images to appear not quite horizontal on the CCD. This results in a larger number of pixel lines being required for readout of the whole array (48 instead of 32) increasing the readout time by 50%. The resulting frame rate in this case is approximately 33 frames per second. Figure 5.6 shows a single short exposure and a long exposure image taken with this configuration. A few of the spots appear to be slightly shifted away from their nominal positions. This is believed to be due to the failure of the optical contacting process, so that the respective prisms have moved, a hypothesis confirmed by much larger shifts appearing in the

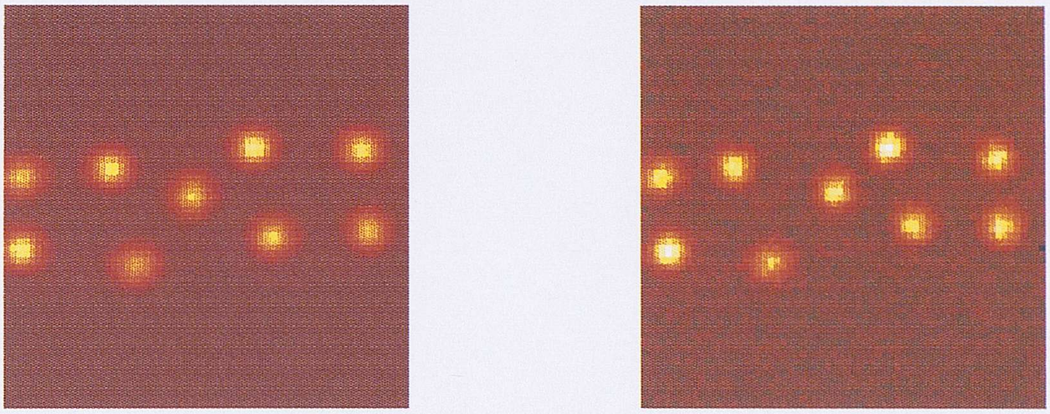


**Figure 5.7**: Multiple-exposure image (average of 3600 frames, left) and a single image (right) acquired at UKIRT using the 64x64 pixel readout mode.

months subsequent to this observing run. However, it is not thought that these prisms would have moved during the period of a single observation. Furthermore, these shifts probably represent very small movements in the pupil plane.

The second configuration which has been programmed to be read out is a $64 \times 64$ pixel region in the centre of the array. The 8 prisms in the central ring of the Shack-Hartmann array form their images within this region of the CCD, together with the central image resulting from the hole in the mask. In this case, frames can be recorded at intervals of 16ms. Multiple and short exposure images taken in this configuration are shown in Fig. 5.7. Failure to align the telescope correctly with the reference star would sometimes lead to two of the spots at the edge of the region being lost.

## 5.4 Temporally Resolved Visible Wavefront Measurements

The intention was to make some temporally resolved Shack-Hartmann measurements in the visible range of the spectrum. Although the detector technology is more readily available at visible wavelengths, due to the wavelength dependence of the seeing (see Eq. 2.122) the sampling rate required at $0.55\mu m$ is approximately 5 times higher than at $2.2\mu m$.

### 5.4.1 Description Of The System

A DALSA CCD is used, containing a CA-D1128 chip (128 square frame transfer CCD), which can be clocked at up to 16MHz pixel readout rate. This is a fast system, with the disadvantage of only 8-bit precision and not particularly good noise characteristics (these are not known exactly). The maximum typical frame rate obtainable with this system is 850 frames per second. The detector is intensified to increase the overall quantum efficiency. The intensifier was manufactured by Instrument Technology Ltd to a z-stack MCP design developed at ICSTM [68]. It incorporates an S20 Photocathode with a sapphire window and a P20 Phosphor Screen, whose nominal peak quantum efficiency occurs at a wavelength of $0.655\mu m$. The CCD has a digital output, which is connected to a bit-flow raptor PCI based frame-grabber. A DELL PCI PC is used to read out the data, utilising a PCI-based SCSI controller, and running a WINDOWS VISUAL C program which is capable of transferring data onto magnetic disk at 400 frames per second. Approximately 2000 frames (limited by the size of the disk) can be acquired in this way before the data is required to be transferred elsewhere. DAT is chosen as the medium for permanent data storage. During the back-up operation a large amount of overhead time is lost between observations. Therefore, each observing run consists of blocks of data (typically 5 seconds in length) separated by large amounts of dead time (approximately 40 seconds). It is possible to reduce the readout rate if the chosen reference star is relatively dim.

The lenslet array used was the same as that described for the high-precision wavefront measurements in section 5.2. A 60mm focal length collimator is used in the f/8 beam (created by the telescope optics) so that the subaperture size is kept below 10cm in the telescope pupil. Relay optics are used to reimage the intensifier output onto the CCD and are configured so that the subaperture image diameter is approximately 3 pixels on the detector. The optical alignment procedure is also similar to that described earlier for the system used at La Palma.

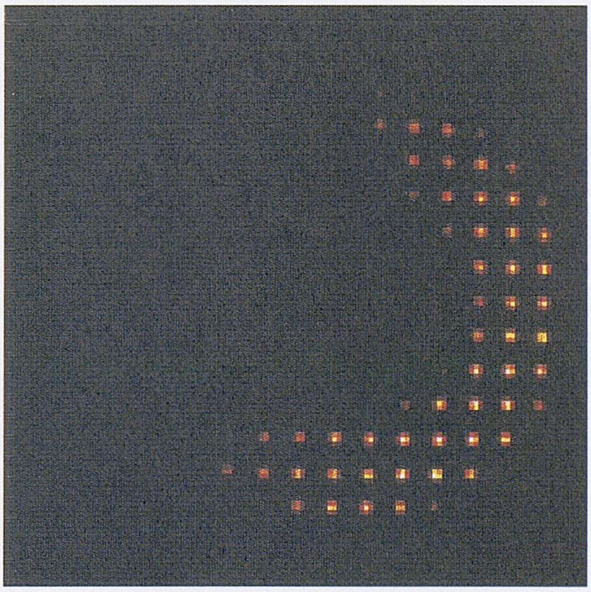


**Figure 5.8**: An example of a multiple-exposure image (average of 2400 frames - set L) acquired at Calar Alto.

The measurements were made at the 1.23-metre telescope situated at Calar Alto in Spain, an observing site accommodating several telescopes at an altitude of 2170 metres.

The telescope is similar in design to the JKT, although the dome is considerably larger. At the time of the observations, the reflective coating on the primary mirror had been accidentally partially removed, leaving a non-symmetric pupil. A typical multiple-exposure Shack-Hartmann image is shown in Fig. 5.8. The apodisation is clearly visible. Some images are completely invisible, while a few images, which are at the edge of the reflecting region, are partially reduced in intensity. There is a danger that the non-uniformity of the mirror may cause inaccuracies in the centroid determination, particularly for these dimmer images. Hence, a decision has had to be made as to which images should be be discarded. A threshold of 55% was found to be satisfactory (images without at least one pixel with a signal level higher than this fraction of the maximum intensity found in the reference frame were ignored). A further problem is caused by the blocked nature of the data acquisition. The failure of the automatic telescope tracking mechanisms leads to a shift in the reference coordinates from one block to the next. This effect is so severe that these coordinates must be recalculated at the start of each block. In order to preserve the identity of each subaperture from one block to the next (subaperture images are numbered arbitrarily), the reference coordinates from the previous block are used to calculate the new coordinates for any particular block. The second stage of the centroid process (see section 4.2.3) must be run on a multiple-exposure image formed from all images within a particular block to provide the reference coordinates for that block, using the set of reference coordinates from the previous block as its input.

## 5.5 Summary of Data Acquired

A number of data sets were recorded at each site. A summary of the sets for which stars close to the sky zenith were used to provide the illumination is given in table 5.2. A number of other sets were also recorded using stars at large declination angles from the zenith in order to study the seeing variation from one part of the sky to another. These are not listed here.

Centroids have been found for each set, using thresholds (see section 4.2.3) which reflect the noise levels and image sizes for each configuration. Analysis of the centroid motions has been made and the results have been assessed in a number of ways. Firstly, the distribution of $r_0$ values from all the possible subaperture pairs has been examined. Secondly, the dependence of seeing on subaperture separation has been analysed. Thirdly, samples of the $r_0$ evolution over time have been plotted, and an examination of their dependence on declination angle has been made. Finally, the temporal power spectra of centroid motions have been calculated and examined for the temporally resolved measurements.

| Set | Site | Night Of | Universal Time | $\lambda$ ($\mu$m) | Exposure Time (ms) | Number of Frames | Centroids Located |
|---|---|---|---|---|---|---|---|
| A | La Palma | 26/12/93 | 20:38 | 0.60 | 15 | 1000 | 86 |
| B | La Palma | 26/12/93 | 22:39 | 0.60 | 15 | 1000 | 86 |
| C | UKIRT | 02/08/94 | 06:46 | 2.2 | 30 | 3500 | 25 |
| D | UKIRT | 02/08/94 | 06:54 | 2.2 | 30 | 3500 | 25 |
| E | UKIRT | 02/08/94 | 07:01 | 2.2 | 30 | 3500 | 25 |
| F | UKIRT | 02/08/94 | 07:28 | 2.2 | 16 | 3600 | 9 |
| G | UKIRT | 02/08/94 | 07:31 | 2.2 | 16 | 3600 | 9 |
| H | UKIRT | 03/08/94 | 05:31 | 2.2 | 16 | 5000 | 7 |
| J | UKIRT | 03/08/94 | 07:23 | 2.2 | 16 | 19000 | 7 |
| K | Calar Alto | 12/07/95 | 21:20 | 0.65 | 3.9 | 20400 | 42 |
| L | Calar Alto | 12/07/95 | 23:10 | 0.65 | 2.6 | 16320 | 43 |
| M | Calar Alto | 12/07/95 | 01:55 | 0.65 | 3.9 | 16320 | 42 |

**Table 5.2**: Summary of data sets for which centroids have been found.

$r_0$ values were calculated assuming a wavelength of $2.2\mu m$ for the infrared data acquired at UKIRT and $0.55\mu m$ for the visible data acquired at the JKT and Calar Alto. Although the measurements were recorded in the red part of the visible spectrum in the latter two cases, an adjustment was made because $0.55\mu m$ is often chosen as a standard for expressing the results of seeing measurements.

## 5.6 Analysis of Correlation of Measurements

As discussed in section 4.3.2, the calculation of the statistical errors involved in $r_0$ measurements depends critically on the degree of correlation between measurement points. Therefore, an assessment of this correlation must be made for every data set before error bounds can be placed on the calculations. The temporal correlation between measurements taken from series of frames or the spatial correlation between differential $r_0$ estimators (or both) must be calculated depending on the behaviour which is being studied.

The effective number of independent measurements of $r_0$ made by a single differential subaperture pair is often less than the number of frames acquired due to the temporal correlation between measurements in successive frames. This correlation will, inevitably, depend on the separation of the two apertures making up the differential pair, since wavefront tilts over larger regions decorrelate more slowly. However, the calculation of a temporal decorrelation is a time-consuming task even for one subaperture pair, since the measurement taken in one frame must, in principle, be cross-correlated with the measurement in every other frame. Therefore, for each data run, the temporal correlation from a single selected pair will be taken as representative of all pairs during that run. If $\alpha'(t)$ is the angular separation (longitudinal or transverse; relative to the mean separation) between the centroids found for that pair at time $t$, and $\delta t$ is the time interval between frames in the run of $n$ frames which began at time $t_0$, then the covariance is calculated as:

$$C_{\alpha'\alpha'}(j\delta t) = \frac{1}{n-j}\sum_{i=0}^{n-j-1} \alpha'(t_0 + i\delta t)\,\alpha'(t_0 + [i+j]\,\delta t)\,. \tag{5.1}$$

This correlation has been calculated for a range of frame separations, $j$, and has been plotted for values of $j$ up to an arbitrary cut-off value (where the correlation is adjudged to have become insignificantly small). Graphs of covariance against frame number for a set chosen from each of the UKIRT and Calar Alto observing runs are shown in Figs. 5.9 and 5.10. As expected, there was found to be no significant temporal correlation in the data taken with the Princeton detector at La Palma.

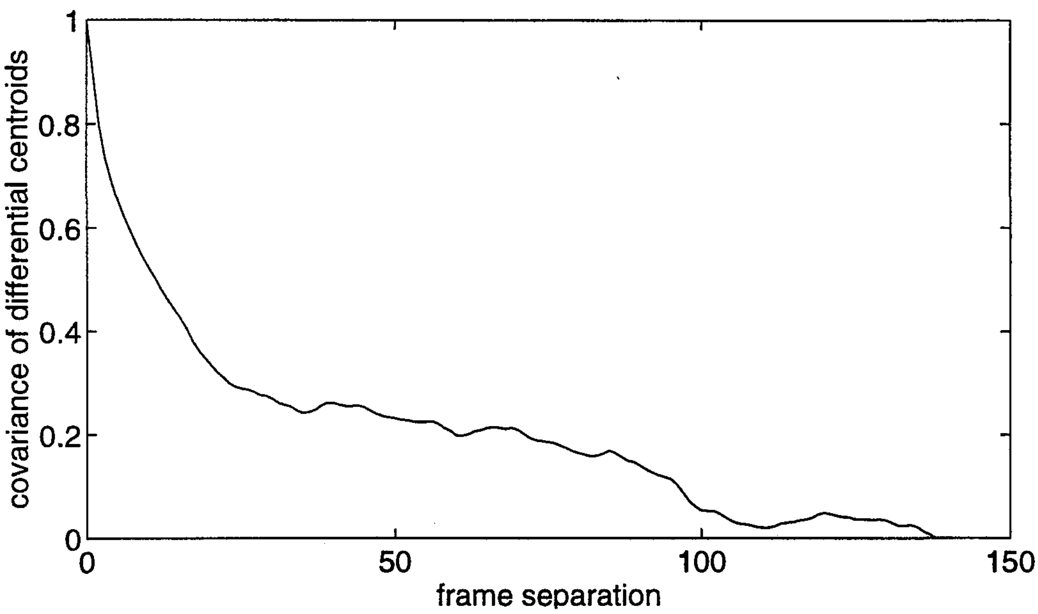



**Figure 5.9**: Temporal covariance of the centroid position difference for one selected pair of subaperture images (set C: UKIRT). The values are normalised so that the covariance is unity for zero frame separation.

The redundancy factor (defined in Eq. A.11), which is the factor by which the number of measurements actually taken must be reduced in order to take into account the temporal correlation when calculating error bounds, can then be calculated as follows:

$$\mathcal{F}_t = 2\sum_{j=0}^{n-1}\left(\frac{C_{\alpha'\alpha'}(j\delta t)}{C_{\alpha'\alpha'}(0)}\right)^2 . \tag{5.2}$$

For the two sets whose correlations are plotted, $\mathcal{F}_t = 21.8$ for set C and $\mathcal{F}_t = 20.6$ for set L.

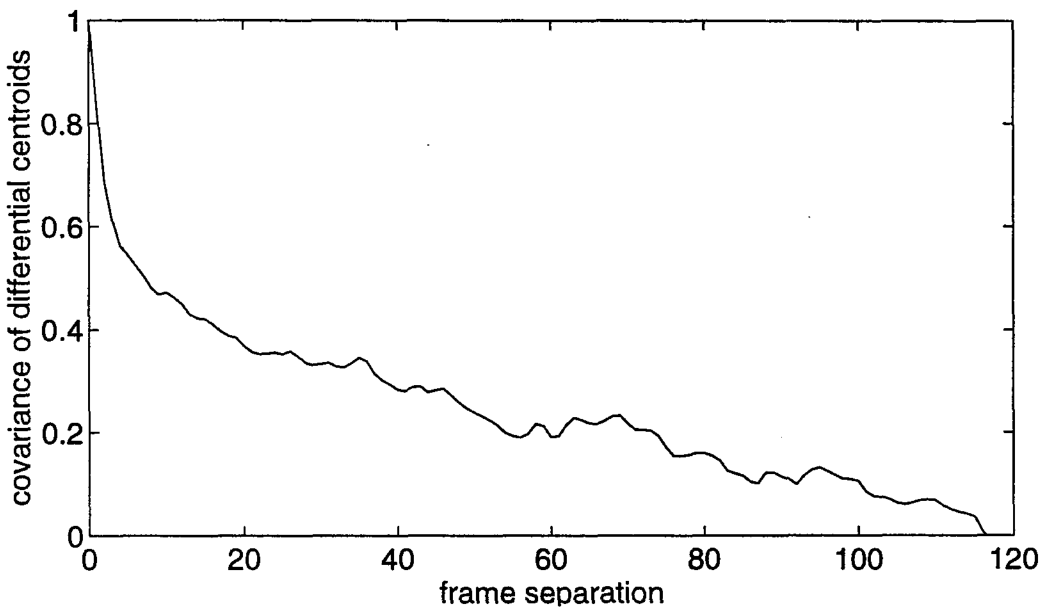



**Figure 5.10**: Temporal covariance of the centroid position difference for one selected pair of subaperture images (set L: Calar Alto). The values are normalised so that the covariance is unity for zero frame separation.

To use the Shack-Hartmann sensor as a differential image motion monitor, each subaperture image is taken in combination with other images to form a number of differential pairs. The spatial correlation between these pairs also causes a reduction in the effective number of independent measurements. The fact that in any particular frame, the displacement of one centroid may be used in the calculation of differential centroid motion variance for more than one pair contributes to the spatial correlation in a way which is a function of the subaperture geometry. A second major contribution is the correlation between the angle-of-arrival measured in one part of the aperture and another, which is a property of the turbulence itself. Once again, the measurement of the correlation between each pair and every other pair is impractical, due to the large number of pairs involved. For this reason, the degree of spatial correlation between pairs over a complete set of frames has been estimated by calculating the covariance between the measurements taken by a selected pair and measurements from all the other pairs. This has been done separately for each run of observations. If the two subapertures selected for this purpose are labelled $l$ and $m$, then the covariance between the angle-of-arrival difference measured by this pair, $\alpha'_{lm}$, and that measured by the pair formed from subapertures $j$ and $k$ (denoted by $\alpha'_{jk}$) is calculated as follows:

$$C_{\alpha'_{jk}\alpha'_{lm}} = \frac{1}{n}\sum_{i=0}^{n-1} \alpha'_{jk}(t_0 + i\delta t)\, \alpha'_{lm}(t_0 + i\delta t)\,. \tag{5.3}$$

Again, $t_0$ is the start time of the run with $n$ frames recorded at intervals of $\delta t$. The spatial redundancy (*i.e.* the factor by which the number of pairs of apertures must be reduced to obtain the effective number of independent $r_0$ estimators) is estimated as follows:

$$\mathcal{F}_s = 2\sum_{i=1}^{N}\sum_{j=i}^{N} C'_{\alpha'_{lm}\alpha'_{ij}}{}^2. \tag{5.4}$$

where $N$ is the number of subaperture images located in each frame. In practice, the sum is taken only over subaperture pairs $i, j$ where their separation is larger than twice their diameter. Table 5.3 summarises the results of such investigations for various sets

| Run | Site | Number of subapertures | Number of subaperture pairs | Redundancy Factor | Effective number of independent $r_0$ estimators |
|---|---|---|---|---|---|
| A | La Palma | 86 | 3283 | 129.2 | 25.4 |
| B | La Palma | 86 | 3126 | 104.0 | 30.1 |
| C | UKIRT | 25 | 289 | 14.2 | 20.1 |
| D | UKIRT | 25 | 289 | 11.6 | 24.9 |
| E | UKIRT | 25 | 289 | 14.8 | 19.5 |
| K | Calar Alto | 42 | 745 | 31.2 | 23.9 |
| L | Calar Alto | 42 | 783 | 37.2 | 21.1 |
| M | Calar Alto | 42 | 748 | 33.8 | 22.1 |

**Table 5.3**: Summary of the spatial correlation factors found for various sets of data.

of data. The effective number of independent measurements is found to be smaller than the number of subaperture images in each case.

## 5.7 $\mathbf{r_0}$ Variation over the Aperture

Over an entire data run, a value of $r_0$ is obtained for each of the pairs of subaperture images. It is of interest to examine the spread of the $r_0$ values so obtained and also to observe the dependence of $r_0$ on subaperture separation in the pupil in order to test the stationarity property and the power laws predicted by Kolmogorov theory. The consequence of stationarity in this case is that $r_0$ should be independent of subaperture pair location. The distribution of $r_0$ estimates has been investigated for a representative run taken at each observing site. The distributions are shown in the form of histograms in in Figs. 5.11, 5.12 and 5.13.

The standard error on the differential variance calculated for a particular pair is found from Eq. A.2 ($\chi$ represents the *square* of the differential image measurements in this case). A value of $n$ is used which is equal to the effective number of independent frames (which is the actual number of frames recorded during the run divided by the temporal redundancy factor, $\mathcal{F}_t$, calculated earlier). The standard error on the estimate of $r_0$ (from a single pair) can be calculated using the error propagation formula (Eq. A.13;

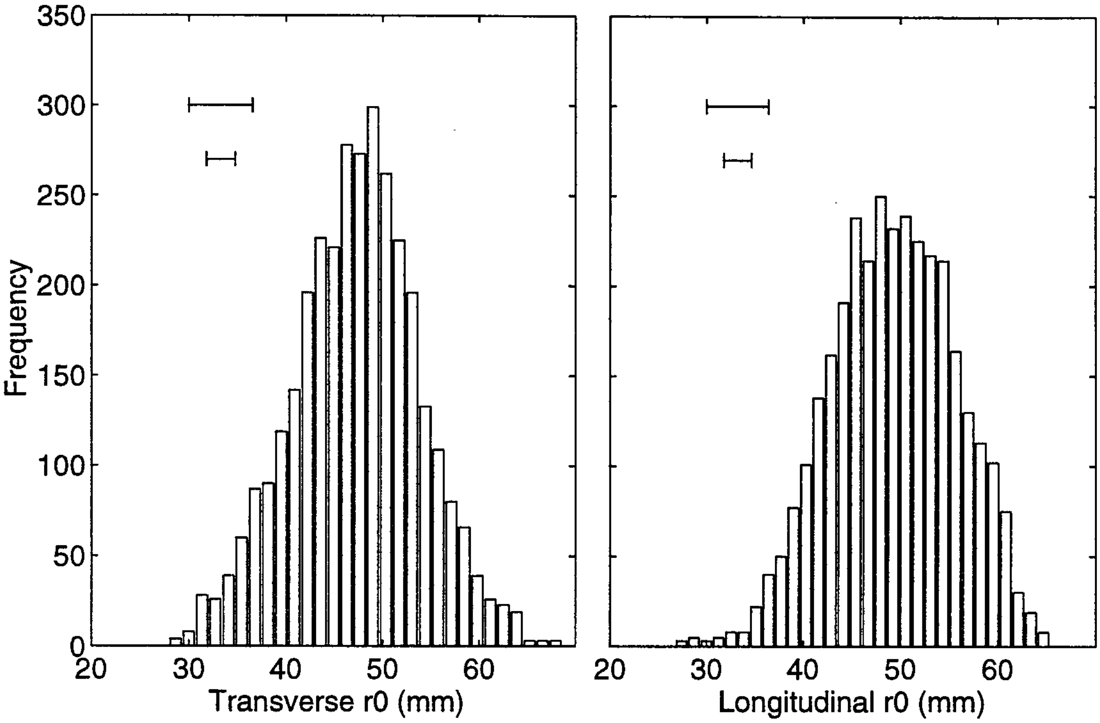



**Figure 5.11**: Histograms of $r_0$ calculated for all pairs of subapertures in the pupil (set A: La Palma). The top line is the standard deviation of this distribution of mean $r_0$ values. The lower line is the mean of the standard errors calculated from individual measurements of $r_0$ (made by individual subaperture pairs). The relatively small values of $r_0$ should be noted (see section 5.8).

with $m = 0.6$). A very important question is this: does the variation in estimated $r_0$ over the aperture merely represent a statistical variation (*i.e.* random errors due to the measurement inaccuracies only)? The histogram is a distribution of estimated $r_0$ values, each found from a single subaperture pair. Each of these $r_0$ estimates may be considered to have an accuracy equal to one standard error (calculated as described above). Under the assumption of stationarity, the typical error on a single $r_0$ estimate (*i.e.* from one differential subaperture pair) ought to be equivalent to the standard deviation of the distribution of $r_0$ estimates taken from many pairs (*i.e.* the distribution shown in the histograms). Two lines are drawn on each of the histogram plots. The length of the lower line is equal to the mean of the standard errors calculated for each and every pair. The upper line is equal in length to the standard deviation of the $r_0$ distribution. In each case, the upper line is approximately twice the length of the lower line, indicating that there is a significant variation in $r_0$ values across the aperture. It should also be noted that any variations in $r_0$ with time (which affect the entire aperture) would lead to a lengthening of the lower line but would not be expected to change the length of the upper line.

It is of further interest to compare the errors observed in practice with theoretical predictions. Because $r_0$ values are calculated from a variance of a measured quantity,

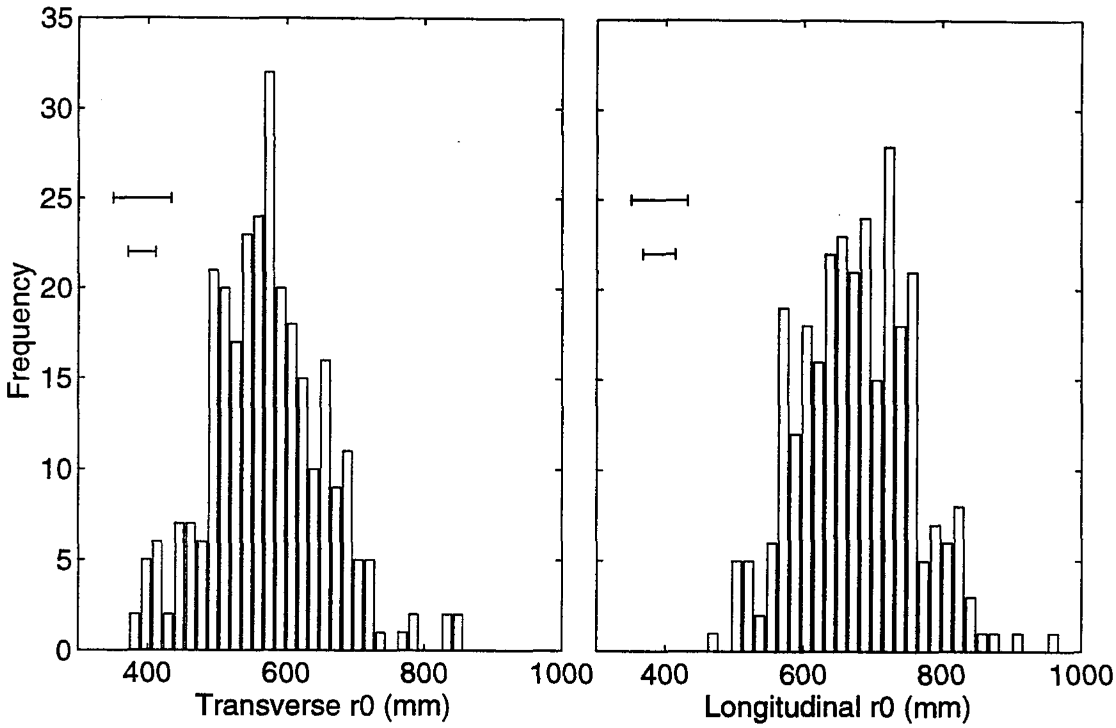



**Figure 5.12**: Histograms of $r_0$ calculated for all pairs of subapertures in the pupil (set C: UKIRT). For further explanation see Fig. 5.11.

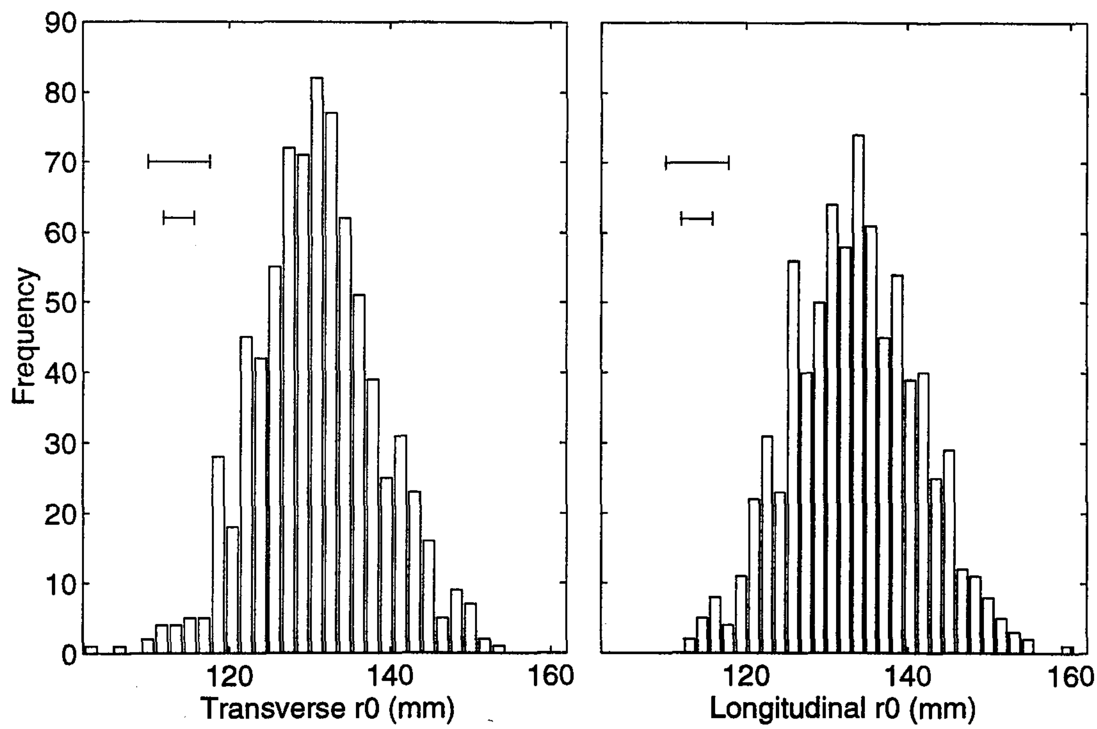



**Figure 5.13**: Histograms of $r_0$ calculated for all pairs of subapertures in the pupil (set L: Calar Alto). For further explanation see Fig. 5.11.

it is possible to express the error on the estimate in terms of the expected value of the variance being measured and the number of independent measurements only. Equation A.4 gives the expected error in the differential variance estimate by one pair due to the finite sample size (*i.e.* finite number of frames). This can be considered to be the best accuracy obtainable from a size of sample $n$, neglecting the effects which cause errors on the centroid measurement, such as photon and readout noise. The mean standard error (the lower line in the histograms) is compared with the expected error in table 5.4. For the Calar Alto and UKIRT data, the error calculated from experiments is of a similar size as (but not smaller than) the theoretically predicted error. However, the JKT data yields a much larger error than predicted; this may be attributed to variations in $r_0$ during the observation period, which would lead to a greater spread in the measurements obtained (see section 5.8, in particular Fig. 5.17).

| Data set | site | measured fractional error in $r_0$ | predicted fractional error in $r_0$ |
|---|---|---|---|
| A | La Palma | 0.06 | 0.03 |
| C | UKIRT | 0.08 | 0.06 |
| L | Calar Alto | 0.03 | 0.03 |

**Table 5.4**: Errors estimated for $r_0$ measurements taken with individual subaperture pairs compared with the expected values.

The data have been investigated further in order to examine a possible source of these errors, namely any significant correlation between subaperture separation and measured $r_0$ value. For each run, the complete set of subaperture pairs is divided into subsets, each containing pairs with approximately equal separations. Corresponding subsets of $r_0$ values (one value for each pair) are thus defined. A mean and standard error is calculated using the individual $r_0$ values from all pairs in each subset. Care has been taken to calculate the standard error correctly by estimating the effective number of independent measurements involved in each subset (see the discussion on spatial correlation given in section 5.6). The procedure is as follows: firstly the number of actual subaperture images involved in each subset is calculated (this will be equal to or less than the number of subapertures in the pupil). However, this number cannot be considered to be equal to the number of independent measurements, since this may be larger than the effective number of independent measurements found for all the pairs in the pupil (*i.e.* the value shown in the sixth column in table 5.3. This number is, rather arbitrarily, multiplied by a factor which is the ratio of the effective number of independent measurements when considering all possible pairs together (the sixth column in table 5.3) and the total number of subapertures (the third column in the same table). This factor represents the degree by which the measurements taken involving one subaperture can be considered independent of those taken by all the others, and is an attempt to account for the spatial correlation between measurements taken by subaperture pairs within an individual subset. It should also be recognised that there will be a correlation between the measurements in different subsets, since many subsets will rely on angle-of-arrival measurements made by a particular subaperture pair.

The mean $r_0$ is plotted against subaperture separation for the three sets in question in Figs. 5.14, 5.15 and 5.16, where the error bars represent one standard error either side of the mean. The graphs show a significant variation of $r_0$ with aperture separation, including some clear trends. The UKIRT data also displays a large difference between the $r_0$ values calculated using the transverse and longitudinal differential variances. This will be discussed in section 5.8. The larger variation in the seeing for this data set must be considered alongside the fact that the UKIRT pupil is considerably larger than that of the other telescopes, so any variations might be expected to be more marked.

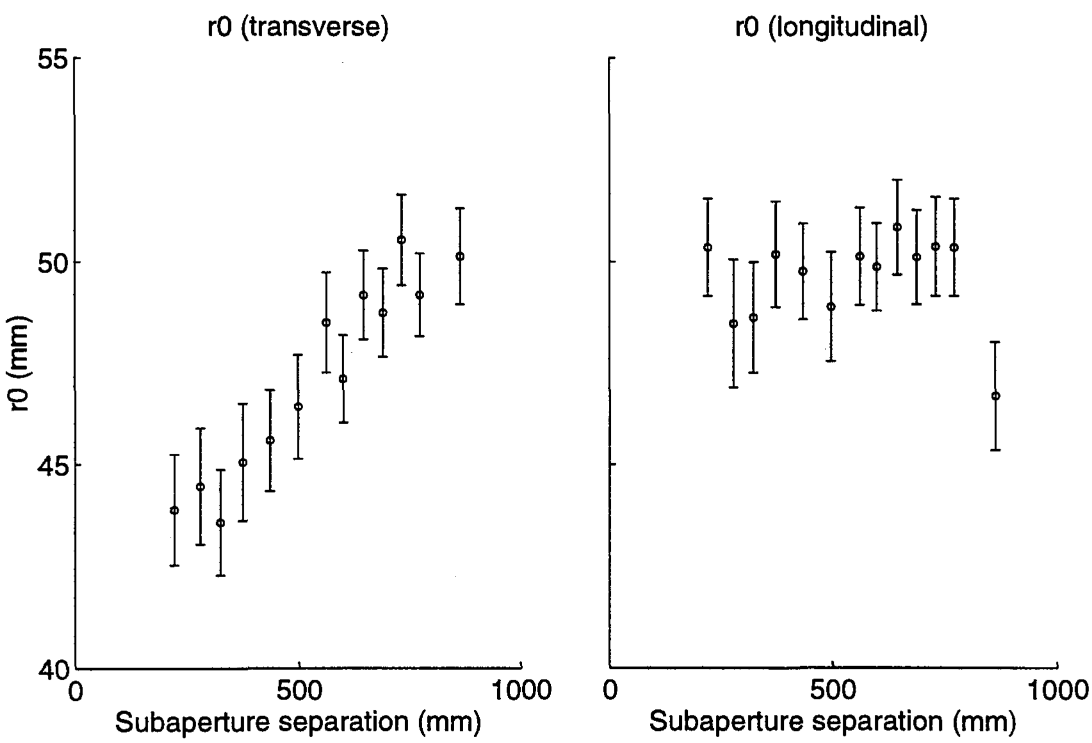



**Figure 5.14**: $r_0$ plotted against subaperture separation (set A: JKT); each point represents the mean $r_0$ value from a group of subaperture pairs with similar separations. The error bars show the standard error calculated from the $r_0$ estimates of all the pairs in the group.

Other data sets have shown similar trends. There appears to be a contribution from turbulence which is violating the Kolmogorov theory in at least one of a number of ways. Firstly, there may be a failure of the theory of stationarity, *i.e.* the structure function of phase may actually vary from one location to another, although this is unlikely to manifest itself here, because each subset of subaperture pairs involves individual subapertures from all parts of the pupil. Secondly, the Kolmogorov power laws may

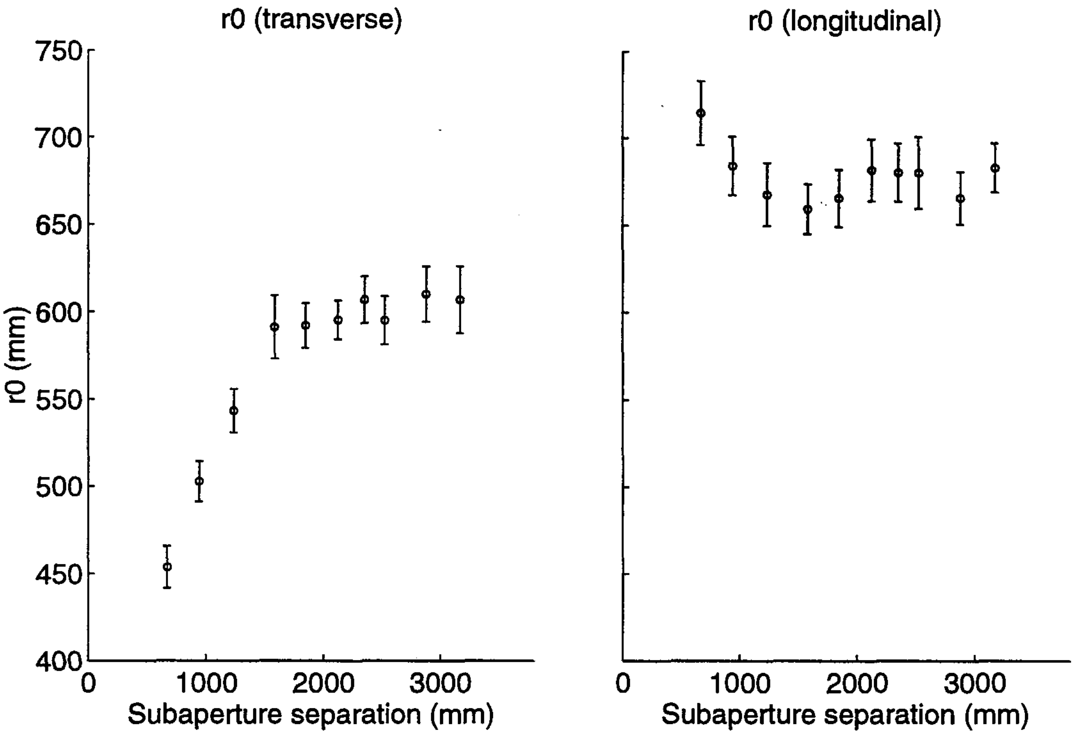



**Figure 5.15**: $r_0$ plotted against subaperture separation (set C: UKIRT); for further explanation see Fig. 5.14).

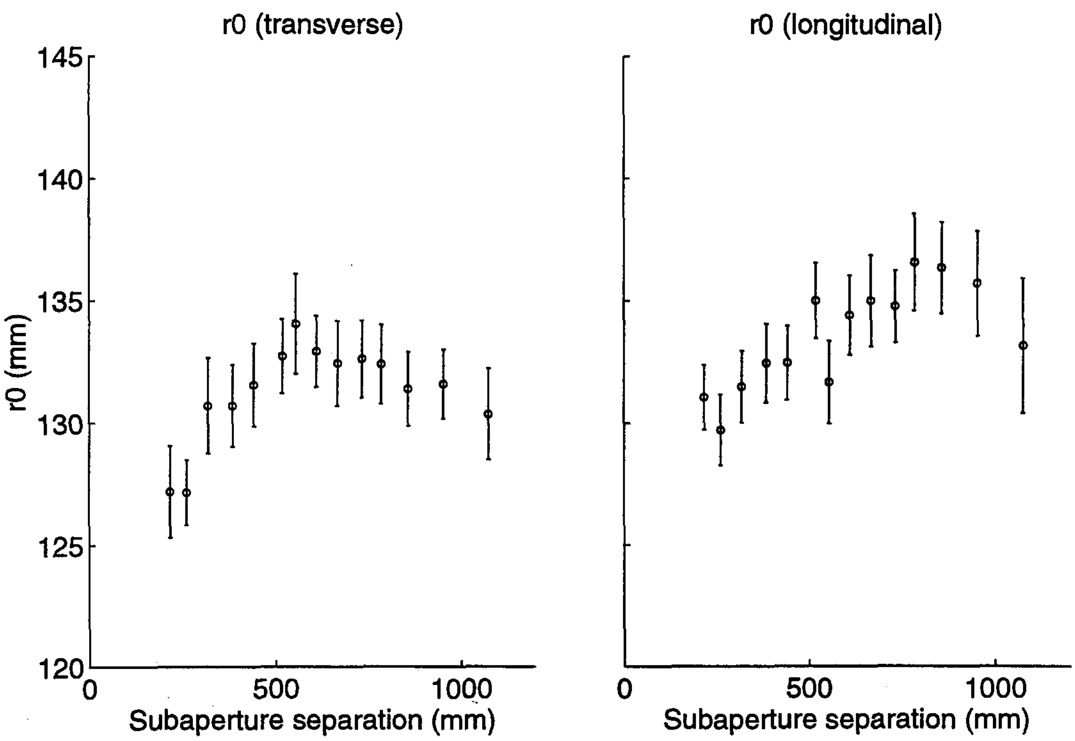



**Figure 5.16**: $r_0$ plotted against subaperture separation (set L: Calar Alto); for further explanation see Fig. 5.14).

not be valid. In this case, the structure function may not follow exactly a 5/3 power law (see Chapter 6; in particular Fig. 6.3). This non-Kolmogorov turbulence may be local turbulence which is affecting different parts of the aperture in different ways.

Some of the subapertures may possibly be partially obscured by the pupil edge or the central obscuration; it was initially thought that the tailing off of the longitudinal $r_0$ at the largest separations might be explained by such subapertures having an artificially small size. The subsets with the largest mean subaperture separations would contain more pairs from which one or both of the subapertures would be at the pupil edge. Furthermore, the primary mirror may contain some curvature, or 'roll off', at its perimeter. Images formed from subapertures at the pupil boundary may consequently wander further than expected, resulting in a higher variance and an $r_0$ estimate which is biased low. However, a further investigation, in which subaperture images from these regions were ignored, showed that the trends occur regardless.

## 5.8 $r_0$ Evolution Over Time

An essential part of any seeing monitor is to examine the evolution of $r_0$ over time. If Kolmogorov theory can be assumed, a value of $r_0$ determines all the spatial wavefront statistics, such as the variance of each of the individual Zernike modes. It is of great interest to have a knowledge of how $r_0$ varies over all timescales; from night to night and also from moment to moment within a particular night. An adaptive compensation system would be optimised to function according to the conditions occurring at the time of observation. Statistical information about the probability of observing a particular resolution at any one time would allow the performance of a proposed system to be evaluated. Also a knowledge of whether $r_0$ varies significantly within the time period of a typical astronomical exposure would determine whether any optimisation is possible.

$r_0$ itself is a statistical quantity, that is, it must be taken as an average over 'all time' and 'all space'. Since seeing conditions are known to vary over time, some compromise must be sought between the necessity to sample a sufficient number of values (particularly to avoid the biasing effects described in 3.7.5) and the requirement to sample regularly enough to keep up with any actual changes in the conditions. A choice must be made of the period of time over which a single $r_0$ estimate will be obtained. Of course, for each block of frames defined in this way, a large number of pairs of subapertures will each be used to make an estimate of $r_0$; the mean of these $r_0$ values will be taken as the estimate of $r_0$ for that block. Each run consists of a number of consecutive blocks, so that a time series of $r_0$ estimates is built up. The purpose of these measurements is to observe the variations in $r_0$ over different timescales. In order to infer the significance of any such variations, it is important to perform a proper error

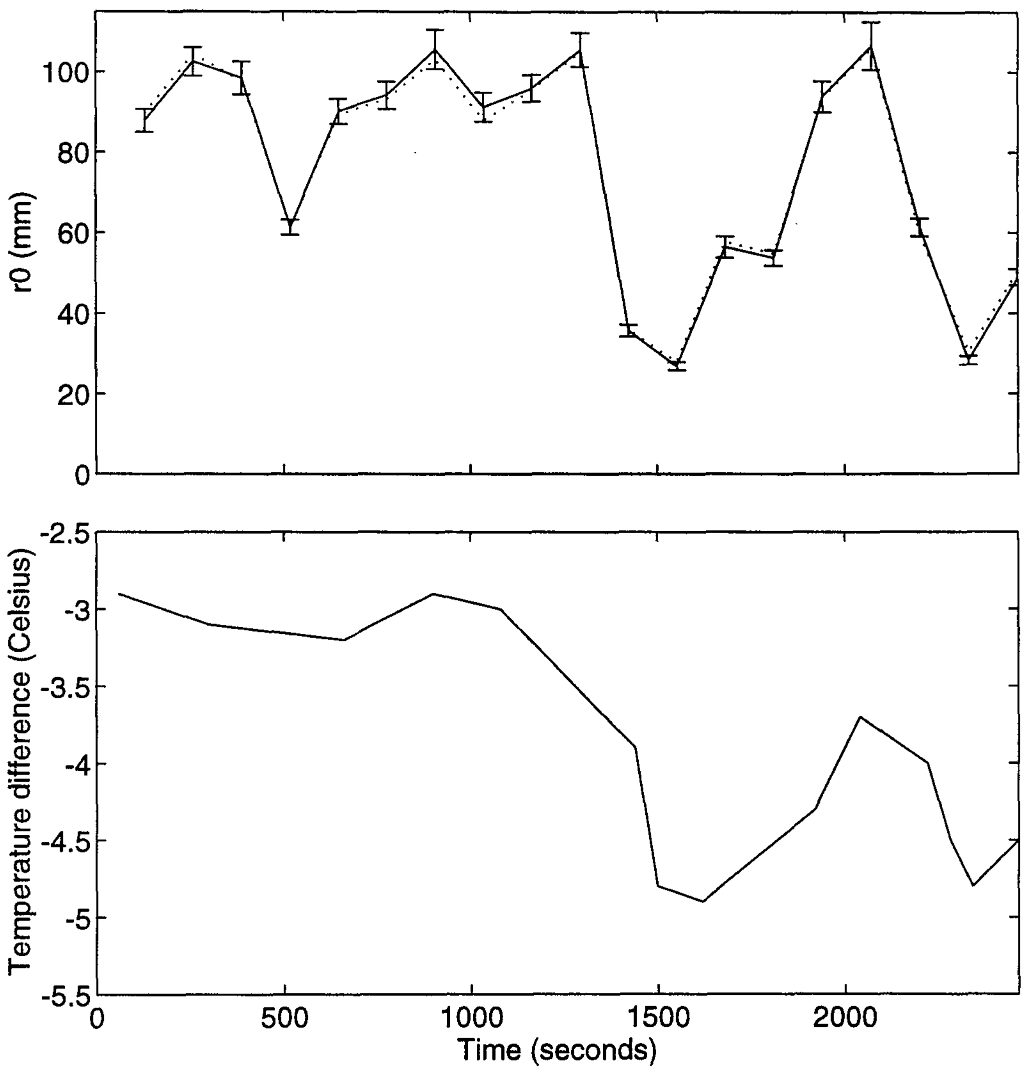



**Figure 5.17**: $r_0$ and temperature plotted against time for set A (La Palma). $r_0$ calculated from the transverse differential variance is plotted as the solid line, with errorbars; $r_0$ calculated from the longitudinal differential variance is plotted as the dotted line (error bars are not shown, but are similar in size to the transverse case). The temperature shown is the difference between that recorded outside the dome and that recorded near the mirror.

analysis.

The combined Shack-Hartmann and image frames acquired with the Princeton CCD were recorded at approximately two second intervals, so it was decided to take runs of 1000 frames (approximately 33 minutes of data). The size of the block used to make each estimate was chosen to be 50 frames; according to Fig. 3.11, the resulting bias in $r_0$ should be less than 3%. The UKIRT data was recorded in runs of 3000 frames or more. A block size of 200 was chosen in this case to provide a sufficiently large number of blocks within a run to observe the time evolution of $r_0$ satisfactorily. However, when one considers that the redundancy factor caused by the temporal correlation for set C is approximately 20, it is apparent that each subaperture pair effectively only makes 10

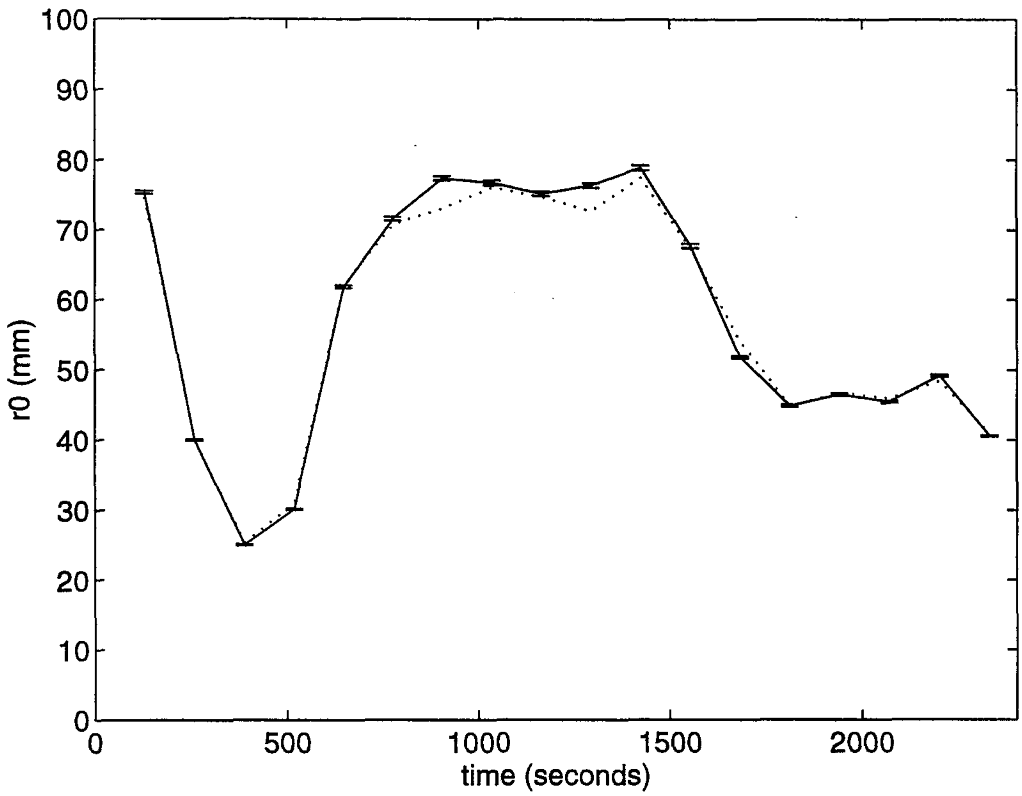



**Figure 5.18**: $r_0$ plotted against time for run B (see Fig. .5.17 for an explanation of the markers).

or so independent estimates of the differential variance per block of 200 frames. With reference to Fig. 3.11, this should lead to a bias in the $r_0$ estimates of approximately 20% (biased high). Due to the nature of the data acquisition at Calar Alto, the turbulence has been measured for, typically, 5 seconds, with a gap of a minute between each of these blocks. For this reason, and because the length of a single observation was quite large, typically 20000 frames, it has been decided to make one $r_0$ estimate per recorded block (normally 2040 frames; this is large enough to avoid any significant biasing effects).

Graphs of $r_0$ plotted against time are shown for samples taken at each of the sites in Figs. 5.17, 5.18, 5.19 and 5.20. In each case, the $r_0$ value calculated from the transverse method is plotted as the solid line. Error bars, which are displayed only for this estimate, have been calculated using the distribution of $r_0$ estimates from all the subaperture pairs. These errors are calculated from the standard error of this distribution, taking into account the redundancy due to the spatial correlation between one subaperture pair and another (see section 5.6; the number of effective $r_0$ estimators in the aperture is given in the sixth column of table 5.3). The $r_0$ values calculated from the longitudinal differential variance are shown as the dotted lines. Since only a small number of nights were available for seeing measurements, the actual values of the seeing obtained cannot be taken as representative. However, properties such as the variation in $r_0$ over long and short timescales are perhaps more instructive. The quality of the

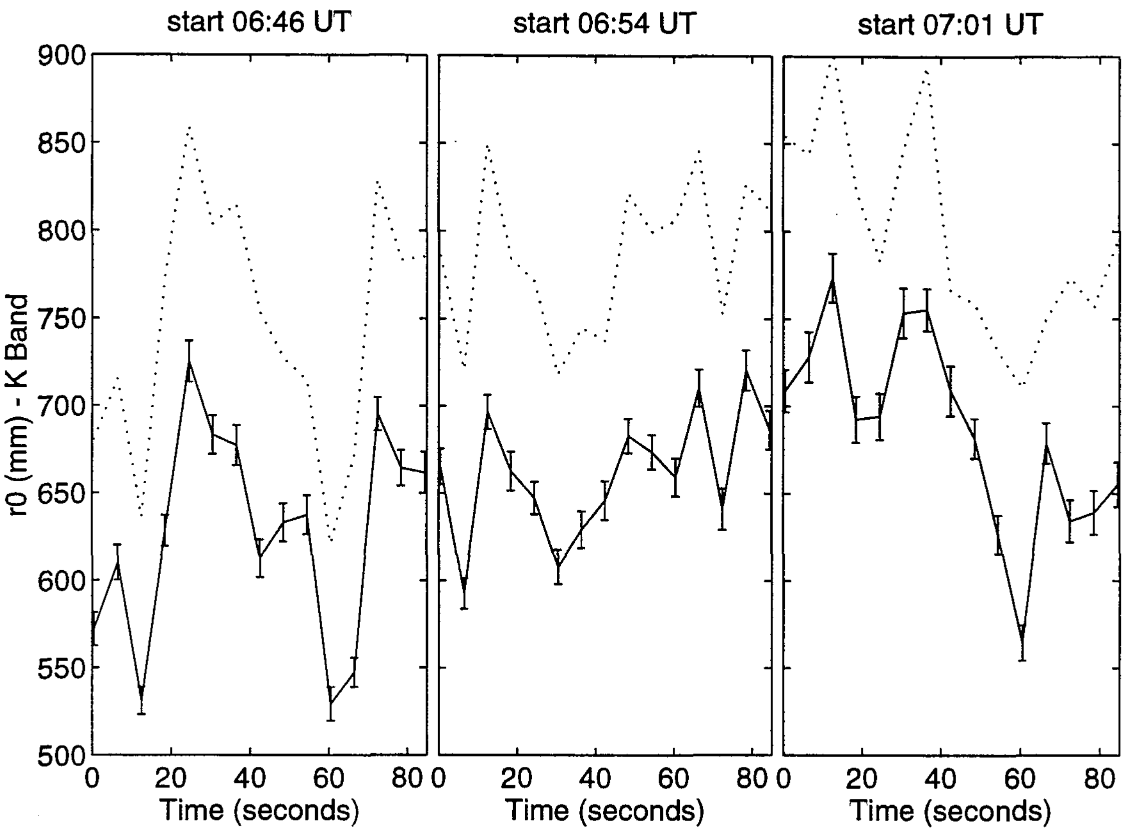



**Figure 5.19**: Sets C,D and E (UKIRT): Plot of $r_0$ against time from three consecutive runs of data, calculated from transverse differential variance (solid line) and longitudinal differential variance (dotted line). Error bars are not shown for the longitudinal case, but are approximately three times larger than the error bars in the transverse case.

seeing shown here can be compared with the measurements of seeing around the sky discussed in section 5.9.

Temperature measurements were recorded at regular intervals during the run of one data set (JKT: set A). The temperatures detected by a sensor outside the telescope dome and another placed close to the primary mirror are displayed as part of the user instrumentation at the telescope. Figure 5.17 shows the evolution of $r_0$ and the temperature difference between the mirror and the external surroundings plotted on the same time axis. The seeing drops by a factor of three within three minutes at one stage. The trend corresponds remarkably well with variations in the temperature differences between inside and outside the dome (which occur when the temperature of the atmosphere changes but the partially enclosed atmosphere in the dome cannot react quickly). One explanation for the correlation between seeing and temperature is that some artificial turbulence is being created as the temperature gradient between the dome contents and the outside atmosphere is upset. Curiously, the changes in temperature appear to follow the variations in $r_0$, as though the latter are driving the former. A slow response of the temperature sensors might cause this. A plot of $r_0$

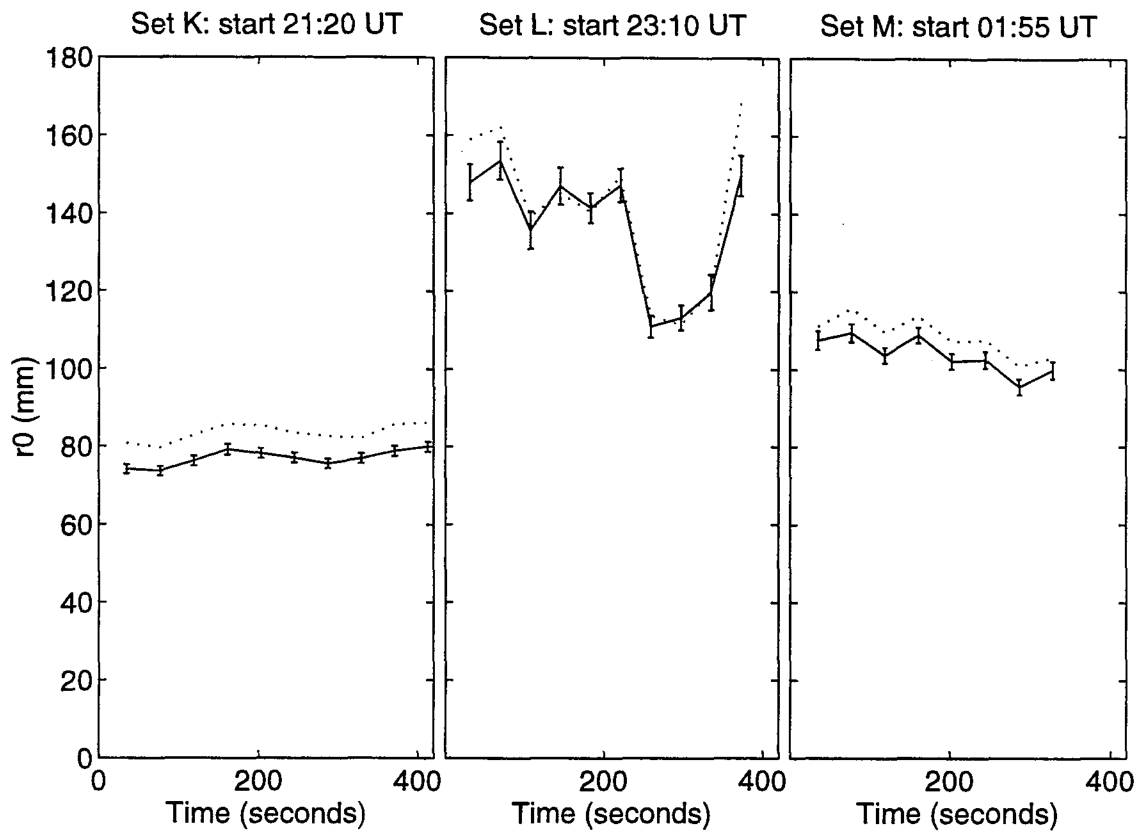



**Figure 5.20**: Sets K,L and M (Calar Alto): $r_0$ vs time measured at different times of the same night, calculated from transverse differential variance (solid line) and longitudinal differential variance (dotted line). Error bars are not shown for the longitudinal case, but are approximately three times larger than the error bars in the transverse case.

against time is shown for a second run (set B) taken on the same night in Fig. 5.18. Once again, rapid variations in $r_0$ are visible.

Some conclusions may be drawn from the small sample of data sets shown here. The seeing appears to be fairly stable at some times and undergoes large variations at others, as is demonstrated in the three sets from Calar Alto (Fig. 5.20). The larger error bars in set L probably reflect that $r_0$ is continuously varying, even within the period of a few frames. The seeing measured by sets K and M, in contrast, is relatively stable. Whether the changes are partially attributable to artificial effects associated with the telescope surroundings is unknown; only an independent seeing monitor could verify this. If the results from sets A and C are compared with the histograms of Figs. 5.11 and 5.12, there is an apparent discrepancy between the mean values of $r_0$ obtained (which is not observed in set L). This is because, if the seeing is varying, a single estimate of $r_0$ over a particular time period will give a different result from the mean value from multiple measurements over the same period, due to the non-linear relationship between centroid motion variance and $r_0$.

The two methods of calculating $r_0$ based on the differential centroid variance in the two transverse and longitudinal directions can sometimes give significantly different results (particularly in the data from UKIRT). This is clearly some sort of non-Kolmogorov effect; the analysis methods of Chapter 6 would possibly give more insight into such features. From the data sets shown, it is not unreasonable to expect to attain seeing values in the range of $10 - 20cm$ (at $0.55\mu m$). However, it is no coincidence that many of these measurements were taken during periods of good seeing, since the alignment of the system is assisted by good weather (see section 5.2.2), which is often associated with good seeing.

## 5.9 $r_0$ Measured for Different Sky Positions

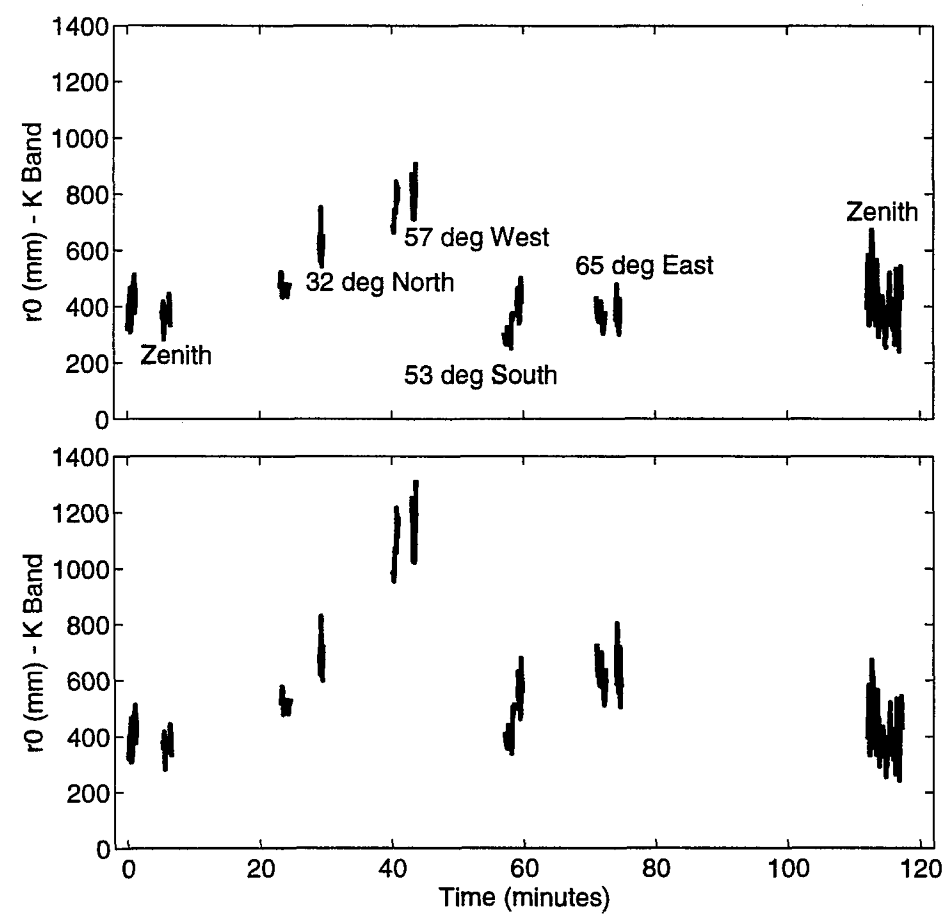



**Figure 5.21**: Seeing around the sky measured at UKIRT from 05:31 UT on 3rd August 1994. Angles are in degrees from the zenith (declination angle). Top: actual seeing measured; bottom: the same seeing compensated for the declination angle assuming Kolmogorov turbulence.

Experiments have been performed which attempt to test the expected dependence of seeing on declination angle (given in Eq. 2.60 for Kolmogorov turbulence). The purpose of these is to attempt to observe any changes in seeing, after altering the position of the reference star in the sky, over and above any natural variations in the seeing at

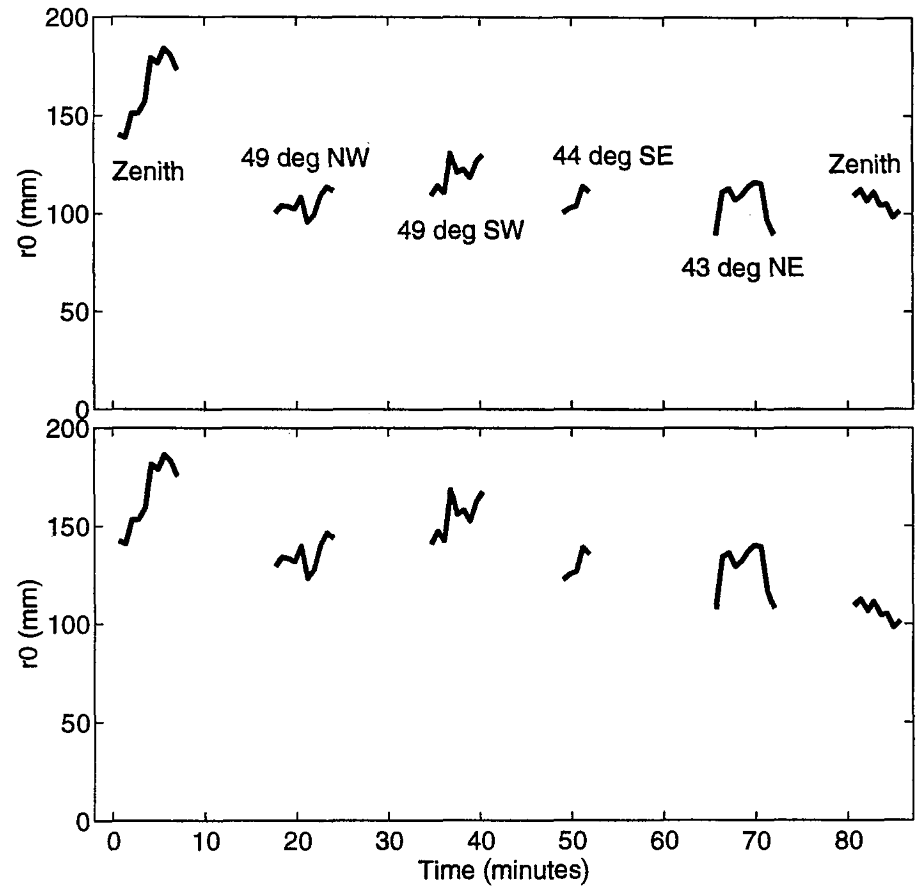



**Figure 5.22**: Seeing around the sky measured at Calar Alto from 00:34 UT on 13th July 1995. Angles are in degrees from the zenith (declination angle). Top: actual seeing measured; bottom: the same seeing compensated for the declination angle assuming Kolmogorov turbulence.

the zenith, and thus provide some verification of the Kolmogorov theory. Ideally, one would require the seeing to be measured simultaneously using a number of reference stars. However, this was not possible with the instrumentation available.

Two such experiments were performed, one at UKIRT and one at Calar Alto, using suitable point sources located in four quadrants of the sky. The stars were close to the declination limit of the telescope concerned in each of the directions. Since all the observations were taken over 120 and 90 minute time periods respectively, stars at the sky zenith were chosen at the beginning and end of the run to provide some sort of check on the variation of $r_0$ over time at the zenith.

At UKIRT, each of the quadrant stars was observed over 3500 frames in both 64x64 and 256x48 pixel readout mode. No significant difference was apparent between the values of $r_0$ obtained at the two exposure rates. At Calar Alto, the exposure time for all the sets involved was 3.9ms, with a corresponding frame rate of 260Hz. The observations involved are not listed in table 5.1. The $r_0$ time series are shown, without error bars, in Figs. 5.21 and 5.22. It is perhaps more interesting to compare the seeing which would be expected at the zenith (compensated for declination angle) as predicted by the Kolmogorov model; that is, the value of $r_0$ multiplied by $(\cos\theta)^{-\frac{3}{5}}$, where $\theta$ is the angle of declination. This information is also shown in the same figures.

It could be argued that the results from Calar Alto (Fig. 5.22) support the Kolmogorov model. The compensated plot appears to show a slow decrease in the compensated seeing over the 90 minutes of the run. The raw seeing appears to have been suppressed in the non-zenith observations. The UKIRT data (Fig. 5.21) shows the converse; if anything, the compensation process has increased the variation between observations, despite the seeing at the zenith being similar at the start and finish of the experiment. The variations are most likely due either to instabilities in the seeing itself, or to local effects which cause pockets of turbulence in different parts of the dome. The former could have been verified by returning to the zenith at regular intervals, but time was limited and this was not possible. Other explanations rely on a possible failure of the Kolmogorov model. Perhaps the power spectra leading to the three-fifths power law dependence of seeing on declination angle were incorrect at the time. Alternatively, the of the turbulent layers may not have been large enough in extent to give consistent seeing over large angles in the sky.

## 5.10 Centroid Power Spectra

When examining the temporal characteristics of the centroids, two considerations must be addressed; firstly, whether the exposure rate is sufficiently high in order to be able to observe any correlation (particularly the high frequency behaviour in the power spectra); and secondly, whether the observed behaviour, in terms of the slopes of the temporal power spectra, matches that predicted by the Kolmogorov theory and Taylor hypothesis. Only if these questions are true, can an attempt be made to reduce the statistics to manageable parameters (such as the Greenwood frequency). Clearly, if the expected slopes are not observed, possible explanations are that the sampling rate is not high enough, or that the combined predictions of the Kolmogorov model and the Taylor Hypothesis are invalid.

The centroid motions were first examined to see if any artifacts from the telecope tracking motors can be observed. Figure 5.23 is a plot of centroid position relative to its reference position (in units of arcseconds in the sky) in two orthogonal directions (the two axes of the CCD array) over a selected series of 1200 frames from the UKIRT data. Since UKIRT is an equatorially-mounted telescope, the detector can be positioned so that the orientation of the sky with respect to the detector does not change and it is possible to relate these two directions to the directions parallel and perpendicular to the motion of the star across the sky. Certain components of frequency can clearly

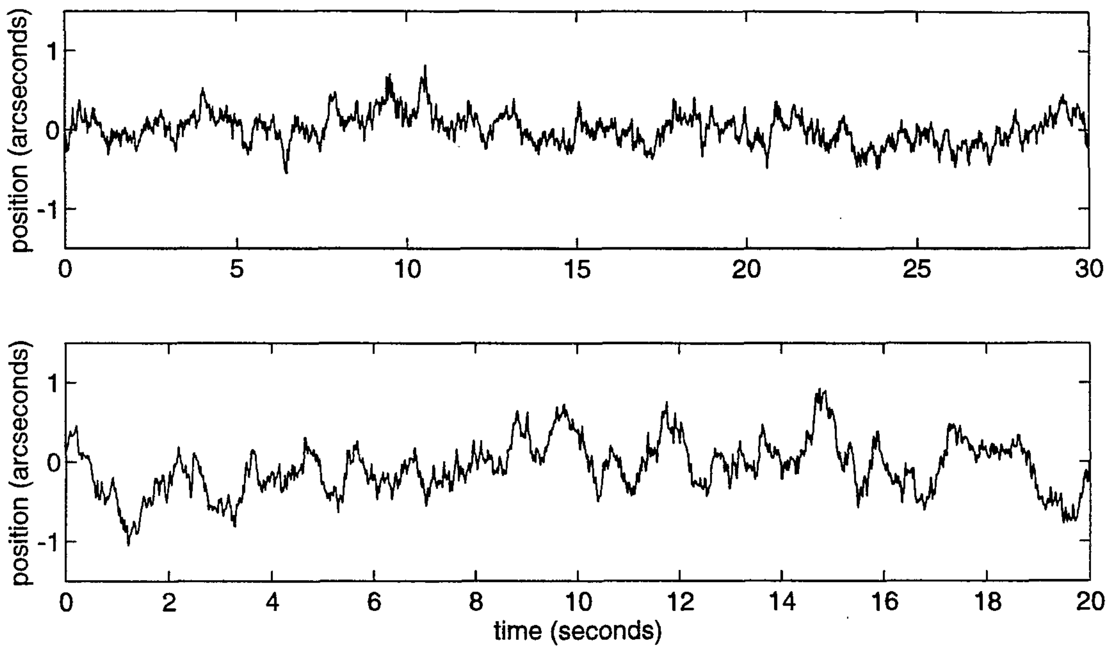



**Figure 5.23**: Centroid position in two orthogonal directions in the sky plotted against time (set E: UKIRT). Bottom: motion in the tracking direction (*ie* motion of star across the sky); top: motion orthogonal to tracking.

be seen in the direction of tracking which are not to be found in the perpendicular direction, particularly between 0.1 and 1 Hz. It should be noted that this set of frames has been deliberately selected to highlight the tracking errors. Nevertheless, this sort of behaviour is intermittently and frequently observed. These components are much more obvious in the plot of centroid power spectra shown in Fig. 5.24. They are almost certainly partly due to jitter in the tracking motors. For this reason, temporal power spectra of centroid motions are only considered here for motion orthogonal to tracking. The relationship between tracking direction and the direction on the detector was, unfortunately, not recorded for the data taken at Calar Alto, so in this case an arbitrary direction was chosen for measurement of the centroid power spectrum.

Such spectra are shown for four data sets acquired at UKIRT and three from Calar Alto plotted on a log-log scale (Figs. 5.25 and 5.27 respectively). These were all calculated from data which were recorded using stars close to the zenith over observations of different lengths (the longer the run, the larger is the statistical sample, but also the greater is the chance that the seeing conditions have varied significantly during the run). One would expect to observe a steeper curve as the frequency increases until the signal disappears into a 'sea' of noise, caused by inaccuracies in the centroids, which in turn result mainly from detector readout noise. In order to estimate the slope at the high frequency end, it is necessary to make a visual estimate of where this noise begins. Clearly, the noisy region of the spectrum is useless, so an attempt is made to fit a straight line to part of the line at frequencies below this selected cut-off frequency.

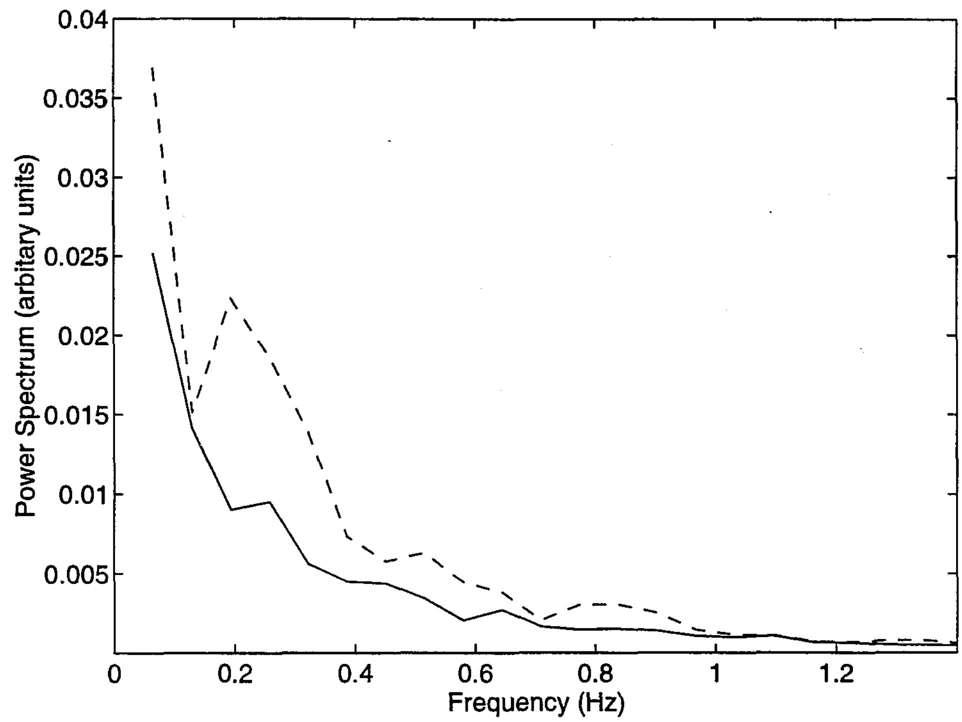



**Figure 5.24**: Temporal Power Spectrum (low frequency end) of centroid motions in the direction of tracking (dashed line) and in the orthogonal direction (solid line; set E: UKIRT). The graph is an average of the spectra of all 9 subaperture images in the 64x64 pixel mode divided into blocks of 512 frames (3600 frames used in total). The power is plotted on a linear axis to highlight the components caused by tracking motor jitter.

A similar fit is made at the low-frequency end of the spectrum. This process requires a certain amount of intuitive selection and may lead to the part of the line with the slope with the largest magnitude (*i.e.* the steepest part) being selected. This slope may indeed be affected if artificial components due to noise in the system, such as tracking errors or electronic noise, are present. It can, however, be concluded that in none of the sets shown can a slope be observed which is close to the expected value of $-11/3$ predicted by the Taylor hypothesis at the high end of the spectrum. Nor do the low-frequency slopes reach their expected values of $-2/3$ (although this part of the spectrum does appear closer to its predicted form than that the high frequency part).

In order to eliminate noise due to tracking errors, the temporal power spectra of differential image motions have been calculated for the same sets. From the data recorded at UKIRT, of the nine subapertures whose images are recorded in the $64 \times 64$ pixel readout mode, eight pairs have been identified whose separations in the pupil plane are equal. Similarly, around 40 pairs of subapertures with equal separation have been identified in each of the Calar Alto data sets. The centre-to-centre axes of the pairs all have a different orientation in the pupil plane; therefore all directional information is lost. By taking the mean of the spectra from all pairs, however, the signal to noise ratio should be increased. The difficulty here is that the variance (and thus the overall

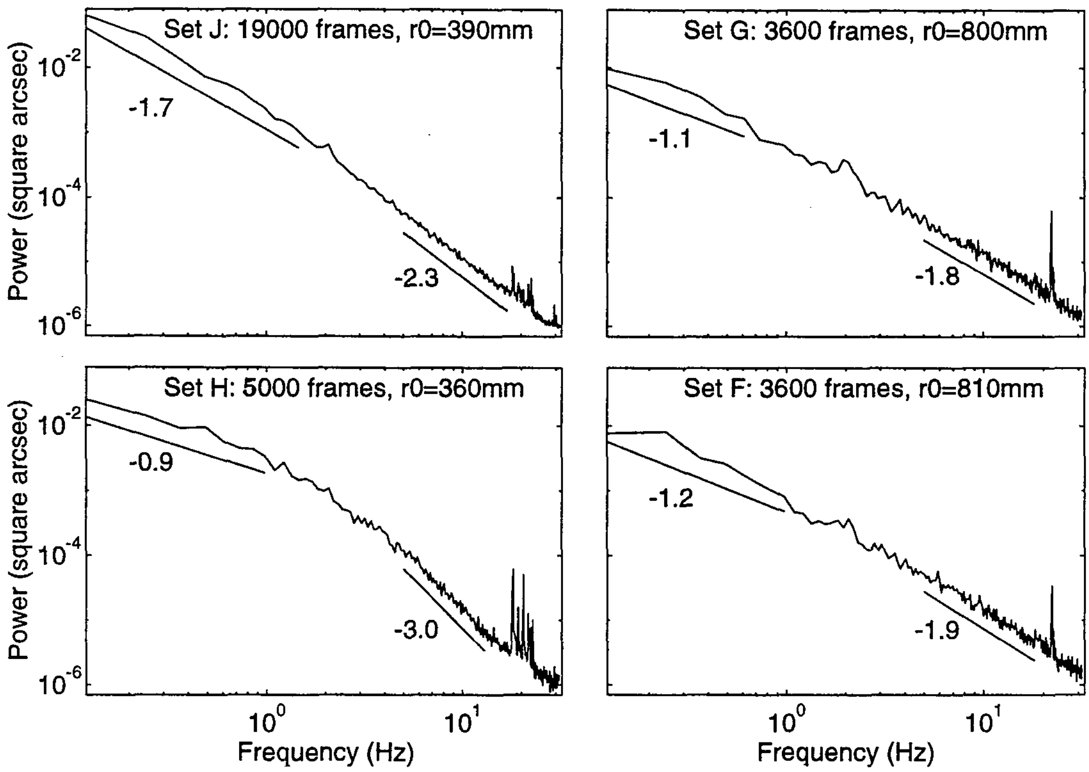



**Figure 5.25**: Temporal power spectra of centroid motions measured in the direction orthogonal to tracking for 4 different data sets acquired at UKIRT. The graph is an average of the spectra calculated from the time series of 9 subaperture images in the 64x64 pixel readout mode divided into blocks of 512 frames. The straight lines have the same slopes as lines of best fit. Their respective slopes are shown.

energy in the power spectrum) of the differential motion of a pair of subaperture images is dependent on their separation (see Fig. 3.4). For this reason, it is undesirable to use the mean of the spectra from pairs with different separations.

The power spectra of longitudinal centroid difference, calculated for the same seven data sets, are displayed in Figs. 5.26 and 5.28. The slopes at the two frequency extremes have been estimated in the same way as before. With reference to the data taken at UKIRT, the first point of interest is that several of the high frequency peaks observed in Fig. 5.25 have been removed. The slopes are more consistent between sets of data than in the case of the raw centroid motion spectra, but again there is no evidence of behaviour consistent with the Taylor hypothesis (under which one would expect slopes of $+4/3$ and $-11/3$ at the low and high frequency ends of the spectrum respectively). There is relatively little difference, in terms of the slopes, between raw centroid and differential centroid power spectra in the Calar Alto data, apart from in the very low frequency regime, where, presumably, some of the tracking errors have been removed.

The most important conclusion from the temporal behaviour examined here is that there appears to be no evidence for any of the behaviour predicted by the Taylor hy-

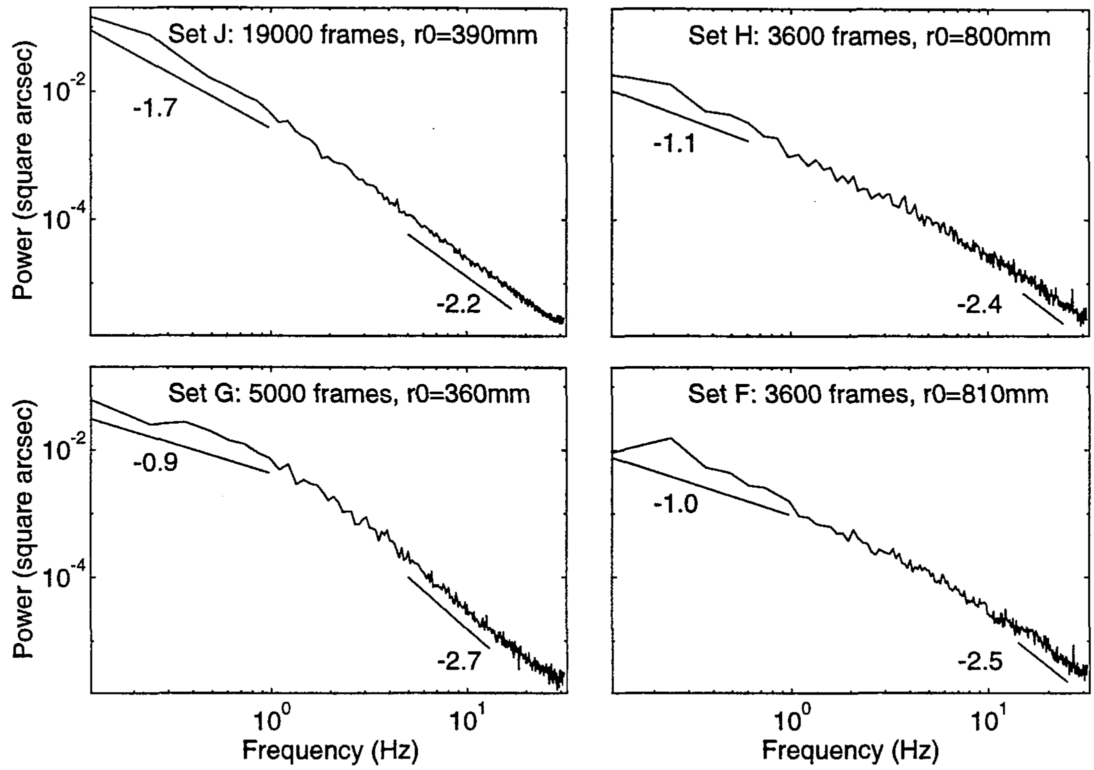



**Figure 5.26**: Temporal power spectra of differential centroid motions for the 4 sets of data shown in Fig. 5.25.

pothesis. Of course, if the sampling frequencies involved are below the cut-off frequency of $\nu_c = 0.3v_x/D$, as defined in section 3.6.5, then the expected high frequency behaviour will not be observed. For the UKIRT data, the highest measurable frequency is approximately 32Hz and the subaperture size is 31cm, which corresponds to a windspeed of $33ms^{-1}$; if there is a turbulent layer of significant strength with a windspeed larger than this, then the the cut-off frequency will not be within the region of measurement, which, in this case, is quite possible. The windspeed which corresponds to the the highest frequency observable in set L ($195Hz$) is approximately $65ms^{-1}$, for a subaperture size of $10cm$. The presence of a layer with this velocity would appear to be less likely, but not impossible. In either event, the apparent failure of the Taylor Hypothesis makes the characterisation of the temporal properties of the atmosphere less straightforward and creates difficulties for the specification of an adaptive optics system. This suggests that a modal analysis of the temporal characteristics of the turbulence may be necessary since a characterisation in terms of parameters such as the Greenwood frequency is not possible.

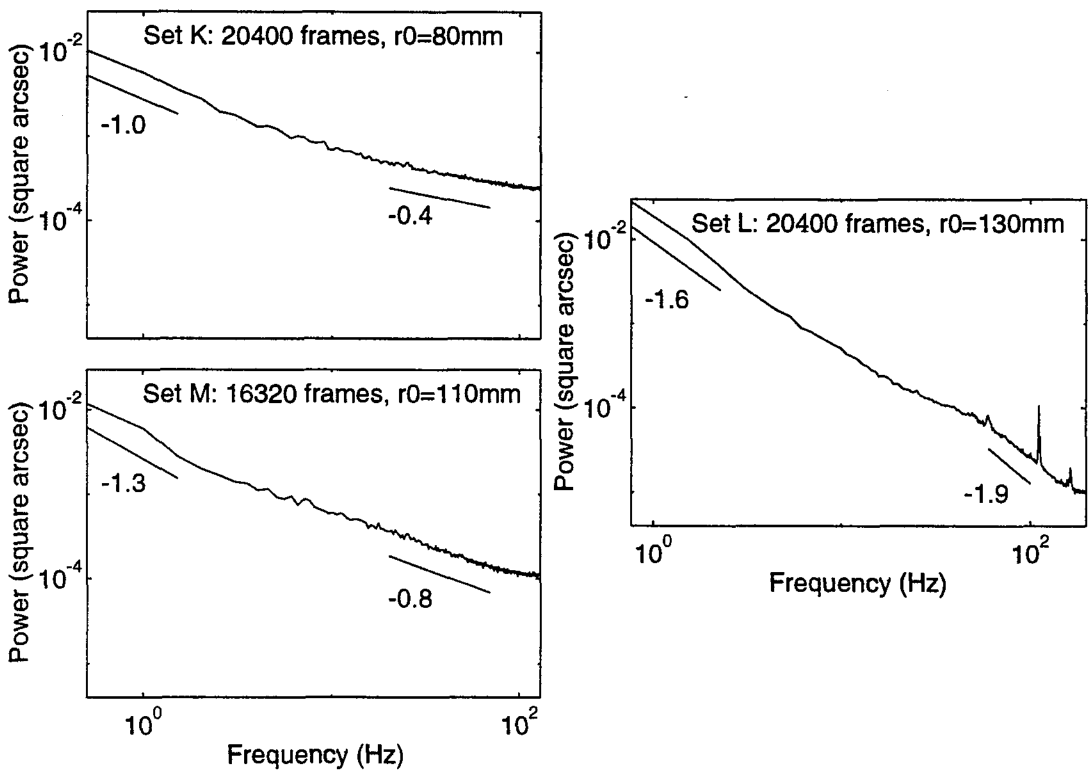



**Figure 5.27**: Temporal power spectra of centroid motions measured in an arbitrary direction for 3 different data sets acquired at Calar Alto. The graph is an average of the spectra calculated from the time series of 42 or 43 subaperture divided into blocks of 512 frames. For further explanation, see Fig. 5.25.

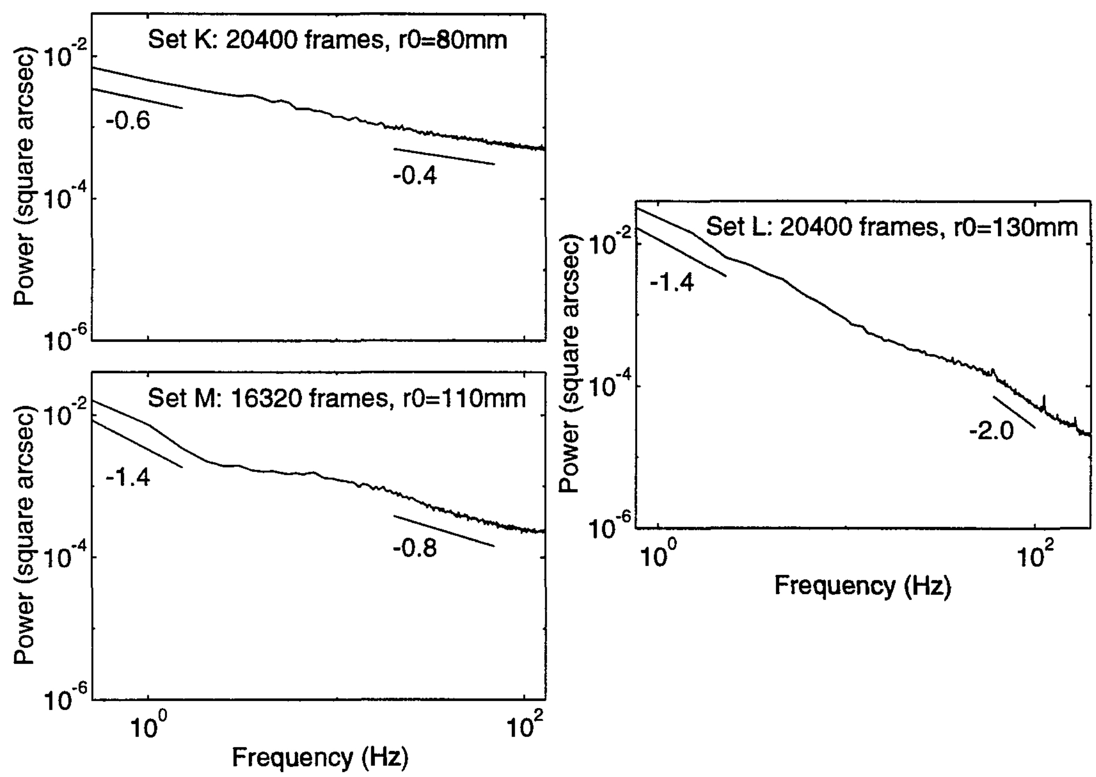



**Figure 5.28**: Temporal power spectra of differential centroid motions for the 3 sets of data shown in Fig. 5.27.

# Chapter 6

# Non-Kolmorogov Turbulence

## 6.1 Evidence for a non-Kolmogorov Structure Function

The weight of scientific evidence which, to some extent, and on some occasions, supports the predictions of the Kolmogorov model of optical propagation through turbulence has led some authors to accept the model without criticism. The model consists of a 5/3 power law, with the appropriate scaling parameter $r_0$, for the structure function of phase fluctuations formed in the telescope aperture. The associated outer scale is accepted to be so much larger than existing telescope apertures to be of little significance to ground-based astronomical imaging. The results of the measurements in Chapter 5, however, are observed to be rarely in complete agreement with the Kolmogorov model. The variation found in $r_0$ values with separation of subaperture pairs within the aperture (see Figs. 5.14, 5.15 and 5.16) is a case in point. Of course, any of these effects could be attributed to the local environment contained within the telescope dome. Nevertheless, a number of other authors have found results not in agreement with the model. Dayton *et al* [69] have made measurements of the phase structure function directly using interferometric methods. They have found good agreement with the 5/3 power law in most cases; however, on one occasion a power law of 1.25 was measured, which is similar to results found by Bester *et al* [70]. Buser *et al* [71] have found that the exponent in the structure function resulting from horizontal propagation experiments performed near to the ground is always less than 5/3. McKechnie has observed that the ratio between image motion and image blurring (interpreted as the ratio between the tilt terms and higher order modes in a modal description of the statistics) is lower than expected for telescope-sized apertures. This is attributed to the the presence of an outer scale which is much smaller than was previously thought plausible, but it could also be explained by deviations in the form of the structure function at scales smaller than the outer scale.

Gurvich and Belen'kii [72] have suggested that the Kolmogorov 11/3 power law for refractive index fluctuations (see Eq. 2.24) which has been used with some success to predict the form of the statistics of phase at ground level (*i.e.* a structure function with a 5/3 power law) may only be valid for tropospheric turbulence. Meteorological measurements in the stratosphere lead to the conclusion that the slope of the refractive index power spectrum is nearer to $-5$ than $-11/3$. The combination of the effects of these two atmospheric layers leads to a phase structure function of the form of $Ar^{\frac{5}{3}} + Br^3$. The wavelength dependence of the seeing produced by the two layers is also consequently different, the stratospheric turbulence being expected to have more influence at longer wavelengths.

## 6.2 Consequences of a non-Kolmogorov Power Law

The observed deviations from Kolmogorov behaviour have led to a more general prediction for the form of the phase spectrum,

$$\Phi_\varphi(f) = \frac{A_\beta f^{-\beta}}{r_0^{\beta-2}}. \tag{6.1}$$

The phase spectrum calculated for Kolmogorov turbulence (see Eq. 2.85), is a specific case with $\beta = 11/3$ and $A_{11/3} = 0.0229$. The phase structure function, defined in Eq. 2.7, can easily be shown to be proportional to $r^{\beta-2}$. Although, in its original form, the parameter $r_0$ has no meaning unless the statistics are Kolmogorov, an analogous quantity $\mathcal{R}_0$ can be defined so that the wavefront variance over an aperture of diameter $\mathcal{R}_0$ averages to one square radian,

$$D_\varphi(\mathbf{r}) = \gamma_\beta \left(\frac{r}{\mathcal{R}_0}\right)^{\beta-2}. \tag{6.2}$$

This definition of $\mathcal{R}_0$ is similar to but not equivalent to that of $r_0$ for the specific case of Kolmogorov turbulence. From the result given in Eq. 3.8, the mean-square phase fluctuations across apertures of diameters $r_0$ and $\mathcal{R}_0$ differ by only 3% in the Kolmogorov case. The parameter $\gamma_\beta$ is a calculable quantity. Thus, for this generalisation of the power law, the two parameters $\mathcal{R}_0$ and $\beta$ describe completely the second-order statistics of the wavefront. With the definition given, $\mathcal{R}_0$ is a direct measure of image quality. However, it is the distribution of the phase power between lower and higher spatial frequencies which has consequences for the design of an adaptive compensation system. The performance of such a system could be evaluated, for example, in terms of the variances of the Zernike modes. The distribution of turbulent phase energy over the individual Zernike modes is given by Noll for a value of $\beta$ of 11/3 (see section 3.2); however, any variation in $\beta$ can distort this distribution.

Boreman and Dainty [73] have recently applied the formulation of Eq. 6.2 to a modal wavefront expansion. The calculations are different for $\beta$ values between 2 and 4 and for values larger than 4; in the latter case, the outer scale must be taken into account. A number of conclusions were made:

- A value of $\beta$ equal to 2 corresponds to the structure function being zero. In this case, all the turbulent energy is contained in the piston term only, and there is no higher order energy.
- $\beta = 4$ corresponds to the case where all the piston-corrected energy is contained in the two tilt terms.
- As $\beta$ approaches 2 from above, the energy is shifted away from the tilt terms into the higher order terms. Thus, as $\beta$ decreases, any compensation system would have to correct more terms to achieve the same degree of compensation. The Zernike variances also peak at lower values of $\beta$ as the order increases, thus preserving the same total variance for a particular value of $D/\mathcal{R}_0$. A formula is provided for the variance of a particular Zernike term as a function of $\beta$.
- For $\beta$ values greater than 4, a Von Karman spectrum [9] (which is a mathematical model which accounts for the presence of an outer scale) is used. The inference is that for an infinite outer scale, all of the energy lies in the tilt terms $\beta \geq 4$. However, a decrease in the outer scale puts more of the energy into the higher order modes. This shift of energy becomes more marked as $\beta \to 4$ from above.

## 6.3 Measurement of a non-Kolmogorov Power Law

The purpose here is to suggest a method of measuring $\beta$ using a Shack-Hartmann wavefront sensor and to use the results to analyse atmospheric data.

The analysis of section 3.6, in which the form of the statistics of the centroid motions was calculated for Kolmogorov turbulence, relied on the fact that the angle-of-arrival statistics may be directly related to the wavefront phase statistics as long as the form of the structure function is known. The analysis methods will now be extended to non-Kolmogorov structure functions.

In section 3.6.2, the covariance of angle-of-arrival fluctuations was related to the structure function of phase via Eq. 3.22, which will be repeated here:

$$C_\alpha(\mu, \eta) = -\frac{\lambda^2}{8\pi^2}\frac{\partial^2}{\partial\mu^2} D_\varphi(\mu, \eta). \tag{6.3}$$

The non-Kolmogorov phase structure function of Eq. 6.2 can be written as

$$D_{\varphi}(\mu, \eta) = \gamma_{\beta} \mathcal{R}_0^{-\beta+2} \left(\mu^2 + \eta^2\right)^{\frac{\beta-2}{2}}. \tag{6.4}$$

Following the same reasoning for the Kolmogorov statistics, by substituting Eq. 6.4 into Eq. 6.3, the generalised results for longitudinal and transverse covariance are:

$$C_{\alpha}(d, 0) = 0.0127 \gamma_{\beta} \lambda^2 \mathcal{R}_0^{-\beta+2} (\beta - 2)(\beta - 3) d^{\beta-4}. \tag{6.5}$$

$$C_{\alpha}(0, d) = 0.0127 \gamma_{\beta} \lambda^2 \mathcal{R}_0^{-\beta+2} (\beta - 2) d^{\beta-4}. \tag{6.6}$$

Hence the ratio of these two quantities is given by:

$$\frac{C_{\alpha}(d, 0)}{C_{\alpha}(0, d)} = \beta - 3. \tag{6.7}$$

This analysis produces a very elegant result; however for $\beta$ values much less than 4, aperture averaging effects become significant even for large subaperture separations. Furthermore, as already discussed, covariance measurements are extremely sensitive to telescope tracking errors and a differential method is preferable.

The differential image motion has been used to measure the statistics of the wavefront (see section 3.6.4). In this case, aperture averaging effects must be taken into account and for this reason it is very difficult to make an analytic calculation. Equations 3.50 and 3.51 are an attempt by Sarazin and Roddier to provide an semi-analytical formula for the Kolmogorov case, but Fried's result (Eq. 3.52) is more useful because it can be applied to any form of structure function and relative position of (circular) apertures.

The non-Kolmogorov structure functions introduced earlier in this chapter can now be applied to Eq. 3.52. The transverse and longitudinal differential angle-of-arrival variances ($\mathcal{D}_{\bar{\alpha}_{\perp}}$ and $\mathcal{D}_{\bar{\alpha}_{\parallel}}$ respectively) are found by using values of 0 and $\frac{\pi}{2}$ for $\psi$. It is useful to obtain the ratio between these two quantities, since several factors cancel and the result depends only on $\beta$ and $S$. In particular, the ratio does not depend on the actual value of the seeing at the time (or the factor $\gamma_{\beta}$).

Since there is no analytic method of integrating Eq. 3.52, a numerical integration has had to be performed for each of a range of $\beta$ values and for each of a range of subaperture separations $S$. This was done using the mathematical package *MATLAB*. Predicted curves of $\mathcal{D}_{\bar{\alpha}_{\perp}} / \mathcal{D}_{\bar{\alpha}_{\parallel}}$ against $S$ for a number of $\beta$ values are shown in Fig. 6.1. Values calculated from Sarazin and Roddier's formulae for Kolmogorov turbulence ($\beta = 11/3$, dotted line) are also shown (presented in a similar fashion as in Fig. 3.2).

The differential variance ratio has been calculated for centroids obtained from 1000 simulated Kolmogorov phase screens. An algorithm was used which simulates the

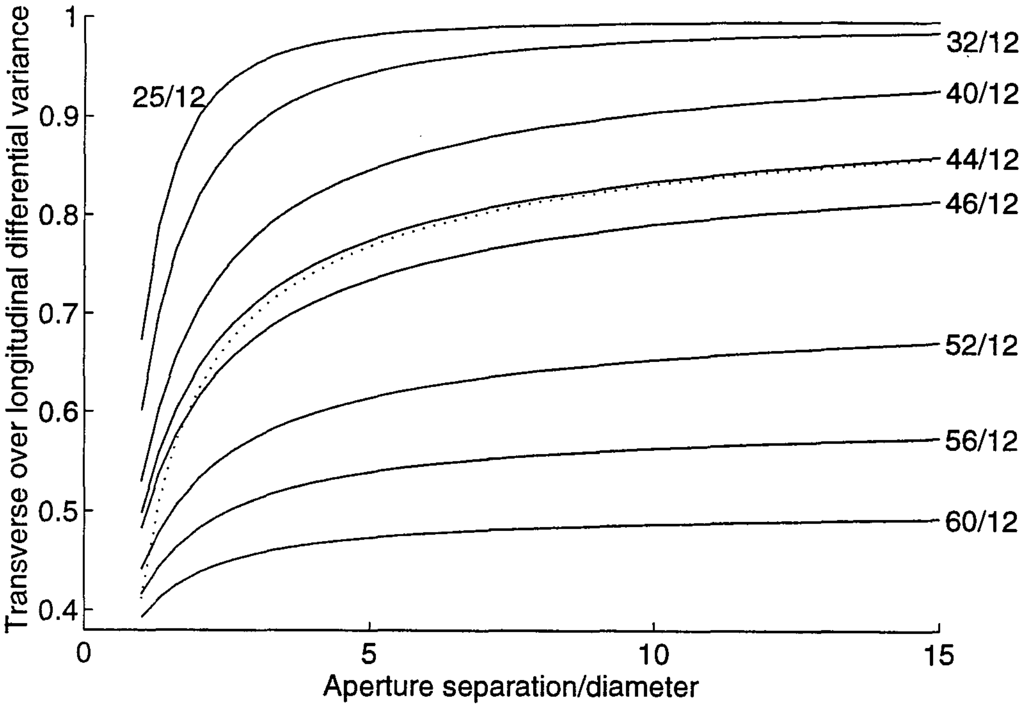



**Figure 6.1**: Theoretical dependence of the ratio of transverse to longitudinal differential angle-of-arrival variance on the exponent ($\beta$, as labelled - solid lines) of the power spectrum of phase, calculated from Fried's formula. Sarazin and Roddier's approximate formula is plotted as the dotted line.

behaviour of a Shack-Hartmann with square subapertures. Each phase screen is divided into square regions which represent the wavefronts sampled by each of the subapertures. The phase screens have been created by software written by P.Negrete-Regagnon [5] using a method which 'folds in' different frequencies. Each phase screen is required to be equivalent in size in its physical representation to the size of the telescope aperture. The phase is sampled across this region in a grid of, for example, $128 \times 128$ phase values. In the first instance a phase array, obeying Kolmogorov statistics, is created which is equivalent in physical size to the nominated outer scale. The sampling frequency, however, is very low, the number of pixels contained in the array being the same size as when reduced to the size of the telescope aperture (*e.g.* $128 \times 128$). The centre quarter of this array is then removed (the rest being thrown away) and interpolated to bring it back to the original size (in pixels). This process effectively reduces the physical size represented by the diameter of one pixel to a half. The higher frequencies, currently missing from the array, are then added in keeping with the Kolmogorov power law. This process is repeated until a phase array is obtained which is the size of the telescope aperture. The advantage of this method is that it does not underestimate the contribution of the lower spatial frequencies. The Shack-Hartmann simulator was also provided by P.Negrete-Regagnon. The software divides each screen into regions representing subapertures and simulates the imaging process at high light levels using

Fourier transforms. In this way, a set of centroid positions is obtained for each phase screen.

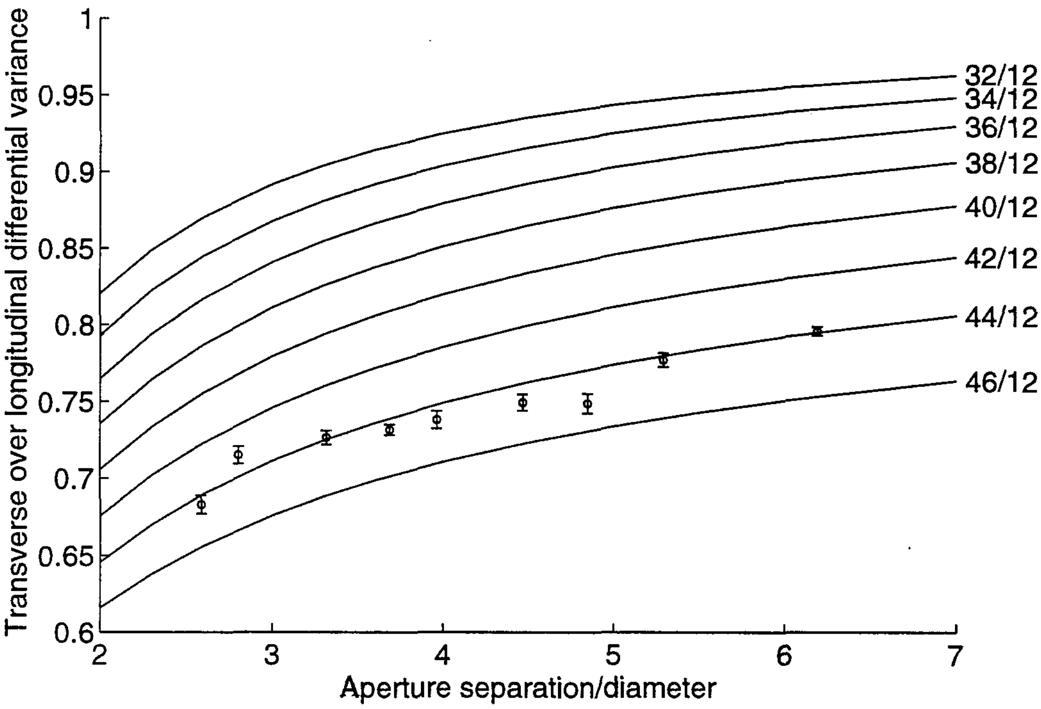



**Figure 6.2**: Comparison of the simulated centroid variance ratio with theoretical calculations.

The centroid differential variances $\mathcal{D}_{\bar{\alpha}_\perp}$ and $\mathcal{D}_{\bar{\alpha}_\parallel}$ have been calculated from the whole set of wavefronts for each pair of subapertures. For the purpose of visualising the data, subaperture pairs with similar separations are grouped together. The mean variance ratio, $\mathcal{D}_{\bar{\alpha}_\perp}/\mathcal{D}_{\bar{\alpha}_\parallel}$, and an associated standard error have been calculated for each group of subaperture pairs. The results are plotted in Fig. 6.2 together with theoretical curves for different $\beta$ values. It can be seen that the points lie close to the curve predicted for a Kolmogorov spectrum (*ie* the 44/12 line). The significant deviations may be due to the fact that the simulated subapertures were square in shape (for the calculation, the diameter of the subapertures was taken to be that of a circle with equivalent area; see section 3.6.3) and also partly due to failures in the simulation process, e.g. the discretisation of the phase in pixel values. However, since most real data is taken using square Shack-Hartmann subapertures, the closeness of the fit means that even moderately small deviations from the 11/3 power law should show up in plots of experimental data.

The theoretical curves have been compared with experimental data taken using the 1-metre JKT telescope in La Palma, as shown in Fig. 6.3. The data set used is set A, as referred to in section 5.2. Each point represents the mean value over 1000 frames of the variance ratio for a number of spot pairs with similar separations. The graph would suggest that $\beta$ is much closer to 3 than the Kolmogorov prediction of 11/3 for

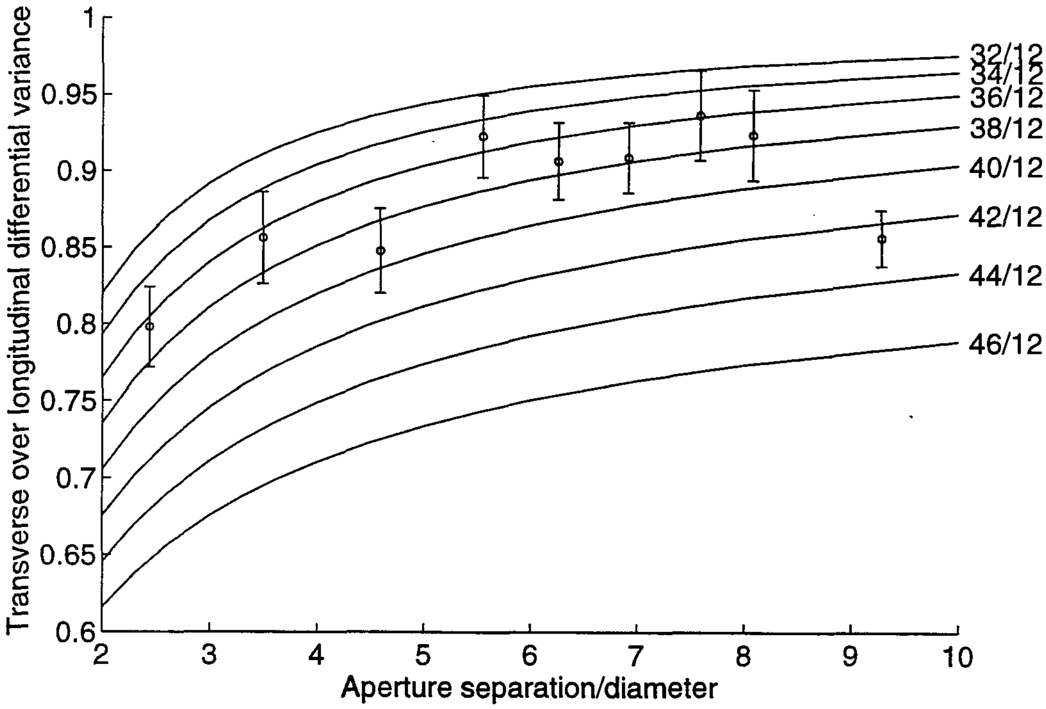



**Figure 6.3**: The ratio of differential variances found in experimental data (set A: JKT) compared with theoretical curves for varying $\beta$.

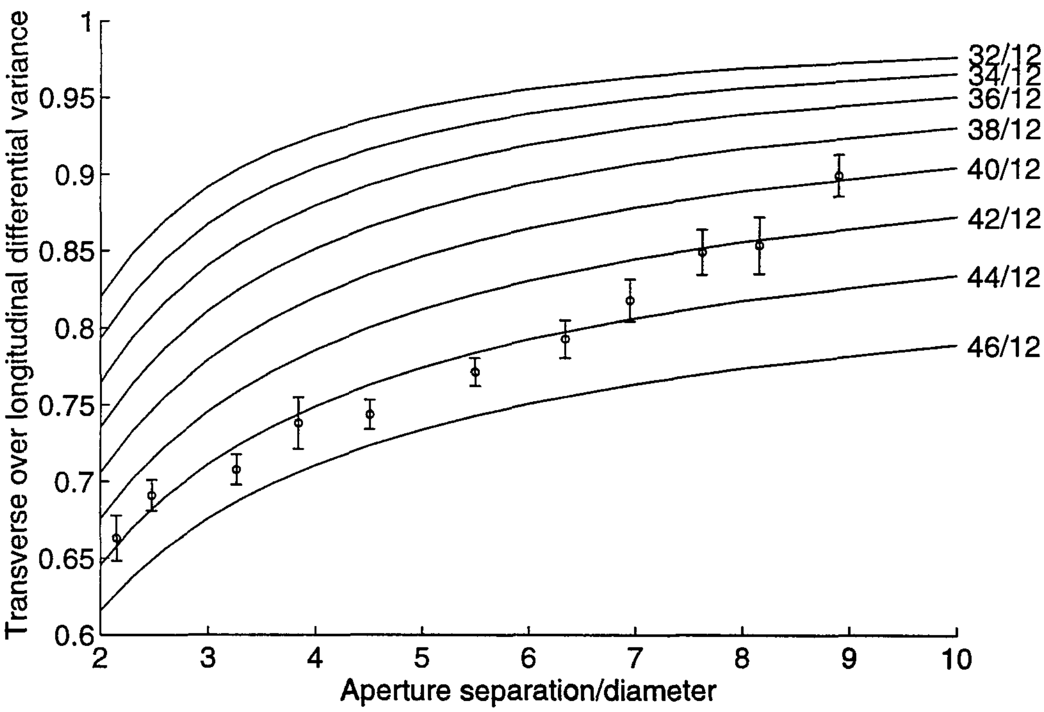
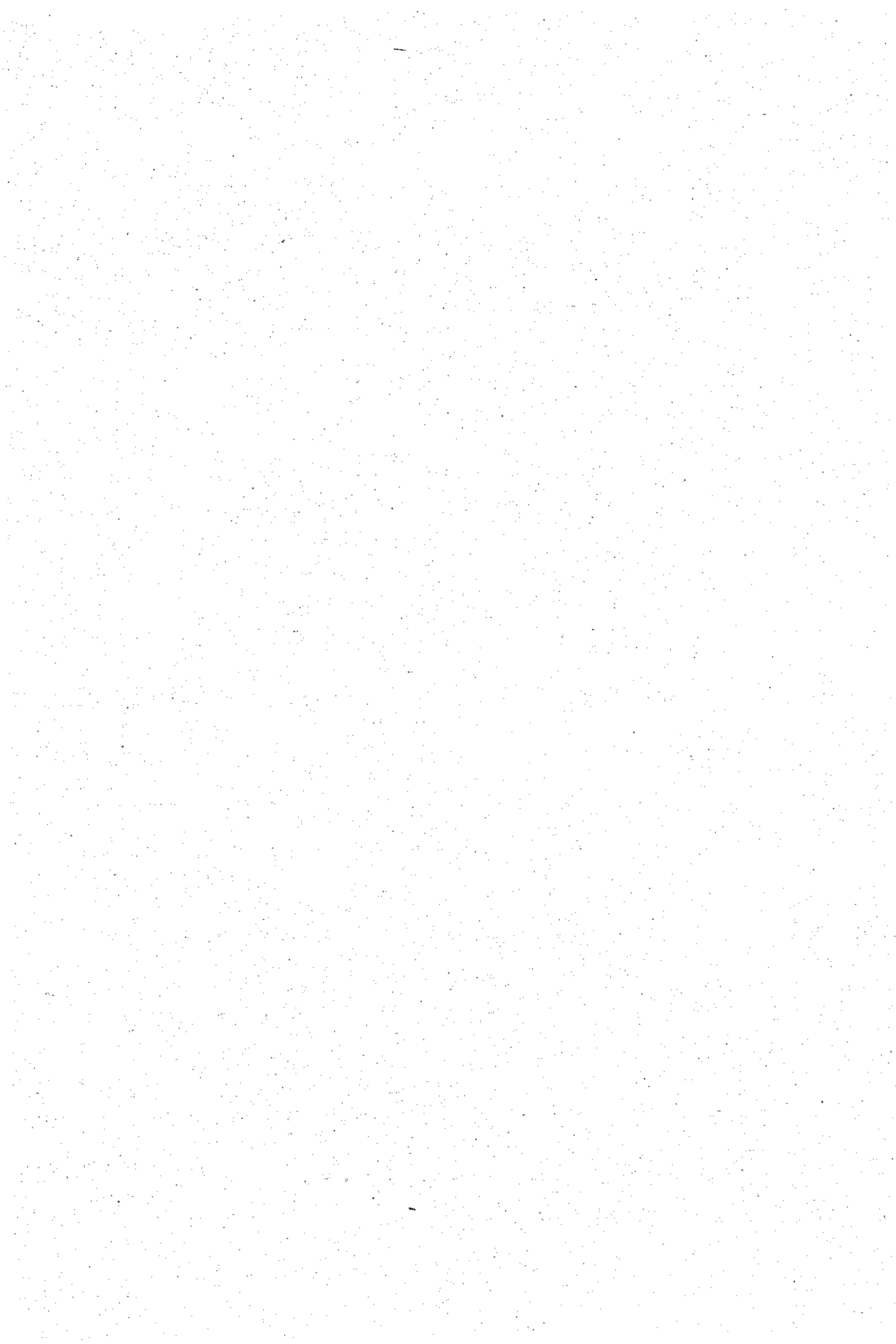



**Figure 6.4**: The ratio of differential variances found in experimental data (set L: Calar Alto) compared with theoretical curves for varying $\beta$.

this data. It should be recognised that this graph is by no means intended to be representative of the properties of the spectra associated with atmospheric turbulence in general. Fig. 6.4 shows an equivalent plot from 20000 frames taken at the 1.23m telescope at Calar Alto. In this case, the behaviour is much closer to that expected from Kolmogorov turbulence.



# Chapter 7

# Concluding Remarks

This thesis has evaluated the potential of the Shack-Hartmann wavefront sensor in performing measurements of the atmospheric 'seeing'; that is, the statistical properties of the distribution of turbulence in the atmosphere which have a direct effect on the quality of an astronomical image. The major results of the Kolmogorov model, proposed for refractive index fluctuations, have been described; these have been used as the starting point for a comprehensive review of an analysis of optical propagation through an atmosphere containing turbulence which behaves in accordance with the model. This analysis enables the statistical description of turbulence in the atmosphere to be directly related to the quality of an image formed by a telescope. The structure function of (spatial) phase fluctuations due to Kolmogorov turbulence can be completely described in terms of a single time-dependent parameter, $r_0$. It has also been shown that the temporal properties of the phase may be predicted using the Taylor hypothesis.

Simple relationships between the seeing itself and the output of the Shack-Hartmann sensor, namely the centroids of images formed by its subapertures, have been reviewed. In certain cases, such as the dependence of the angle-of-arrival variance (as observed through square and circular apertures) on $r_0$, the derivation of previously established results has been simplified using original methods. These relations have been used in the analysis of samples of wavefront data. Due regard has been given to the procedures necessary to avoid obtaining biased results and to obtain results with the prescribed level of accuracy. In particular, the effect of any correlation between measurements on this level has been investigated. The experimental practices involved have also been extensively discussed, as have the computer algorithms necessary to process the large volumes of data. A critical analysis of the results obtained has been made, with particular emphasis on those aspects of the results which suggest deviations from the Kolmogorov model.

Perhaps the most original contribution of this work is towards the investigation and characterisation of non-Kolmogorov turbulence. The conventional methods of analysing Shack-Hartmann data require that the strict relationship between the statistics of the phase and those of its first derivative, which is predicted by the Kolmogorov model, be used in order to extract $r_0$. It has been shown here that simple Shack-Hartmann centroid measurements can, in principle, be used to measure the complete form of the structure function itself. A possible generalised form for this function is given in terms of two variables: $\mathcal{R}_0$, analogous to $r_0$ as defined in the Kolmogorov case, and $\beta$, which is the exponent of the power law describing the variation of phase over distance ($\beta$ is non-varying and equal to 11/3 for Kolmogorov fluctuations). Centroid measurements have been used, with some success, to measure these parameters using some of the wavefront data already discussed.

The UK Joint Observatory Site Evaluation (JOSE) programme, which is currently underway, has as its objective regular seeing measurements over an extended period of time. A knowledge of the statistical behaviour of the structure function over the typical period of an astronomical long-exposure, over the course of one night, and from one night to the next is necessary for the specification of an adaptive optics system. The JOSE project is intented to provide such information in order that adaptive optics systems can be incorporated within two existing large-aperture telescopes at the UK sites in La Palma and Hawaii. The project will use Shack-Hartmann sensors, including the coordinate-remapping design outlined in this thesis, to perform the seeing measurements; therefore many of the conclusions made here have relevance to this project.

It is further intended to use data from the JOSE project to measure the form of the atmospheric structure function on a large number of nights. It is hoped that a knowledge of the distribution of values of $\beta$ over a long period can be gathered. This should enable an adaptive system to be modelled accurately in terms of $\mathcal{R}_0$ and $\beta$. It is apparent that the measurement of $\beta$ may be very sensitive to averaging effects resulting from a finite exposure time [40]. A full investigation into the effects of temporal averaging on the measurement is currently being carried out. This is being done under the assumption of the Taylor hypothesis for typical maximum velocities thought likely to occur for turbulent layers in the atmosphere. Such an analysis ought to set a limit on the accuracy of any measurement of $\beta$ made using the method described in this thesis.

# Appendix A

# Error Analysis Formulae

The purpose of this appendix is to present the standard formulae which are involved in the error analysis described in the earlier chapters. All formulae in this section have been derived assuming Gaussian statistics.

## A.1 Error Resulting from Independent Estimates

Using a sample of $n$ independent estimates of a stationary quantity $\chi$, the mean is estimated as:

$$\mu_\chi = \mathcal{E}\left[\chi\right] = \frac{\sum_{i=1}^{n} \chi_i}{n}. \tag{A.1}$$

where $\mathcal{E}$ represents the expectation value of the quantity. Its variance can be estimated as:

$$\sigma_\chi{}^2 = \mathcal{E}\left[\chi^2\right] - \left(\mathcal{E}\left[\chi\right]\right)^2 = \frac{\sum_{i=1}^{n} \chi_i{}^2}{n-1} - \mu_\chi{}^2. \tag{A.2}$$

$\sigma_\chi$ is known as the standard deviation. The standard error, $\sigma_e$, can be used as a measure of the inaccuracy of estimating of a statistic of the population using a finite sample size. In the case of the mean, the standard error is given by [74]:

$$\sigma_e\left(\mu_\chi\right) = \frac{\sigma_\chi}{\sqrt{n}}. \tag{A.3}$$

The standard error on the estimate of the variance is found to be [74]:

$$\sigma_e\left(\sigma_\chi{}^2\right) = \sqrt{\frac{2}{n}}\sigma_\chi{}^2. \tag{A.4}$$

## A.2 Error Resulting from Correlated Estimates

In general, if estimates of a time varying (but stationary) quantity, $\chi(t)$, are made at finite intervals in time, some degree of correlation may exist between successive measurements. In this case, the measurements cannot be considered to be independent. In order to account for this, so that an unbiassed estimate of the error on a statistic can be made, it is useful to consider the autocovariance of the measured quantity, defined as

$$C_{XX}(\tau) = \mathcal{E}\left[\chi(t)\,\chi(t+\tau)\right] - {\mu_X}^2 \tag{A.5}$$

The standard error on the mean estimated over a finite time interval $T$ is given by [74]:

$$\sigma_e(\mu_X) = \frac{1}{T}\int_{-\infty}^{\infty} C_{XX}(\tau)\,d\tau, \tag{A.6}$$

and, similarly, for the variance [74]:

$$\sigma_e\left({\sigma_X}^2\right) = \frac{2}{T}\int_{-\infty}^{\infty} {C_{XX}}^2(\tau) + 2{\mu_x}^2 C_{XX}(\tau)\,d\tau. \tag{A.7}$$

Both these formulae assume that $T$ is large, the necessary requirement being that $C_{XX}(\tau)$ is negligible for values of $\tau$ comparable with $T$.

One can extend this formulation to deal with the case where estimates of the stationary quantity $\chi$ are made using samples at different locations. If these measurements are partially correlated then the standard errors due to spatial correlation can be calculated as:

$$\sigma_e(\mu_X) = \frac{1}{n}\cdot\frac{1}{2}\sum_{i=1}^{n} C_{XXi}, \tag{A.8}$$

and, similarly, for the variance:

$$\sigma_e\left({\sigma_X}^2\right) = \frac{2}{n}\cdot\frac{1}{2}\sum_{i=1}^{n} {C_{XXi}}^2 + 2{\mu_x}^2 C_{XXi}. \tag{A.9}$$

In this case, $C_{XXi}$ represents the covariance between the measurements taken at location 1 and location $i$. The factor of $1/2$ is due to the fact that the integral is taken from $-\infty$ to $\infty$, but the sum is taken only over positive coordinate differences.

In the particular application where the mean of the quantity being measured is zero (*i.e.* $\mu_X = 0$) and the variance is required to be estimated, one can rewrite Eq. A.4 in terms of an effective number of independent measurements, $n'$:

$$\sigma_e\left({\sigma_X}^2\right) = \frac{2}{n'}{\sigma_X}^2. \tag{A.10}$$

This effective number can be written as $n' = n/\mathcal{F}$ where $\mathcal{F}$ is a redundancy factor caused by the correlation between measurements:

$$\mathcal{F} = 2\sum_{i=1}^{n} {C'_{XXi}}^2. \tag{A.11}$$

The quantity $C'_{XXi}$ is simply the covariance normalised so that $C'_{XX1} = 1$.

$\mathcal{F} = 1$ for entirely independent measurements.

## A.3 Error Propagation Formulae

If $x$ is a measured quantity with associated error $\sigma_x$ and $y$ is some function of $x$, it is possible to estimate the error $\sigma_y$ on $y$ using the following formulae ($m$ is a real constant):

$$y = mx : \qquad \frac{\sigma_y}{y} = \frac{\sigma_x}{x}. \tag{A.12}$$

$$y = x^m : \qquad \frac{\sigma_y}{y} = m\frac{\sigma_x}{x}. \tag{A.13}$$

# Bibliography


[1] A. Labeyrie. Attainment of diffraction limited resolution in large telescopes by Fourier analysing speckle patterns in star images. *Astronomy and Astrophysics*, 6:85–87, 1970.

[2] J.C. Dainty. Speckle interferometry in astronomy. In L. Johnson and C. Allen, editors, *Symposium on Recent Advances in Observational Astronomy*, pages 95–111, 1981.

[3] J.C. Dainty. Stellar speckle interferometry. In *Topics in Applied Physics, Volume 9: Laser Speckle and Related Phenomena*, chapter 7. Springer-Verlag Berlin Heidelberg, 1984.

[4] J.C. Dainty and J.R. Fienup. *Image Recovery: Theory and Application, Phase Retrieval and Image Reconstruction for Astronomy*, chapter 7, pages 231–275. Academic Press, 1987.

[5] Pedro Negrete-Regagnon. *Bispectral Imaging in Astronomy.* PhD thesis, Imperial College, London SW7 2BZ, 1995.

[6] J. Primot, G. Rousset, and J.C. Fontanella. Deconvolution from wave-front sensing: A new technique for compensating turbulence-degraded images. *Journal of the Optical Society of America A.*, 7(9):1598–1608, September 1990.

[7] W.T. Rhodes and J.W. Goodman. Interferometric technique for recording and restoring images degraded by unknown aberrations. *Journal of the Optical Society of America*, 63(6):647–657, June 1973.

[8] A.H. Greenaway. Terrestrial optical aperture synthesis technique (TOAST). *Optics Communications*, 58(3):149–154, June 1986.

[9] V.I. Tatarski. *Wave Propagation in a Turbulent Medium.* McGraw-Hill, 1961. Translated from the Russian by R.A. Silverman.

[10] R.R. Beland. Propagation through atmospheric optical turbulence. In F.G.Smith, editor, *The Infrared and Electro-Optical Systems Handbook Volume 2: Atmospheric Propagation of Radiation*, chapter 2. SPIE Optical Engineering Press, 1993.

[11] T.S. McKechnie. Atmospheric turbulence and the resolution limits of large ground-based telescopes. *Journal of the Optical Society of America*, 9(11):1937–1954, November 1992.

[12] F. Roddier. The effects of atmospheric turbulence in optical astronomy. *Progress in Optics*, 19:281–377, 1981.

[13] J. Vernin and M. Azouit. Traitement d'image adapte au speckle atmospherique. *Journal of Optics (Paris)*, 14(3):131–142, 1983.

[14] J. Vernin and F. Roddier. Experimental determination of two-dimensional spatiotemporal power spectra of stellar light scintillation. Evidence for a multilayer structure of the air turbulence in the upper troposhere. *Journal of the Optical Society of America*, 63(3):270–273, March 1973.

[15] J. Vernin and C. Muñoz-Tunon. Optical seeing at La Palma observatory II. intensive site testing campaign at the Nordic optical telescope. To be published in Astronomy & Astrophysics, November 1993.

[16] R.E. Hufnagel and N.R. Stanley. Modulation transfer function associated with image transmission through turbulent media. *Journal of the Optical Society of America*, 54(1):52–61, January 1964.

[17] J.W. Goodman. *Statistical Optics*. John Wiley & Sons Inc., 1985.

[18] R.S. Lawrence and J.W. Strohbehn. A survey of clear-air propagation effects relevant to optical communications. *Proceedings of the IEEE*, 58(10):1523–1545, October 1970.

[19] R.K. Tyson. *Principles of Adaptive Optics*. Academic Press Inc., 1991.

[20] A. Papoulis. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill Inc, 1965.

[21] V.I. Tatarski. *The Effects of the Turbulent Atmosphere on Wave Propagation*. Israel Program for Scientific Translations Ltd, 1971.

[22] D.L. Fried. Statistics of a geometric representation of wavefront distortion. *Journal of the Optical Society of America*, 55(11):1427–1435, November 1965.

[23] D.L. Fried. Optical resolution through a randomly inhomogeneous medium for very long and very short exposures. *Journal of the Optical Society of America*, 56(10):1372–1379, October 1966.

[24] R.J. Noll. Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America*, 66(3):207–211, March 1975.

[25] D.L. Fried. Probability of getting a lucky short-exposure image through turbulence. *Journal of the Optical Society of America*, 68(12):1651–1658, December 1979.

[26] A.T. Young. Seeing: Its cause and cure. *The Astrophysical Journal*, 189:587–604, May 1974.

[27] M. Born and E. Wolf. *Principles of Optics*. Pergamon Press, 1980.

[28] S.F. Clifford. The classical theory of wave propagation in a turbulent medium. In J.W.Strohbehn, editor, *Laser Beam Propagation in the Atmosphere*, chapter 2. Springer-Verlag, New York, 1978.

[29] D.L. Fried. Atmospheric turbulence optical effects: Understanding the adaptive optics implications. In D.M Alloin and J.-M. Mariotti, editors, *Adaptive Optics For Astronomy*, Dordrecht, Netherlands, July 1993. NATO Advanced Study Institute, Kluwer Academic Publications.

[30] J-M. Conan, G. Rousset, and P-V. Madec. Wavefront temporal spectra in high resolution imaging through turbulence. *Journal of the Optical Society of America*, 12(7):1559–1570, July 1995.

[31] D.P. Greenwood. Bandwidth specification for adaptive optics systems. *Journal of the Optical Society of America*, 67(3):390–393, March 1977.

[32] D.L. Fried. Greenwood frequency measurements. *Journal of the Optical Society of America A.*, 7(5):946–947, May 1990.

[33] M.M. Colavita, M. Shao, and D. Staelin. Atmospheric phase measurements with the Mark 3 stellar interferometer. *Applied Optics*, 26(19):4106–4112, October 1987.

[34] D.L. Fried. Time-delay-induced mean-square error in adaptive optics. *Journal of the Optical Society of America A.*, 7(7):1224–1225, July 1990.

[35] J.Y. Wang and J.K. Markey. Modal compensation of atmospheric turbulence phase distortion. *Journal of the Optical Society of America*, 68(1):78–87, January 1978.

[36] R.G. Lane and M. Tallon. Wave-front reconstruction using a Shack-Hartmann sensor. *Applied Optics*, 31(32):6902–6908, November 1992.

[37] V.N. Mahajan. Zernike circle polynomials and optical aberrations of systems with circular pupils. *Applied Optics*, 33(34):8121–8124, November 1994.

[38] V.N. Mahajan. Zernike annular polynomials and optical aberrations of systems with annular pupils. *Applied Optics*, 33(34):8125–8127, November 1994.

[39] N. Roddier. Atmospheric wave-front simulation using Zernike polynomials. *Journal of Optical Engineering*, 29:1174–1180, 1990.

[40] H.M. Martin. Image motion as a measure of seeing quality. *Publications of the Astronomical Society of the Pacific*, 99:1360–1370, December 1987.

[41] C. Roddier and F. Roddier. Influence of exposure time on spectral properties of turbulence-degraded astronomical images. *Journal of the Optical Society of America*, 65(6):664–666, June 1975.

[42] R.C. Cannon. Global wave-front reconstruction using Shack-hartmann Sensors. *Journal of the Optical Society of America A.*, 12(9):2031–2039, September 1995.

[43] J.W. Goodman. *Introduction to Fourier Optics.* McGraw-Hill, Inc, 1968.

[44] R.N. Bracewell. *The Fourier Transform and its applications.* McGraw-Hill, 1986.

[45] M. Sarazin and F. Roddier. The ESO differential image motion monitor. *Astron. Astrophys.*, 227:294–300, 1990.

[46] D.L. Fried. Differential angle of arrival: Theory, evaluation, and measurement feasibility. *Radio Science*, 10(1):71–76, January 1975.

[47] W.J. Tango and R.Q. Twiss. Michelson stellar interferometry. *Progress in Optics*, 17:258, 1980.

[48] M. Sechaud, G. Rousset, V. Michau, J.C. Fontanella, J.G. Cuby, F. Rigaut, and J.C. Richard. Wavefront sensing in imaging through the atmosphere: A detector strategy. *Active and Adaptive Optical Components, SPIE*, 1543:479–490, 1991.

[49] G. Rousset. Wavefront sensing. In D.M Alloin and J.-M. Mariotti, editors, *Adaptive Optics For Astronomy*, Dordrecht, Netherlands, July 1993. NATO Advanced Study Institute, Kluwer Academic Publications.

[50] K.A. Winick. Cramer-Rao lower bounds on the performance of charged-couped-device optical position estimators. *Journal of the Optical Society of America, A.*, 3(11):1809–1815, November 1986.

[51] J.M. Beckers. Interpretation of out-of-focus star images in terms of wave-front curvature. *Journal of the Optical Society of America, A.*, 11(1):425–427, January 1994.

[52] F. Roddier. Wavefront sensing and the irradiance transport equation. *Applied Optics*, 29(10):1402–1403, April 1990.

[53] C. Roddier and F. Roddier. Wave-front reconstruction from defocused images and the testing of ground-based optical telescopes. *Journal of the Optical Society of America, A.*, 10(11):2277–2287, November 1993.

[54] F. Roddier. Curvature sensing: A diffraction theory. Technical Report 87-3, NOAO Advanced Development Program, December 1987.

[55] F. Roddier. Curvature sensing and compensation: A new concept in adaptive optics. *Applied Optics*, 27(7):1223–1225, 1988.

[56] F. Roddier, C. Roddier, and N. Roddier. Curvature sensing: A new wavefront sensing method. *SPIE*, 976:26, 1988.

[57] P. Hickson. Single-image wavefront curvature sensing. *Journal of the Optical Society of America*, 11(5):1667–1673, May 1994.

[58] F. Roddier. Error propagation in a closed-loop adaptive optics system - a comparison between shack-hartmann and curvature wave-front sensors. *Optics Communications*, 113(4–6):357–359, 1995.

[59] D.L. Fried. Least-square fitting a wave-front distortion estimate to an array of phase-difference measurements. *Journal of the Optical Society of America*, 67(3):370–375, March 197.

[60] J.E. Graves and D.L. McKenna. The University of Hawaii adaptive optics system: Iii the wavefront curvature sensor. *SPIE*, 1542:262–271, 1991.

[61] F. Roddier, J.E. Graves, and E. Limburg. Seeing monitor based on wavefront curvature sensing. *Advanced Technology Optical telescopes IV, SPIE*, 1236:474–479, 1990.

[62] T.W. Nicholls, C.J. Solomon, M. Wells, N.J. Wooder, and J.C. Dainty. Experiments in wavefront sensing at La Palma. In M.A. Ealey and F. Merkle, editors, *Adaptive Optics in Astronomy, SPIE*, volume 2201, pages 555–561, March 1994.

[63] J-Y Zhang. *High Angular Resolution Speckle Imaging.* PhD thesis, Imperial College of Science, Technology and Medicine, London SW7 2BZ, 1994.

[64] J.M. Mariotti. Experimental results on atmospheric turbulence obtained with an infrared speckle-interferometer. *Optica Acta*, 30(6):831–840, 1983.

[65] M.J. Selby and R. Wade. Speckle interferometry in the near-infrared. *Monthly Notices of the Royal Astronomical Society*, 187:553–566, 1979.

[66] P. Puxley. ALICE update and IRCAM3 performance. ROE report., 1993.

[67] M. Wells. An infrared Shack-Hartmann sensor, February 1994. Institute of Astronomy, University of Edinburgh.

[68] T.J. Norton. *Development of a Microchannel Plate Image Intensifier for an Astronomical Photon-Counting Detector.* PhD thesis, Imperial College, 1995.

[69] D. Dayton, B. Pierson, B. Spielbusch, and J. Gonglewski. Atmospheric structure function measurement with a Shack-Hartmann wave-front sensor. *Optics Letters*, 17(24):1737–1739, December 1992.

[70] M. Bester, W.C. Danchi, L.J. Degiacomi, and C.H. Townes. Atmospheric fluctuations: Empirical structure functions and projected performance of future instruments. *The Astrophysical Journal*, 392:357–374, June 1992.

[71] R.G. Buser. Interferometric determination of the distance dependence of the phase structure function for near-ground horizontal propagation at 6328 angstroms. *Journal of the Optical Society of America*, 61(4):488–491, April 1971.

[72] A.S. Gurvich and M.S. Belen'kii. Influence of stratospheric turbulence on infrared imaging. *Journal of the Optical Society of America, A.*, 12(11):2517–2522, November 1995.

[73] G.D. Boreman and J.C. Dainty. Zernike expansions for non-Kolmogorov turbulence. *The Journal of the Optical Society of America, A.*, 1996 (in press).

[74] J.S. Bendat and A.G. Piersol. *Random Data - Analysis and Measurement Procedures*. John Wiley & Sons, 1986. Second Edition.