

PHOTON LIMITED IMAGING

# **Imaging at Low Light Levels using the Triple Correlation**

Malcolm John Northcott

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

Supervised by;  
Professor J.C.Dainty,  
Applied Optics Group, Department of Physics.

Imperial College of Science, Technology and Medicine  
University of London

June 1989

## Acknowledgements

My sincerest thanks to Professor Chris Dainty for his encouragement and guidance over the period in which this research was carried out. Thanks to his efforts to find funding, my colleagues and myself have been able to attend a number of international conferences. He has also organised several collaborative observing programs, which have lead to my participation in a number of interesting and stimulating field trips. Working with Chris has been a great pleasure and privilege.

Also many thanks to Dr Geoffrey Ayers who helped immensely, in understanding many aspects of this subject, and for his atmospheric simulation program, which provided me with valuable test data. I should also like to thank Nick Wooder, whose computer programs and help in maintaining the computers, has saved me a great deal of time and effort, Angela Darling, for reading and criticising much of the contents of this thesis and Luiz DeFretas for playing the devil's advocate.

I would also like to thank my parents for their patience, encouragement and financial support during this period of my life. Many thanks to Jackie, my friends and colleagues, for their support and encouragement, especially in the closing phases of writing this thesis.

### Statement of original work

In chapter two section 2.2.1 is a re-written and simplified version of the proof given by Goodman in "Statistical Optics", Wiley 1985. The remainder of the work in this chapter is review material. Chapter three consists mainly of review material taken from many sources. The exceptions are sections 3.3 and 3.4 which do not appear in the literature in this form and sections 3.6 and 3.9 which are wholly my own work. Chapter four consists of my own work. The detailed implementation of the algorithms presented in chapter five is my own work. However, the direct and the radon method have been mentioned in the literature. The sub-plane method is wholly my own work and is probably the most important contribution to the field contained in this thesis. The material making up chapter six is mainly review material taken from a variety of sources. The phase closure arguments are of my own devising, but have a basis in published work. Section 6.7 represents work carried out in collaboration with G.Ayers <sup>19</sup>.

## Abstract

With the advent of imaging photon counting detectors, it is now possible to view dynamic processes at extremely low light levels. There has therefore arisen a need for techniques to extract image information from a stream of photon limited data. The triple correlation allows us to recover an image from a series of photon limited frames, in spite of image translation or the effects of turbulence on the imaging system.

In this work the semi-classical theory of photo-detection is reviewed and current technologies for producing imaging photon detectors are discussed. We then discuss of the properties of the triple correlation and its Fourier transform, the bispectrum.

It is shown that the triple correlation can be used to image a randomly translating object at extremely low light levels. This is demonstrated by experiment, using data from an early model of the ITL Imaging Photon Detector. In spite of the many shortcomings of this particular detector, reasonable results were achieved for a one dimensional imaging problem.

In optics, one of the main applications of the triple correlation has been in the domain of speckle interferometry. However this application presents us with the problem of manipulating the triple correlation or bispectrum of large two dimensional image arrays. Three techniques for dealing with large arrays of data are presented; each has been coded and tested with simulated translating object data and speckle interferometry data.

It then shown why the triple correlation is useful in the field of speckle interferometry and brief analysis of signal to noise ratio (SNR) is given. Finally a simple reconstruction of a binary star from real speckle data is presented.

|                                                      |          |
|------------------------------------------------------|----------|
| <b>1 Introduction</b>                                | <b>7</b> |
| <br>                                                 |          |
| <b>2 Photon Detectors</b>                            |          |
| 2.1: Photon counting                                 | 14       |
| 2.2: Semi-classical theory                           | 14       |
| 2.2.1: The perfect photo-multiplier                  | 15       |
| 2.3: Dead time                                       | 19       |
| 2.3.1: Non polarisable dead time                     | 21       |
| 2.3.2: Polarisable dead time                         | 22       |
| 2.3.3: The effect of dead time on count rate         | 22       |
| 2.4: Imaging photon detectors                        | 24       |
| 2.4.1: Time framed devices                           | 25       |
| 2.4.2: Time tagging devices                          | 29       |
| 2.5: The IPD detector                                | 33       |
| 2.6: The PAPA detector                               | 35       |
| 2.7: The MAMA detector                               | 39       |
| 2.8: The future; avalanche photo-diodes              | 40       |
| <br>                                                 |          |
| <b>3: The Triple Correlation</b>                     |          |
| 3.1: Notation                                        | 42       |
| 3.2: Relationship to the auto-correlation            | 42       |
| 3.3: Symmetry of the triple correlation              | 45       |
| 3.4: Symmetry of the bispectrum                      | 45       |
| 3.5: Calculation of the TC using photon differencing | 47       |
| 3.6: Photon bias                                     | 53       |
| 3.7: Reconstruction of phase from the bispectrum     | 59       |
| 3.8: Photon noise                                    | 65       |

|                                                              |    |
|--------------------------------------------------------------|----|
| 3.9: Photon differencing with a continuous time window ..... | 70 |
| 3.10: TC of an object on a uniform background .....          | 75 |

#### **4: Imaging A Randomly Translating Object**

|                                        |    |
|----------------------------------------|----|
| 4.1: Introduction .....                | 77 |
| 4.2: The shift and add algorithm ..... | 77 |
| 4.3: The experimental setup .....      | 82 |
| 4.4: Image translation .....           | 85 |
| 4.5: Data storage .....                | 87 |
| 4.6: Experimental results .....        | 90 |

#### **5: Calculation of the Triple Correlation for Large Arrays**

|                                  |     |
|----------------------------------|-----|
| 5.1: Introduction .....          | 93  |
| 5.2: The direct method .....     | 96  |
| 5.3: The radon method .....      | 103 |
| 5.4: Sub plane calculation ..... | 110 |

#### **6: Imaging through Atmospheric Turbulence.**

|                                                         |     |
|---------------------------------------------------------|-----|
| 6.1: Introduction .....                                 | 116 |
| 6.2: Speckle interferometry .....                       | 116 |
| 6.3: Recovery of phase information .....                | 122 |
| 6.4: Triple correlation .....                           | 122 |
| 6.5: Phase closure and triple correlation .....         | 124 |
| 6.6: Signal to noise .....                              | 131 |
| 6.7: Reconstruction from a simulated point source ..... | 136 |
| 6.8: Reconstruction of more complex objects .....       | 138 |

|                      |     |
|----------------------|-----|
| 7: Conclusions. .... | 141 |
|----------------------|-----|

## Appendices

|                                                                                                      |     |
|------------------------------------------------------------------------------------------------------|-----|
| A: Proof that the inter event statistics of a Poisson source<br>are negative exponential .....       | 143 |
| B: Deriving the form of the bispectrum from the<br>triple correlation .....                          | 145 |
| C: 68010 machine code routine for calculating the triple<br>correlation by photon differencing ..... | 147 |

## **Introduction.**

In classical optics light is assumed to be described as a pure wave phenomenon. However we know from the more modern and more accurate quantum mechanical formulation of electro-magnetic theory that this simple description is not complete. In particular, the quantum theory suggests that the electromagnetic field is quantized, or in other words, that it transports energy in discrete packets, which are given the name photons. The most convincing demonstration of this quantization is provided by the photo-electric effect.

As in many other fields, one of the most profound consequences of the quantization of the electro-magnetic field is that it implies an absolute upper bound on the amount of information that can be carried by a beam of light of a given mean intensity and wavelength. That is, it imposes a fundamental limit to the finesse with which measurements can be made.

Modern instrumentation allows us in some circumstances to approach very closely to this fundamental sensitivity limit. In particular the photo-multiplier tube and similar devices allow us at low flux levels to count individual photon events; these photon events can in some sense be associated with individual photons, i.e. quanta of the electro-magnetic field. In these circumstances the photon noise (analogous to shot noise) associated with the discrete nature of the signal can contribute a significant proportion of the overall noise on the detected signal.

A traditional photo-multiplier tube allows us to measure the time varying intensity of an optical field at a single point. By contrast the new generation of imaging photon detectors such as the "IPD" and "PAPA" detectors, allow the measurement of time varying optical intensity at many points simultaneously.

Strictly speaking none of these detectors allow simultaneous measurement, since only one photon event at a time can be detected. With these detectors we can image rapidly changing intensity patterns at low light levels. However the typical bandwidths achievable by imaging detectors, between  $10^5$  to  $10^7$  events per second, place an upper limit on the speed of phenomena which we can record. When integrated over time periods short enough to freeze an intensity pattern, a frame of data may consist of only a small number of photo-events. There thus exists a need to develop techniques which can be used to extract information from a series of images which are severely degraded by photon noise. Figure 1.1 on page 10 shows a computer simulation of what we would expect to see when imaging a object at different photon event count levels. We see that even with 200,000 photons, the photon noise on the image is significant (about a 5 percent variance on the signal in each pixel) for this particular image. In the range of  $2 \rightarrow 200$  photons this image is virtually unrecognisable.

In Chapter two, we discuss the semi-classical theory of photo-electric detection. This theory is a quantum extension of the traditional classical theory. It can be shown to give the same predictions as full quantum electromagnetic theory, at least in as far as results dependent upon photo-electric detection are concerned. The rest of the chapter consists of a review of the currently available imaging photon detectors, including their strengths and weaknesses.

With the general availability of imaging photon detectors, there is now a need for methods of processing and extracting information from photon limited images. Imagine that we are observing an image which is undergoing continuous evolution in time. If we were to use an imaging photon detector to record such an image, then we could imagine partitioning the photon event stream into

a series of time frames. Provided that the temporal length of each of these frames were short compared with the time scale of the image evolution, we would thereby generate a series of snap-shots of the evolving image. Unfortunately due either to the limited data-rate of these detectors, or the intrinsic faintness of the source under observation, the resulting snap-shot might be severely degraded by photon noise. If we wish to extract useful information from such a time series, we would probably have to average information from several frames of data. In many cases we would need to transform the data in order that a meaningful average can be formed. If the time evolution of the image is due to random translation of the object (linear wavefront tilts), or the degrading effects of atmospheric turbulence, then the auto-correlation and triple correlation transformations are useful in this respect. The bulk of this thesis is concerned with the application of the triple correlation to imaging translating objects and imaging through turbulence at low light level. Chapter 3 describes the properties of the triple correlation in detail.

For the case of a randomly translating object the most obvious way to process the data, would be to sum the data frames, after making their centres of gravity coincident. Unfortunately, at very low light level this process leads to distortion of the image.

Since the auto-correlation of a function is independent of the absolute coordinates of that function we may usefully average the auto-correlations of each photon limited image. It is straight-forward to show that apart from a trivial bias term this average will converge toward the auto-correlation of the high light level image. However the auto-correlation of an image contains only information about its Fourier modulus. As a consequence of this, recovery of an image from its auto-correlation is in general a difficult problem, especially in the presence of noise.

### Photon limited images

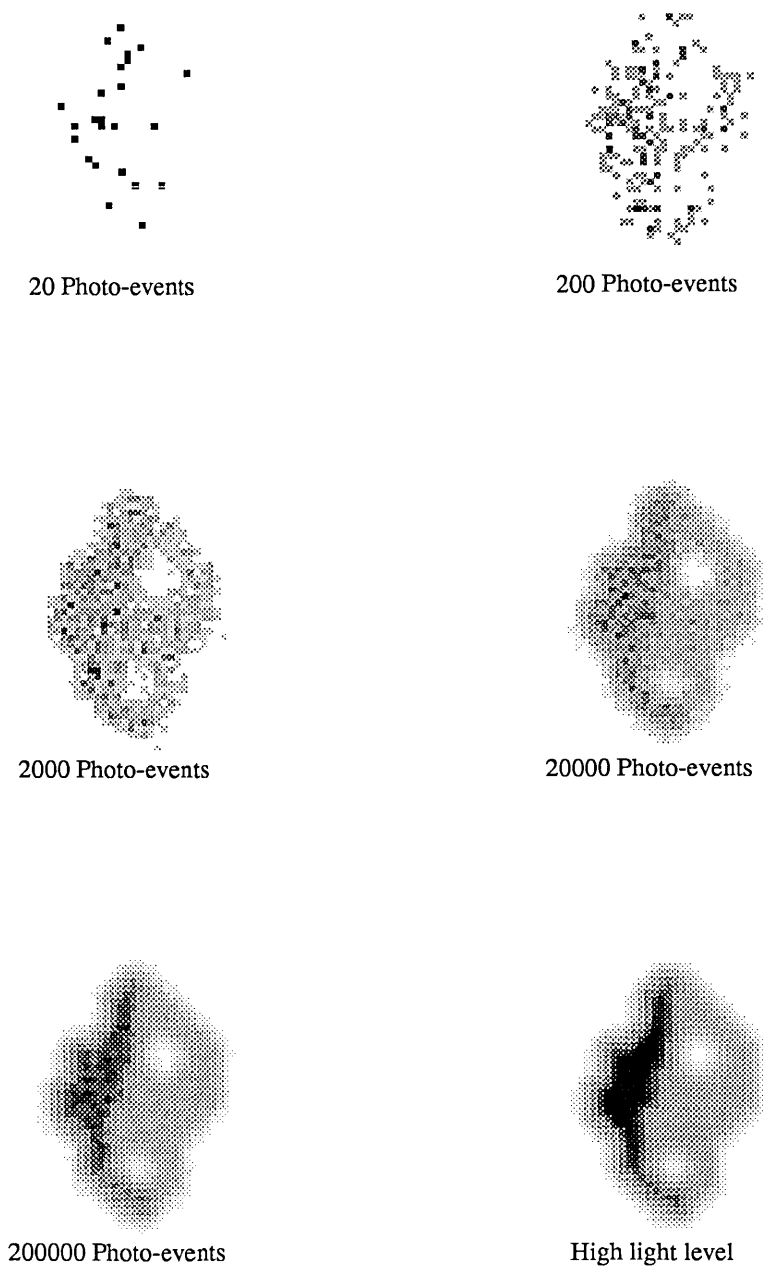


**Fig. 1.1**

*The above pictures show an example of an image recorded at different mean photon rates; even with 200000 photon events the image is degraded by a significant amount of photon noise.*

---

The triple correlation of a function, like its auto-correlation is also independent of the absolute co-ordinates of the function being correlated. In contrast to the auto-correlation the triple-correlation of an image contains both Fourier phase and modulus information about the original function. The ability to obtain this phase information from the triple correlation makes reconstruction of a function from its triple correlation a trivial task in the absence of noise <sup>1</sup>. We have used this technique to construct an image from a computer simulation of a randomly moving object images at low photon levels. In Chapter four an experiment to demonstrate imaging a randomly translating object at extremely low light level is described, using data from an imaging photon detector.

The triple correlation of a one dimensional function is two dimensional, and that of a two dimensional function is four dimensional. This presents a major practical difficulty when trying to process large two dimensional arrays of data, mainly because computers do not usually have enough memory to store the whole of the triple correlation of a large two dimensional array. In Chapter five we present three different approaches to solving this problem. Reconstructions from simulated data are presented for all three approaches.

As mentioned previously, the triple correlation and auto-correlation are also useful for imaging through turbulence. The triple correlation has already been used to process speckle data from large astronomical telescopes <sup>2,3</sup>. This use of the triple correlation can be viewed as an extension of the of the auto-correlation technique first proposed by Labeyrie <sup>4</sup>.

In Chapter six some theoretical foundation for the use of the triple correlation for astronomical imaging is presented. In particular we look at the connection between the radio astronomy technique of phase closure and the use of the triple correlation in optical speckle interferometry. Recently an optical analogue of the phase closure technique has been demonstrated by Baldwin et al <sup>5</sup>.

who used an aperture mask in a large optical telescope to simulate several smaller telescopes. This demonstrates the strong connection between some aspects of radio astronomy and high resolution imaging in optical astronomy. In the last part of chapter six some reconstructions from real speckle data are presented.

During the course of the research the following papers have been published in journals and conference proceedings.

J.C. Dainty, M.J. Northcott. "Imaging a randomly translating object using the triple correlation.", *Opt. Comm.*, **58**, 11-16 (1986).

J.C. Dainty, M.J. Northcott. "Imaging a randomly translating object using the triple correlation.", *O.S.A Topical Meeting on Quantum limited Imaging & Image processing*, Hawaii (1986).

G.R. Ayers, J.C. Dainty, M.J. Northcott, "Comparison of speckle imaging techniques." *O.S.A Annual Meeting*, Seattle, (1986). (Abstract: *J.Opt.Soc.Am. A*, **3**, 121 Dec 1986).

G.R. Ayers, J.C. Dainty, M.J. Northcott, "Photon limited imaging through turbulence.", *Proc. SPIE*, **808**, 19-25 (1987).

G.R. Ayers, J.C. Dainty, M.J. Northcott, "Triple correlation imaging through turbulence.", *O.S.A Annual Meeting*, Rochester, (1987). (Abstract: *J.Opt.Soc.Am. A*, **4**, 108 Dec 1987).

G.R. Ayers, J.C. Dainty, M.J. Northcott, "Photon limited imaging at low light levels.", *Proc. NOAO-ESO Conference on; High resolution imaging by interferometry*, Garching (1988).

G.R. Ayers, M.J. Northcott, J.C. Dainty "Knox-Thompson and triple correlation imaging through atmospheric turbulence.", J.Opt.Soc.Am. A, **5**, 964-985 (1988).

M.J. Northcott, G.R. Ayers, J.C. Dainty, "Algorithms for image reconstruction from photon limited data using the triple correlation.", J.Opt.Soc.Am. A, **5**, 986-992 (1988).

## 2

### **Photon detectors.**

#### **2.1. Photon counting.**

Up to 1900 the wave theory was thought to be sufficient to explain all observable optical phenomena. In 1900 Max Planck postulated the quantization of light in order to explain the spectral distribution of black body radiation and in 1905 Einstein used quantization of the electromagnetic field to explain the photo-electric effect; it thus became clear that the simple wave model of light did not offer a complete description of optics.

One of the more profound consequences of the quantization of light is to place a theoretical limit on the amount of information that a light beam of given intensity can carry. Indeed in many applications, for instance, where photon counting photo-multipliers or similar devices are used, or where very fine metrology is to be performed, this theoretical limit becomes of great practical significance.

#### **2.2. Semi-classical theory.**

For problems involving the detection of light using the photo-electric effect it is unnecessary to use the full quantum electrodynamic (QED) description of electromagnetism <sup>6</sup>. Instead we can use the semi-classical theory, which is based upon the following three axioms;

- 1: The probability of a photon event occurring in a space-time volume small compared with the coherence volume of a light field is proportional to the classical intensity of the light field within that space-time

volume, and to the size of the space-time volume.

- 2: The probability of multiple events occurring within the above mentioned space-time volume is negligibly small.
- 3: Individual photon events are statistically independent of one another.

### 2.2.1. The perfect photo-multiplier.

In the following section, we will show that the above three axioms imply that the counting distribution of photon events is described by the well known Poisson formula. For a perfect photomultiplier we can always choose a sampling time small enough to satisfy condition 2 above. Assume for the present that the optical field has a constant intensity over a counting window of time length  $t$ . We now ask the question; what is the probability of observing  $N$  photons in a counting window of length  $t + \delta t$ . Over the small additional time interval the chance of observing a photo-electron is  $KI\delta t$ , where  $I$  is the classical intensity and  $K$  is a conversion constant whose value will depend upon quantum efficiency of the detection system. This assertion makes use of axioms 1 and 2. If  $P(N; t)$  is the probability of observing  $N$  photon events in a time interval  $t$ , then using axiom 3. we can write;

$$P(N + 1; t + \delta t) = P(N; t) P(1; \delta t) + P(N + 1; t) P(0; \delta t). \quad 2.2.1$$

Equation 2.2.1 represents the only two ways in which we can partition the  $N + 1$  photons into the two time periods.

Using the fact that;

$$P(1; \delta t) = KI\delta t, \quad 2.2.2$$

and

$$P(0; \delta t) = (1 - KI\delta t), \quad 2.2.3$$

we can write;

$$\frac{P(N+1; t+\delta t) - P(N+1; t)}{\delta t} = KI \left[ P(N; t) - P(N+1; t) \right]. \quad 2.2.4$$

Allowing  $\delta t$  to tend towards zero we have;

$$\frac{dP(N+1; t)}{dt} = KI \left[ P(N; t) - P(N+1; t) \right]. \quad 2.2.5$$

If we now take the Laplace transform of Eqn 2.2.5.

$$p(N+1; s)(s + KI) = KI p(N; s), \quad 2.2.6$$

$$\text{where } p(N; s) = \int_0^{\infty} P(N; t) e^{-st} dt.$$

Writing M for N+1 and recursively applying 2.2.6 we get;

$$p(M; s) = \frac{(KI)^M}{(s + KI)^M} p(0; s). \quad 2.2.7$$

The only way in which we can observe zero photo-electrons in the time interval  $t+\delta t$  is if we observe no photoelectrons in the time interval  $t$ , and no photo-electrons in the time interval  $\delta t$ . Again using the fact that the photo-electron detection probabilities in two non-overlapping time intervals are independent we can write;

$$P(0; t + \delta t) = P(0; t)P(0; \delta t), \quad 2.2.8$$

Using Eqn. 2.2.3 and allowing  $\delta t$  to tend to zero,

$$\frac{dP(0; t)}{dt} = -KI P(0; t), \quad 2.2.9$$

thus;

$$P(0; t) = e^{-KI t}. \quad 2.2.10$$

Therefore the Laplace transform of  $P(0; t) = p(0; s) = \frac{1}{(KI + s)}$ .

From Eqn. 2.2.7 and Eqn. 2.2.10,

$$p(M; s) = \frac{(KI)^M}{(s + KI)^{M+1}}. \quad 2.2.11$$

Assume that  $F(M; t) = \frac{(KI t)^M e^{-KI t}}{M!}$ , then  $f(M; s)$ , its Laplace transform is given by;

$$\begin{aligned} f(M; s) &= \int_0^{\infty} \frac{(KI t)^M}{M!} e^{-t(KI + s)} dt, \quad 2.2.12 \\ &= \left[ \frac{-t^M}{KI + s} e^{-t(KI + s)} \right]_0^{\infty} + \frac{KI}{(KI + s)} \int_0^{\infty} \frac{(KI t)^{M-1}}{(M-1)!} e^{-t(KI + s)} dt, \\ &= \frac{KI}{(KI + s)} f(M-1; s), \\ &= \frac{(KI)^M}{(KI + s)^{M+1}}. \end{aligned}$$

Therefore, by comparison of Equations. 2.2.11 and 2.2.12;

$$P(M; t) \equiv F(M; t) = \frac{(KI t)^M e^{-KI t}}{M!}, \quad 2.2.13$$

which is the well known Poisson distribution. The quantity  $(KI t)$  is the mean number of photon events per time  $t$ , hence the more usual form in which the Poisson distribution is presented;  $P(M, \bar{N}) = \frac{\bar{N}^M e^{-\bar{N}}}{M!}$

In practice this result has a limited range of validity, because the small interval of time  $\delta t$  which we used in the proof cannot in reality be taken to zero. This is because real photon detectors suffer from the problem of dead time, which typically limits their minimum  $\delta t$  to between  $10^{-6}$

and  $10^{-9}$  seconds, depending upon the type of detector. Dead time and its effects are discussed more fully in section 2.3.

This proof has assumed that the classical intensity  $I$ , is constant; with a real light source this is never true, although it is approached very closely by the light from a well stabilised single mode laser. For a thermal light source the classical intensity is calculated by summation of the contributions of all the random phasors over the surface of the source. The classical intensity at a point illuminated by a thermal source therefore varies widely (its variance is equal to its mean intensity) on a time scale of the order of the reciprocal of the bandwidth of the detected light. Thus even for the relatively narrow bandwidth of 10nm the classical intensity will be varying at about  $10^{13}$  Hz . However the smallest  $\delta t$  that can be realised with a real detector is of the order of  $10^{-9}$  seconds and therefore the classical intensity from a thermal source at 10 nm bandwidth if integrated over this  $\delta t$  (sampling time) would be constant to a standard deviation of approximately 1 percent. Thus the above proof holds well for 'constant intensity' thermal light provided we do not use too narrow a bandwidth or sample too rapidly.

The most interesting photo-electron distributions occur when the time variation of the classical intensity is comparable to, or longer than the sampling time of the detector. Most of the following work is concerned with the photon limited detection of stochastic classical intensity variations on a long time scale compared <sup>WITH</sup> the sampling time of the detection system. The probability distribution of photon events in this case is doubly stochastic, one contribution coming from the statistics of the classical intensity variation, another from Poisson statistics of photo-electron emission.

In general we can calculate the photo-electron distribution by integrating the Poisson distribution for fixed classical intensity over the statistics of the classical light intensity  $I$ ;

$$P(N; t) = \int_0^{\infty} P(N; t; I) P_c(I) dI. \quad 2.2.14$$

$$= \int_0^{\infty} \frac{(KI t)^M e^{-KI t}}{M!} P_c(I) dI. \quad 2.2.15$$

Where  $P(N; t; I)$  is the conditional photo-electron probability for a fixed time  $t$  and fixed intensity  $I$  and  $P_c(I)$  is the probability distribution for the classical light intensity.

### 2.3. Dead time.

As mentioned above, in order to satisfy axiom number 2 of the semi-classical theory it is necessary that our detector sampling time can be reduced to a time which is small compared to the mean inter-event time of the detected signal. Equivalently, it is necessary that our detector bandwidth is large compared to the mean event frequency. In other words, in order to satisfy condition number 2 in general, an ideal photon event detector should be able to distinguish two photon events occurring within an arbitrarily short time of each other.

In reality all photon detectors suffer from the problem of dead time, which means that after a photon event has been detected there is a small interval of time over which the detector is insensitive to further photon events. This effect places a lower limit on the sampling time for which a particular detector will be useful. It is important to realise that this dead time is in general not a function of the detector photocathode, but rather

due to the internal gain mechanism and external electronics. That is, during a dead time period the probability of a photo-electron being emitted by a photocathode is unaffected; it is merely that during this period such a photo-electron will not be registered by the detection electronics as a photon event, whereas outside a dead time it will have some finite probability of being detected as a photon event. This means that the underlying photo-electron distribution is still Poisson; however the detected photon event distribution may show deviations from the Poisson distribution, Eq. 2.3.13.

It is possible to calculate  $P(N; t)$  under these circumstances; for instance for a detector exhibiting non-polarisable dead time the Poisson form of  $P(N; t)$  becomes <sup>7</sup>;

$$P(N; t) = \frac{\gamma(n-1, \beta_{n-1})}{(n-1)!} - \frac{\gamma(n, \beta_n)}{n!}. \quad 2.3.15$$

where

$$\gamma(n, \beta_n) = \int_0^{\beta_n} v^n e^{-v} dv,$$

is the incomplete gamma function,  $\beta_n = KI(1 - n\delta)$  and  $\delta$  is the ratio of the detector dead time to the interval  $t$ . Chang et al have developed a method for correcting the moments of such a distribution for the effects of dead time.<sup>8</sup>

An ideal or theoretically perfect photon counting detector, would have an output count rate, linearly proportional to the integrated classical intensity over its photo-cathode. Provided the count rate is kept low enough a real photon counting detector will behave very much as an ideal detector. Thus in the low photon event regime we can write;

$$\overline{R}_I = \epsilon I. \quad 2.3.16$$

where  $\overline{R}_I$  is the mean photon event rate for an ideal detector,  $I$  is the classical intensity integrated over the photo-cathode and  $\epsilon$  is a constant. Thus in an ideal photon counting device each photo-electron emitted by the photocathode has the same probability of being detected as a photon event. In photomultipliers and micro-channel plate (MCP) devices the bulk of this probability is associated with the probability of a photo-electron striking the first dynode in a photomultiplier, or entering a channel in a MCP.

The dead time of a photon counting detector is usually modelled as being of either the polarisable or non-polarisable type. In most cases one or other of these ideal forms of dead time behaviour can be used to describe a photo-electron detector to reasonable accuracy, provided the mean inter-event time (computed from  $\epsilon I_e$ ) is much greater than the detectors dead time.

### 2.3.1. Non-polarisable dead time.

The non-polarisable model assumes that once a photo-electron has been registered, the detector is unable to register further events for a time equal to the dead time of the detector. Also it assumes that any unregistered events occurring within such a dead time period do not extend the length of time for which the detector is inactive. The meaning of 'unregistered event', in this context, should be taken to mean a photon event which would have been registered if the detector in question had been an ideal photon counting detector. This model is typically applicable to standard photomultiplier tubes.

### 2.3.2. Polarisable dead time.

In order to understand this model let us start with an ideal photon counting detector. After every photon event which this ideal detector counts a polarisable detector will be unable to count further events for a dead time period. The crucial difference between this behaviour and that of a non-polarisable detector is that the polarisable detector will be disabled for a further dead time period by a new photo-electron, even if it is already in a dead time period. This type of behaviour describes well the behaviour of the local dead time of MCP, when they are used in 'saturated gain mode'.

### 2.3.3. The effect of dead time on count rate.

In order to appreciate the two dead time models more fully, it is instructive to calculate the count rate of detectors exhibiting dead time, as a function of light intensity, or equivalently the count rate of an ideal detector. We can see immediately that the asymptotic behaviour of a polarisable detector is for its count rate to approach zero, and the asymptotic behaviour of a non-polarisable detector is for its count rate to approach the reciprocal of the detector dead time. Of course any real detector may suffer severe damage before coming anywhere near these limits.

Looking first at the non-polarisable dead time model. In this case calculating the count rate as a function<sup>of</sup> an ideal detector's count rate is particularly simple. Assuming that the ideal detector's photon events follow Poisson statistics, the effect of dead time is simply to extend the mean inter-event time by an amount equal to the dead time. This comes directly from axiom number 3 of the semi classical theory, which implies that the

statistics of photon event timing are independent of when we choose to start observing. Thus we may 'choose' to start observing one dead time after each photon event, as opposed to immediately after. If for a perfect detector the mean inter-event time is  $\frac{1}{\bar{R}_I}$ , for a detector with dead time  $\tau$  the mean inter-event time is  $\frac{1}{\bar{R}_I} + \tau$ . The count rate is the reciprocal of the inter-event time and thus becomes  $\frac{\bar{R}_I}{(\bar{R}_I\tau + 1)}$ .

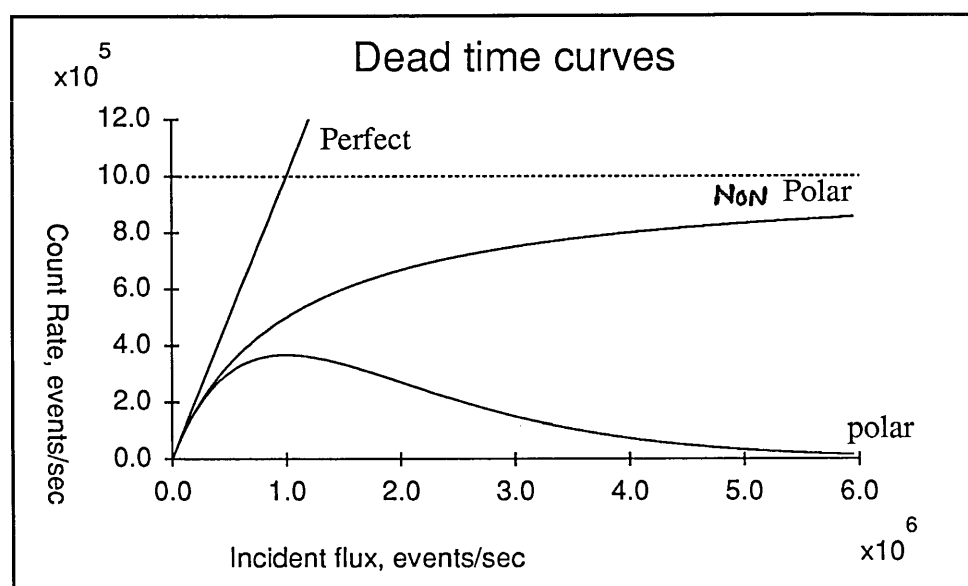
Calculating the count rate for a polarisable detector is slightly more complicated. Again consider an ideal photon-counting detector; to simulate the polarisable dead time of this detector it is necessary to impose the condition that a photon event is counted if and only if no photon events have occurred in a previous dead time interval. Thus in effect we are counting the number of times that the inter-event time is greater than the detector's dead time. In other words the count rate is the probability of obtaining an inter-event period of greater than the dead time  $\tau$  multiplied by the mean event rate for the same detector with zero dead time under the same conditions. We know that for Poisson statistics the inter-event statistics are negative exponential (see Appendix A), therefore;

$$\text{Rate} = \bar{m} \int_{\tau}^{\infty} \bar{R}_I e^{(-\bar{R}_I t)} dt . \quad 2.3.17$$

$$= \bar{R}_I e^{(-\bar{R}_I \tau)} \quad 2.3.18$$

Figure 2.1 shows the actual count rate as a function of the perfect detector count rate, for a perfect detector, a detector with non-polarisable dead time and a detector with polarisable dead time; we assume that the

quantum efficiency of all the detectors is the same.



*Count rate of a perfect detector vs count rate for a detector suffering from either polarisable or non-polarisable dead time.*

**Fig. 2.1**

For both dead-time-affected detectors their dead time is taken as  $1\mu\text{s}$ . The perfect detector count rates vary between 0 and  $6 \times 10^6$  events per second or 0 and 6 per  $\tau$ .

## 2.4. Imaging photon detectors

The most common photon counting devices, namely photomultiplier tubes, give only the time co-ordinate of photon events. The only way to observe a light field at two or more different locations using these devices is to use two or more of them, with appropriate optics. This approach is

used in optical velocimetry, where the temporal cross-correlation of intensity at two or more different positions yields fluid flow velocity measurements. Another example of this approach is the use of large numbers of photomultipliers in neutrino detectors, for instance the Japanese Kamioka detector <sup>9</sup>: in this instance a large collecting area is of just as much importance as spatial resolution.

An Imaging Photon Detector is a device which simultaneously measures the temporal and spatial co-ordinates of a photon event. The price paid for the spatial co-ordinate information is reduced temporal resolution and substantially longer dead times. The use of these detectors has so far been restricted to astronomy and high energy physics areas where the maximum information must be extracted from a weak signal.

#### **2.4.1. Time Framed Devices.**

The first generation of imaging photon detectors were developed for astronomical imaging and spectroscopy. The most sophisticated and well known of these is the IPCS (Imaging Photon Counting System), developed by Boksenberg et al <sup>10</sup>. The very earliest of these devices simply use a multistage magnetically focussed image intensifier, to amplify single photon events into high light level spots on a phosphor. An image of this phosphor is then recorded on a Plumbicon TV camera. On-line or subsequent analysis of the TV picture allows the position of each photon event to be calculated by finding the centroid of each light spot, see Figure 2.2. The intensified TV camera is still used extensively, but the image intensifier is now often of the Micro Channel Plate (MCP) type. The new generation of IPCS cameras are more advanced than this in that they use a CCD camera to image the phosphor.

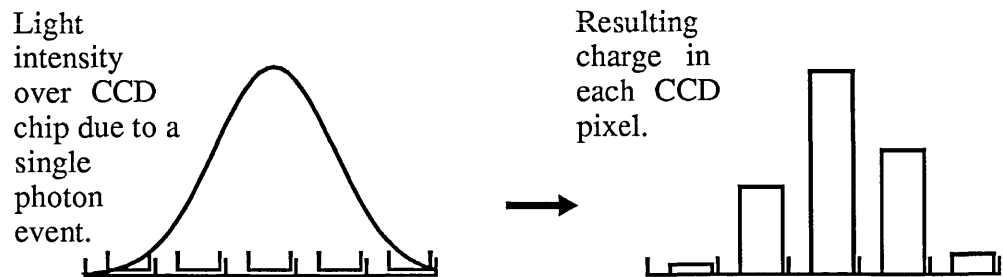


*Illustration of how the extended light envelope due to a single photon event can cover several CCD elements. The charge collected by each CCD element can be used to calculate the centroid of the light envelope to an accuracy much greater than the CCD element spacing.*

---

**Fig 2.2**

One of the reasons that this type of device remains very popular is that, because the photon event co-ordinates are computed by centroiding a Gaussian light spot which covers several scan lines, it is possible to image with high resolution using these devices. This is particularly true of the IPCS which can provide resolution of the order of 2500 pixels in one of its dimensions.

For more sophisticated data analysis this type of detector is severely limited by the fact that its temporal resolution is limited to video rates, or reasonably achievable film frame rates, that is about fifty frames per second. Again the IPCS comes out on top in this category because it is

possible to scan its CCD at higher than standard video rates, but only at the price of reducing its active area and hence spatial resolution. For some types of data processing, for instance speckle interferometry, it is possible to alleviate this problem by gating the frame exposures down to the required period. Of course this leads to loss of signal, which is undesirable when data is photon limited, since in this regime the signal-to-noise ratio is always likely to be a problem. A further problem intrinsic to framed devices, is that the arrival times of the individual photon events within the frame are not known to better than the frame time and thus the order in which the photon events were recorded is not known. For many types of data processing this loss of information leads to a further loss of signal to noise in the recovered signal. In Chapter 3 it is shown how this extra temporal information may be used to gain extra signal to noise when carrying out auto-correlation or triple-correlation processing on photon limited data.

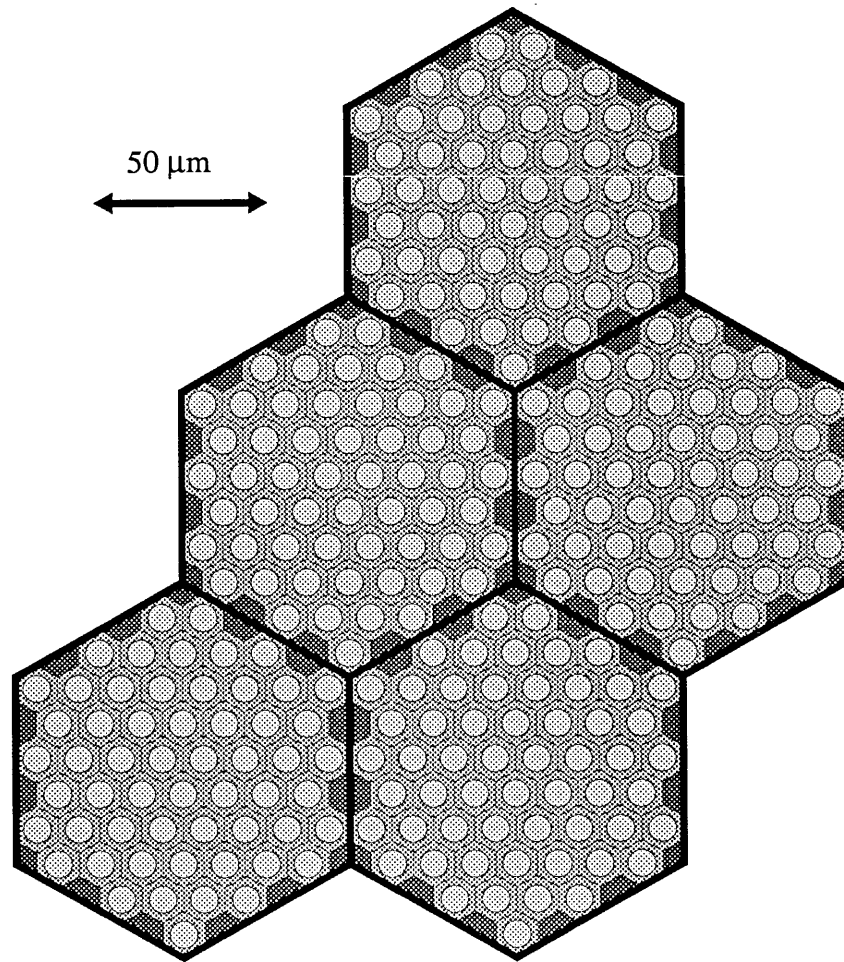
Associated with the low frame rate is a further problem, which leads to a small loss of information in direct imaging uses, but gives rise to serious problems if one attempts to apply auto-correlation or triple-correlation processing to data from this type of detector. As mentioned previously, in these detectors, each detected photon event is imaged as a roughly gaussian profiled light spot covering several scan lines of a video or CCD image. Each gaussian light spot then has a centroiding algorithm applied to it, in order to find the relevant photon event coordinate, see figure 2.2. Unfortunately this method of identifying photon events fails when two or more events occur in close proximity in the same frame. Since the frame time is relatively long, at useful data rates the probability of finding such close proximity events will not be negligible. In this case both photon

events might be counted as one, or if the condition is recognised, both events might be rejected. When carrying out auto-correlation or triple-correlation of such data one finds that because the centroiding algorithm cannot uniquely locate two photons separated by only a few pixels, the correlation contains a hole (the "photon hole") around its origin. That is one finds no correlation at zero spatial lag and reduced correlations at small spatial lags. When carrying out temporal correlation we would refer to points points close to the correlation origin as points of small temporal lag. By analogy we refer to points close to the origin of a spatial correlation as points of small spatial lag. One might think that the photon hole problem could be cured by simply cutting out the step of locating individual photon events and simply processing the intensity distributions using analogue correlation routines. Unfortunately this approach will not work because each photon in the image is weighted by an unknown weighting factor, which makes it impossible to correct for the photon bias, see Chapter 3. One way to alleviate the photon hole problem and solve the photon bias problem is to use data from two separate cameras viewing the same speckle pattern, and then process the cross correlations of the two data sets <sup>11</sup>. Obviously using this arrangement it is impossible for the same photon to appear in both frames, we can thus use analogue correlation routines without having to worry about photon bias. The price paid for this arrangement is that each camera can only receive half of the available light. Obviously this results in a loss of signal to noise (SNR); for instance with auto-correlation processing the SNR loss would be approximately a factor of  $\sqrt{2}$ .

#### 2.4.2. Time Tagging Devices.

Newer types of imaging photon counters are represented by three different types of device, namely the IPD (Imaging Photon Detector), the PAPA (Precision Analogue Photon Address) and the MAMA (Multiple Anode Microchannel Array). What characterises these detectors is the fact that they are not intrinsically framed devices, that is each photon event detected can be given both spatial and temporal co-ordinates. The temporal co-ordinates of each photon event must of course be unique, since these detectors are unable to simultaneously detect two or more photon events. Typically the dead time and hence temporal resolution of these detectors is better than  $10^{-6}$  seconds. These devices (with the possible exception of the PAPA detector) use a Microchannel Plate (MCP) as their primary amplifier.

The Micro Channel Plate functions as a light weight, compact and high gain image intensifier; its structure is illustrated in Figure 2.3. These devices are used at low gain in night sights and intensified surveillance cameras. However they can produce enough gain with low dark current to be used in photon counting detector. An MCP is manufactured by producing a large hexagonal master consisting of an array of hexagonal glass prisms, each about a centimetre or two in diameter, each containing a cylinder of different glass in its centre. The master is then stretched until it is only a couple of centimetres wide, it is then cut into small lengths and re-assembled into a large hexagon. This stretch re-assembly procedure is repeated until the original hexagonal tubes are reduced to a diameter of about  $12\text{ }\mu\text{m}$ . At this stage the resulting glass prism is either cut at an angle into thin sections if Z stacked plates are to be made, or carefully sheared and cut if C section plates are to be made. The glass



*End view of a micro-channel plate. The micro-channel plate consists of a regularly spaced array of microscopic tubes in a glass substrate. Originally formed from two different types of glass, the cylindrical prisms of glass at the centre of each small hexagon, are chemically removed in the final phase of the manufacturing process.*

**Fig 2.3**

---

which originally formed the centre of each hexagonal tube is now chemically removed, leaving a glass plate full of hexagonally packed tubes or micro-channels, on a 12  $\mu\text{m}$  pitch. Each plate is now carefully doped to make it slightly conducting and to reduce its work function.

In use a high potential difference is applied between the two faces of a plate thus setting up a steep potential gradient along the length of each of the micro-channels. Under these circumstances if a photo-electron enters one end of a micro-channel, it will be accelerated down the length of the tube; each time it collides with the side of the tube it will liberate daughter electrons which will in turn liberate more daughter electrons. The maximum charge which a single micro-channel can deliver is of the order of  $10^5 - 10^6$  electrons. When this level of amplification is achieved the gain is said to be saturated. The gain is limited by 'space charge' effects, or in other words, the effect of the mutual repulsion of the electrons in a charge cloud. This means that the gain of a MCP cannot be increased indefinitely simply by lengthening the micro-channels. Furthermore it is not possible to increase the gain by simply increasing the potential gradient along a micro-channel, since this would lead to increased current flow through the plate, and hence increased ohmic heating. Eventually thermal electron emission would lead to increased dark current and eventually a thermal runaway effect. One way to increase the available gain is to use several micro-channel plates. An example of this is the so called Z stacked plate where amplification is spread over 3 or more separate micro-channel plates. The plates are proximity "focussed" which means that the charge from one micro-channel will impinge on several micro-channels when entering the next plate. In this way the final amplification ~~final amplification~~ is spread over several adjacent micro-

channels, thus giving higher over-all gain.

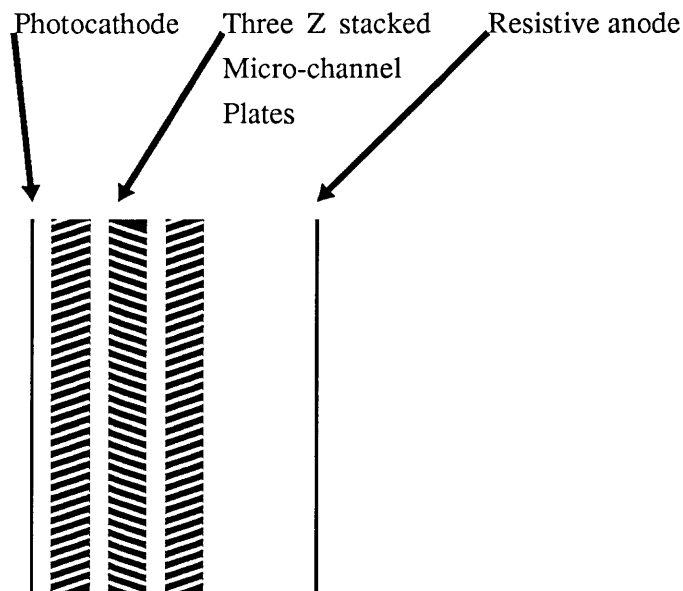
When built into a photon counting detector the first MCP usually has a separate thin photo-cathode held close to it. In order to reduce signal spread and increase the electron capture probability the photocathode is held at a positive potential of a few hundred volts with respect to the first MCP. A potential problem with the MCP device is that this photo-cathode can be very easily damaged by heavy positive ions being accelerated to high velocities in the reverse direction along the micro-channels. This is the main reason why the micro-channels are curved or arranged in Z stacks, since this arrangement reduces the probability of a positive ion attaining sufficient energy to be able to damage the photo-cathode. A further protection method sometimes used is to coat the first MCP with a very thin film of aluminium oxide, which stops heavy ions reaching the photo-cathode but allows most photo-electrons to penetrate through to the MCP.

After a micro-channel has amplified a photo-electron in saturated gain mode its sides will be depleted of electrons, thus it will be unable to immediately amplify a second photo-electron. Since the plate and micro-channels are slightly conducting a small current flows between the two sides of the plate, it is this current which must recharge a depleted micro-channel. The time taken for a micro-channel to recharge is called the local dead time. For typical micro-channel plates this local dead time is of the order of tens of milliseconds when the plates are used in saturated gain mode. One cannot reduce this dead time simply by increasing the recharge current (by increasing the voltage across the plate) since as mentioned previously this is liable to result in thermal runaway. This form of local dead time conforms quite closely to the polarisable model. When

used below saturated gain MCP amplifiers give a negative exponential pulse height distribution. When used in saturated gain mode a well behaved pulse height distribution is observed, this enables stable thresholding to be set, to distinguish single photon events from dark current and from multiple photon events. There is thus a strong incentive for using MCP amplifiers in saturated gain mode.

## 2.5. The IPD detector.

The construction of an IPD detector is illustrated in Fig 2.4. The primary amplification stage in this device is a MCP used in saturated gain mode.



*Essential components in the construction of an IPD, the photo-cathode can be up to 40mm in diameter.*

**Fig. 2.4**

---

The IPD was the first type of imaging photon counter to be developed; the detector itself is very small, about 5cm diameter and 3cm long, and mechanically quite robust.

In this type of detector the charge cloud due to an amplified photon event is sensed directly by the readout electronics. Currently at least three different techniques are available for determining the spatial co-ordinates of a charge cloud due to a photon event. These are resistive anode, wedge and strip <sup>12</sup>, or multiple anode (see MAMA detector).

The IPD detector used for the experiments reported in Chapter 4 made use of the resistive anode readout technique. The charge cloud from the final stage of a Z stacked MCP is proximity focussed onto a square uniformly resistive anode sheet. Each corner of this sheet is connected to an amplifier, the input impedances of which are arranged such that the resistive anode behaves as if it were a infinite resistive sheet. The charge from each corner of the anode is integrated to form the four quantities A,B,C,D, which are then converted into the co-ordinates of the photon event by applying the following transformations using a simple analogue computer;

$$x = \frac{A + B - (C + D)}{A + B + C + D} . \quad 2.5.19$$

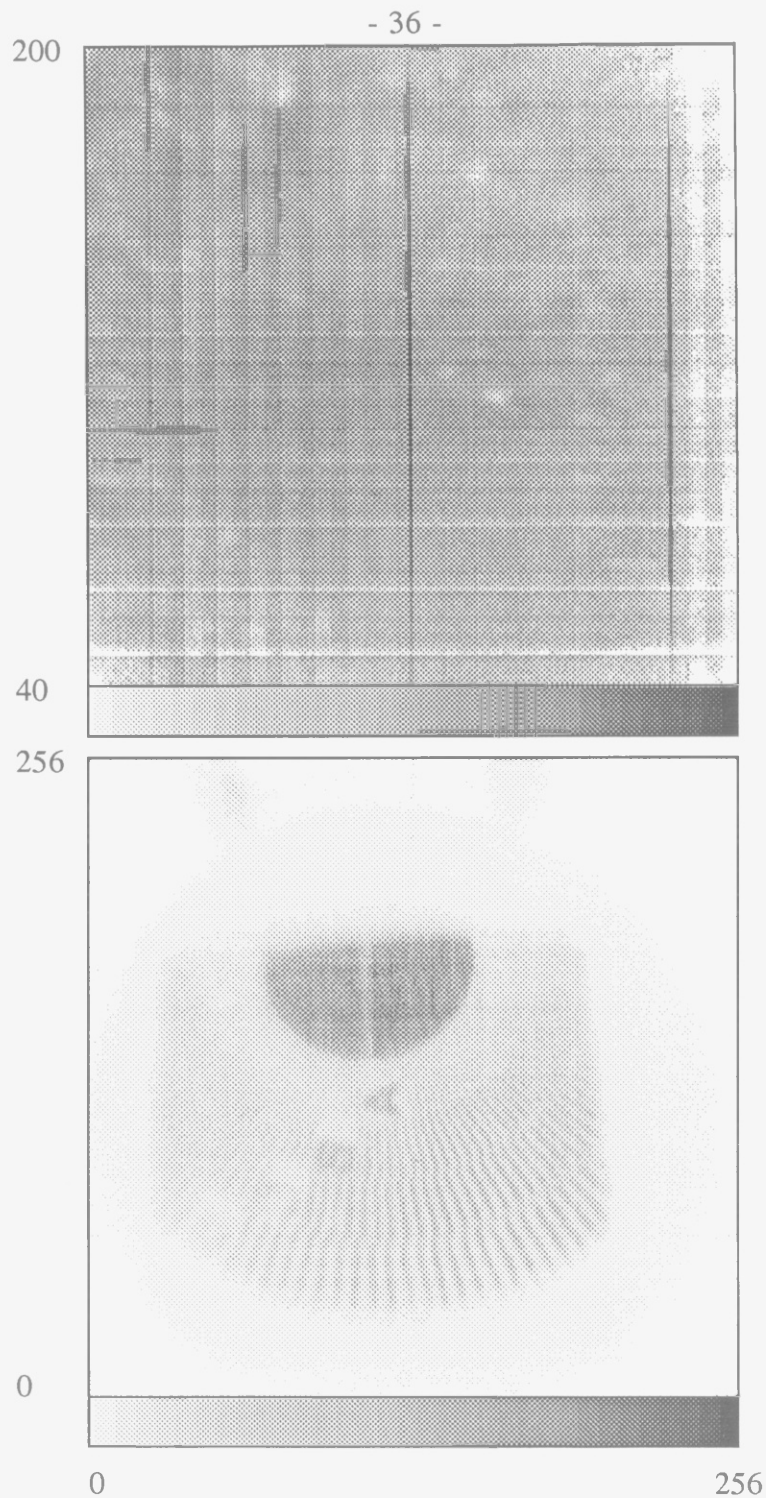
$$y = \frac{A + D - (C + B)}{A + B + C + D} . \quad 2.5.20$$

The maximum count rate for our system is of the order of  $2 \times 10^5$  counts per second, that is the global dead time is about 5 $\mu$ s, which appears to be mainly due to the limited speed of the processing electronics. As indicated in Fig. 2.5 the spatial resolution of these detectors is not very good although more modern versions are claimed to be able to resolve up to 350 pixels across a 25 mm diameter photo-cathode. The main sources of

error in position determination are lack of uniformity in the resistive anode, and electronic noise caused by shot noise and Johnson noise. Thus in order to minimise the position error in an IPD, the gain of the MCP has to be maximised; this means that a large number of micro-channels must be used in the final amplification of a photon event. A side effect of this high amplification is that a large number of micro-channels and thus a large area of the final MCP is depleted by each photon event, which means that local dead time problems are particularly severe in this type of detector. Indeed when an IPD is used to image small bright objects, for instance a star field, it is very common to see that the brightest objects have a hole in their middle. This effect is caused by local dead time which as pointed out earlier is polarisable and thus leads to a count rate approaching zero when the input intensity is high.

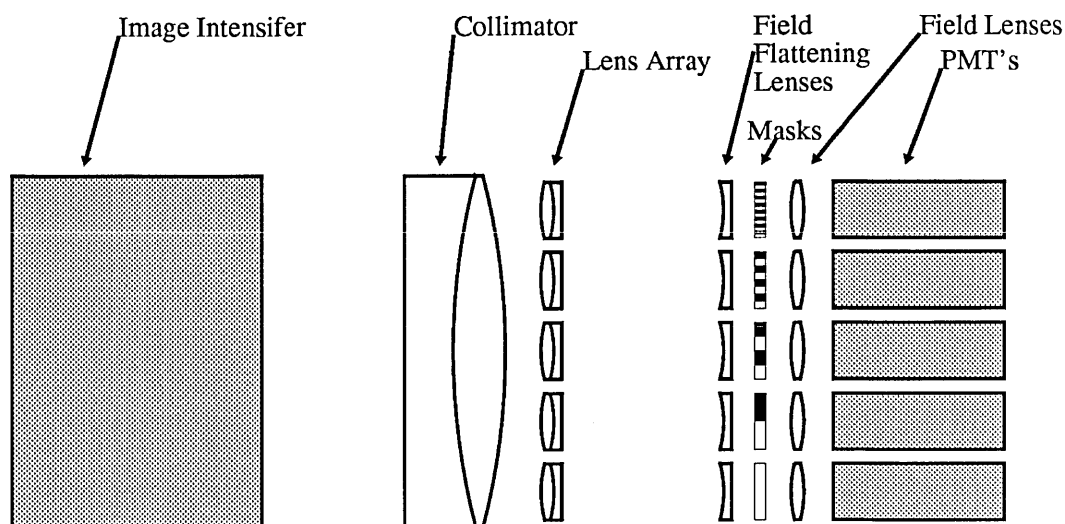
## **2.6. The PAPA detector.**

The basic construction of the PAPA detector is shown in Fig. 2.6. Either an MCP or a conventional magnetically focussed intensifier may be used as the primary amplification device. In either case the charge cloud resultant from a photon event is made to impinge upon a short time-constant phosphor, causing a brief flash of light at the photon event location. This phosphor is imaged into typically 19 separate but identical images. Each of these images is then viewed through a different grey code mask by a photomultiplier. With this arrangement the presence or absence of a signal in a photomultiplier indicates directly whether the Gray code bit encoded by the mask in front of it, is set in the position co-ordinate of the photon event. Thus a photon events co-ordinate can be read directly from the photomultipliers outputs. In a 19 photomultiplier



*The flat field (Top) of the IPD used in Chapter 4. In more modern versions of the device the cross hatching in the flat field is reduced by using Gray coded instead of binary coded analogue to digital converters. An image of a resolution test chart (Bottom) indicates that this IPD has a resolution of about 100 pixels across the 18 mm photocathode.*

**Fig. 2.5**



*Construction of the PAPA detector. This one dimensional version of the detector would have a resolution of 16 pixels.*

**Fig. 2.6.**

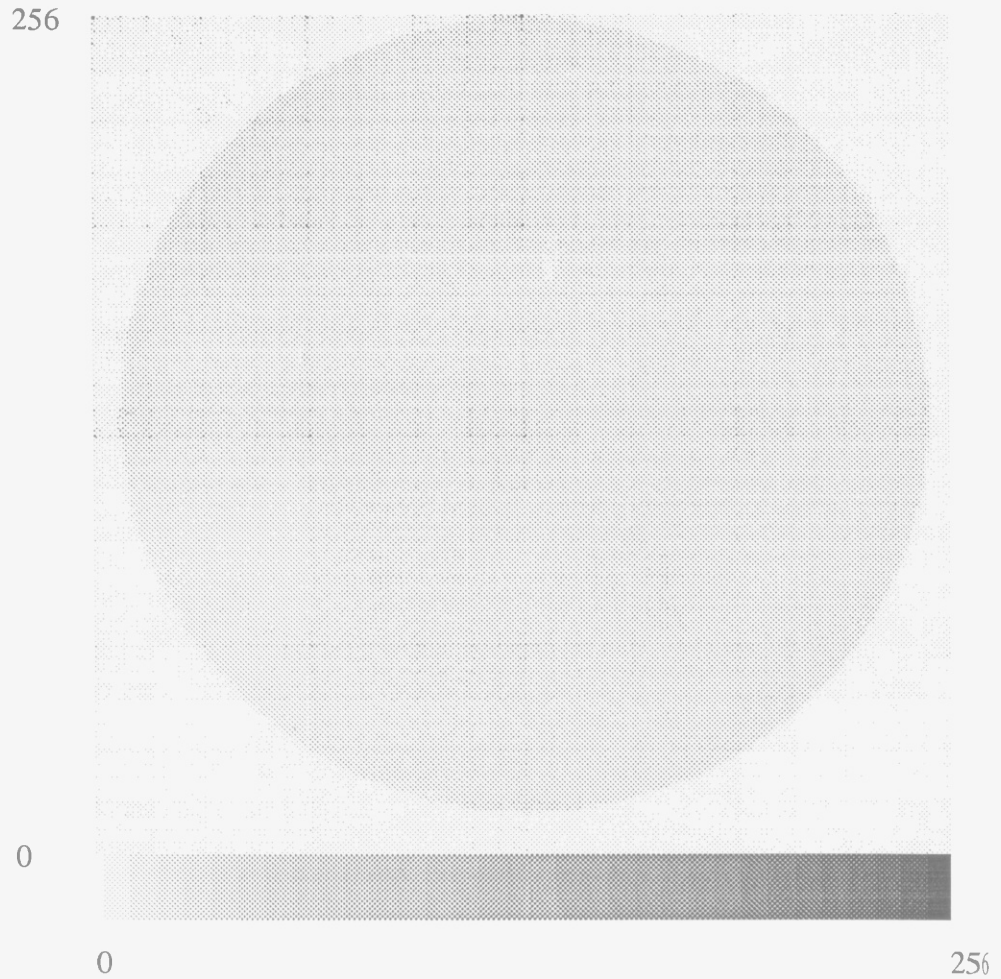
system we would have 9 grey code masks covering each spatial co-ordinate, giving a  $512 \times 512$  co-ordinate digitisation matrix. The output from the 19<sup>th</sup> photomultiplier is used for timing and thresholding purposes.

A Gray code is used instead of a binary code for the following reason. Consider a photon event which is very close to the centre of the photocathode, its binary position code should be either  $256 = 100000000$  binary, or  $255 = 011111111$  binary depending upon which side of the most significant mask it falls. If the photon-event should fall exactly in the middle of the detector the light spot due to the photon event would fall on an edge in every detector. If the masks were perfectly aligned but the thresholding were not set well, this would give a co-ordinate of  $511 = 111111111$  binary, which is obviously completely erroneous. In general,

taking into account the effects of mask misalignment, a "random" event co-ordinate would be produced. Since the pulse height distribution of the PAPA detector's image intensifier is negative exponential, it is difficult to threshold out such bad co-ordinates. With binary coded masks a similar ambiguity occurs at every mask edge. In a Gray coded mask set no two mask boundaries fall in the same position, therefore the maximum positional error that this effect could produce would be plus or minus a least significant bit.

If an MCP is used in a PAPA detector it is generally used with unsaturated gain, which virtually eliminates local dead time problems, but leads to a poor exponential pulse height distribution. It is therefore very difficult to efficiently threshold out dark current events in a PAPA detector. This variable pulse height is the main reason for having an unmasked photomultiplier in the PAPA detector <sup>13</sup>. ~~It is~~ <sup>It is</sup> used to threshold and strobe the whole detector. If an event has been accepted its pulse height is measured by the unmasked photomultiplier, this value is then used to dynamically alter the threshold of the remaining photomultipliers. The purpose of this dynamic thresholding is to increase the detector's resistance to photomultiplier dark current noise. It is important to run this detector well below its maximum count rate since, because thresholding is so problematical, it is difficult to determine if a photon event is in fact composed of two simultaneous or near simultaneous photon events at different spatial locations. It is obviously important to either detect these double events, or to run at a photon event rate where they are rare, since the co-ordinate returned by such an event will be completely meaningless.

At present the main limitation on the speed of the PAPA detector is set by the decay time of the phosphor screen; it is in the region of



*A typical flat field for a PAPA detector. In this case the cross hatch pattern is due to slight mask misalignments. Due to the difficulty in thresholding, the flat field of this type of detector is quite unstable.*

**Fig. 2.7**

$10^6 \rightarrow 10^7$  events per second.

## **2.7. The MAMA detector.**

The MAMA detector is fundamentally very similar to the IPD detector, the main distinguishing feature being the use of a much more

elaborate position readout system, which gains both higher spatial resolution and reduced global dead time. Present versions of the detector have resolutions of  $256 \times 1024$  and  $1024 \times 1024$  and it is hoped eventually to develop a detector with resolution  $2048 \times 2048$  <sup>14</sup>. In a MAMA detector the charge cloud from a photon event in an MCP intensifier impinges upon a grid of electrodes; these electrodes are wired in such a way that only  $2 \cdot (\sqrt{n} + \sqrt{m})$  amplifiers are needed to determine the position of a photon event, where  $m$  and  $n$  are digital resolution of each side of the detector. The first versions of the MAMA detector <sup>15</sup> use single stage 'C' MCP plates, but there is no reason why multi-plate amplifiers could not be used, although this might exacerbate the local dead time problem.

On present devices the maximum count rate appears to be about  $10^6$  events per second, limited by the speed of the electronics.

## **2.8. The future; avalanche photo-diodes.**

There is little doubt that the future of imaging photon detectors lies in the developing technology of arrays of avalanche photo-diodes. Theoretically the maximum quantum efficiency of a vacuum tube device using a conventional photocathode is 50%. This is because a photoelectron has a 50% probability of emerging each side of the photocathode and thus a 50% probability of entering the photon counters gain mechanism, which for example would consist of a dynode chain or MCP. In contrast the depletion region of an avalanche photo-diode combines in the same piece of material both the detectors gain mechanism and the functionality of a photocathode. Thus their efficiency can in theory, be greater than 50%. Indeed real quantum efficiencies of 35% have been measured in these devices. The types of image intensifiers used on conventional

imaging photon detectors typically have a quantum efficiency of the order of 5%. An array of 50% efficient photo-diodes would thus give a signal level ten times higher than any of the previously described systems. Equivalently it would, for example, allow astronomical speckle interferometry to be carried out on objects at least 2.5 magnitudes fainter than is possible at present.

With each element able to operate at  $10^6$  counts per second the potential data rate from such a detector would be truly staggering. Unfortunately it will probably be many years before large arrays of these devices become available.

Studies of single avalanche photodiodes by R. Brown at RSRE Malvern has shown that if cooled they can offer comparable performances to good photon counting photo-multiplier tubes for photon correlation measurements. In some respects they perform better; for instance they do not suffer from significant after-pulsing. At present however their dead time, which is of the order of  $10^{-6}$  seconds cannot match that of a good photomultiplier. One of the most exciting aspects of this technology is that in the future, avalanche photo-diodes are likely to be capable of photon counting in the near infrared, allowing the advantages of photon counting processing to be brought to this wavelength range.

### 3

## The Triple Correlation.

### 3.1. Notation.

In an attempt to keep the meaning of equations as clear and unambiguous as possible, the following conventions have been adopted. Lower case letters will be used to represent functions in the spatial or temporal domains and the same upper case letters will be used to represent the Fourier transforms of these functions (in the frequency domain). We will generally use the letters  $x, y, t$ , to represent co-ordinates in the spatial domain, and the letters  $u, v, \nu$ , to represent co-ordinates in the frequency domain. Any quantity represented in bold type should be taken to represent a vector and symbols in ordinary type represent real or complex scalars. The order of a correlation function or its Fourier transform will be indicated a bracketed numerical superscript.

### 3.2. Relationship to the auto-correlation.

The auto-correlation  $i^{(2)}(x_1)$  of a signal  $i(x)$  is defined as:

$$i^{(2)}(x_1) = \int_{-\infty}^{\infty} i^*(x)i(x+x_1)dx. \quad 3.2.1$$

By analogy the triple-correlation  $i^{(3)}(x_1, x_2)$ , (from now on referred to as TC) is defined as follows :

$$i^{(3)}(x_1, x_2) = \int_{-\infty}^{\infty} i^*(x)i(x+x_1)i(x+x_2)dx \quad 3.2.2$$

Note that the TC of a function has double the dimensionality of the original function. Thus for instance, the TC of a two dimensional image, would occupy a four dimensional space.

One of the basic properties shared by the auto-correlation and TC, and indeed all higher order correlations, is their invariance to linear translation of the co-ordinate system. In other words since we integrate with respect to the vector  $\mathbf{x}$  over all space, we can add any constant offset to  $\mathbf{x}$  without affecting the value of the correlation. We shall be making use of this translational invariance property in chapter four.

The Fourier transform of the triple correlation of a complex function  $i(\mathbf{x})$ , which will be written as  $I^{(3)}(\mathbf{u}, \mathbf{v})$ , is perhaps rather unfortunately called the bispectrum:

$$I^{(3)}(\mathbf{u}, \mathbf{v}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^{(3)}(\mathbf{x}_1, \mathbf{x}_2) e^{-i\mathbf{x}_1 \mathbf{u}} e^{-i\mathbf{x}_2 \mathbf{v}} d\mathbf{x}_1 d\mathbf{x}_2. \quad 3.2.3$$

It is straightforward to show that the following simple relationship exists, between the bispectrum  $I^{(3)}(\mathbf{u}, \mathbf{v})$ , and Fourier transform  $I(\mathbf{u})$ , of a function  $i(\mathbf{x})$  (see Appendix B.);

$$I^{(3)}(\mathbf{u}, \mathbf{v}) = I(\mathbf{u}) I(\mathbf{v}) I^*(\mathbf{u} + \mathbf{v}). \quad 3.2.4$$

This result is analogous to the Parseval's relationship between the power spectrum and Fourier transform of a complex function;

$$I^{(2)}(\mathbf{u}) = I(\mathbf{u}) I^*(\mathbf{u}). \quad 3.2.5$$

It is instructive to write down the phase of the power spectrum and the bispectrum in terms of the Fourier phase of the signal. We represent the phase of the a spectrum by  $\beta$ , that is  $\beta(\mathbf{u}) = \text{phase of } I(\mathbf{u})$  and  $\beta^{(3)}(\mathbf{u}, \mathbf{v}) = \text{phase of } I^{(3)}(\mathbf{u}, \mathbf{v})$ . For the power spectrum we have;

$$\beta^{(2)}(\mathbf{u}) = \beta(\mathbf{u}) - \beta(\mathbf{u}) \equiv 0, \quad 3.2.6$$

and for the bispectrum;

$$\beta^{(3)}(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{u}) + \beta(\mathbf{v}) - \beta(\mathbf{u} + \mathbf{v}). \quad 3.2.7$$

It is evident from the above equations that the phase of the power-spectrum is always identically zero; this means that we can only recover the modulus of the Fourier transform of a function from its power spectrum or auto-correlation. However it is evident from Eqn. 3.2.7, that the phase of the bispectrum, is in general non-zero. It is from this basic difference in behaviour that the utility of the bispectrum arises. We will show later that it is possible to derive the Fourier phase of a function from the phase of its bispectrum. Thus, while the power spectrum retains only the Fourier modulus of the signal, the bispectrum contains **complete** information about the signal, *subject to noise and an uncertainty in absolute position.*

If the original function  $i(\mathbf{x})$  is real then its Fourier transform must be hermitian, that is  $I^*(\mathbf{u}) \equiv I(-\mathbf{u})$ . Using this fact we can rewrite Eqn. 3.2.4, for a **real** function  $i(\mathbf{x})$  only;

$$I^{(3)}(\mathbf{u}, \mathbf{v}) = I(\mathbf{u})I(\mathbf{v})I(-\mathbf{u} - \mathbf{v}), \quad 3.2.8$$

or,

$$\beta^{(3)}(\mathbf{u}, \mathbf{v}) = \beta(\mathbf{u}) + \beta(\mathbf{v}) + \beta(-\mathbf{u} - \mathbf{v}). \quad 3.2.9$$

It is interesting to note that in this case the vector co-ordinates of each of the Fourier terms which make up the bispectrum form a closed triangle in the Fourier plane. This has important repercussions when we consider the use of the triple-correlation for imaging through turbulence this is explained more fully in Chapter 6 where the concept of phase closure is discussed.

### 3.3. Symmetry of the triple correlation.

From equation 3.2.2 we see that the TC is invariant under the following co-ordinate transformation.

$$x_1 \rightarrow x_2 \qquad x_2 \rightarrow x_1. \qquad 3.3.1$$

This implies 2 fold symmetry in the correlation.

In addition if and only if the signal is **real**, a further, but more complicated symmetry condition pertains. Making the substitution  $x \rightarrow x - x_1$  in equation 3.2.2 and remembering that since  $i(x)$  is real it is its own complex conjugate,

$$\begin{aligned} i^{(3)}(x_1, x_2) &= \int_{-\infty}^{\infty} i(x - x_1) i(x) i(x + x_2 - x_1) dx \\ &\equiv i^{(3)}(-x_1, x_2 - x_1), \end{aligned} \qquad 3.3.2$$

giving the TC invariance under the transformation:

$$x_1 \rightarrow -x_1 \qquad x_2 \rightarrow x_2 - x_1. \qquad 3.3.3$$

The above symmetries imply six fold redundancy in the TC of a real function, see Figures 3.1 and 3.2.

### 3.4. Symmetry of the bispectrum.

It is obvious that the bispectrum, in common with the TC, possesses the simple symmetry:

$$u \rightarrow v \qquad v \rightarrow u. \qquad 3.4.4$$

Again if a function is **real**, its Fourier transform is hermitian. In this case the functions bispectrum displays two further symmetry conditions.

### Symmetry of Triple Correlation & Bispectrum

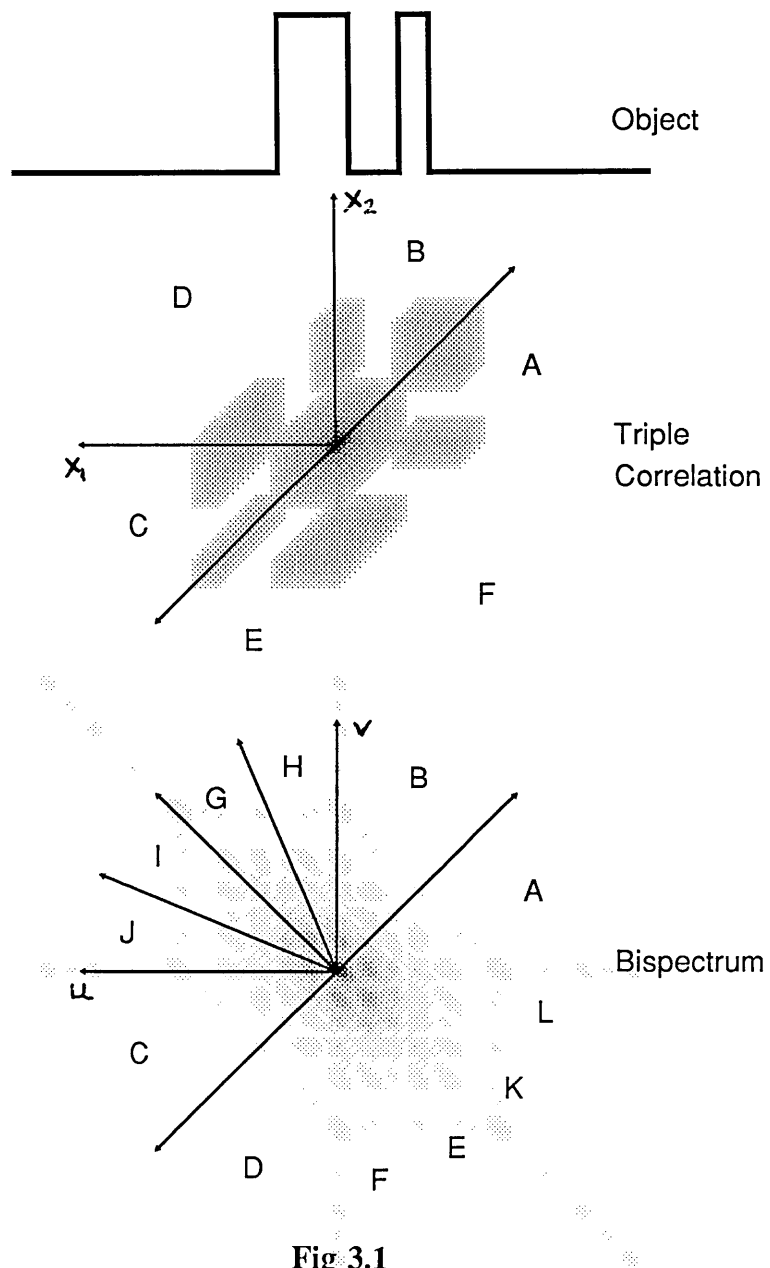


**Fig 3.1**

*Symmetry of the TC and bispectrum of a simple one dimensional signal, regions of symmetry are denoted by heavy lines and given labels A through K.*

| Symmetry of the TC.                                   |                     |    |    |    |                    |     |
|-------------------------------------------------------|---------------------|----|----|----|--------------------|-----|
| Transform                                             | Co-ordinate Mapping |    |    |    | Sector on Fig 3.1. |     |
|                                                       | 2                   | 1  | 3  | 4  | x                  | y   |
| $x_1 \rightarrow x_2$<br>$x_2 \rightarrow x_1$        | 1                   | 2  | 4  | 3  | y                  | x   |
| $x_1 \rightarrow -x_1$<br>$x_2 \rightarrow x_2 - x_1$ | -2                  | -1 | -3 | 1  | -x                 | y-x |
|                                                       | -1                  | 1  | -4 | -1 | -y                 | x-y |
| $x_1 \rightarrow x_2$<br>$x_2 \rightarrow x_1$        | -1                  | -2 | 1  | -3 | y-x                | -x  |
|                                                       | 1                   | -1 | -1 | -4 | x-y                | -y  |

The first column in the table shows the co-ordinate transforms under which the triple correlation is invariant. The second and third columns trace a pair of example co-ordinates through repeated application of transformations. The fourth column traces a generalised co-ordinate through the transformations. The fifth column shows how the region labelled A on Fig. 3.1 is transformed into the other labelled regions.

Fig. 3.2

These are;

$$u \rightarrow -u \qquad v \rightarrow -v. \qquad 3.4.5$$

and

$$u \rightarrow u \qquad v \rightarrow -u-v. \qquad 3.4.6$$

These three symmetry conditions together, give the bispectrum of a real function twelve fold redundancy, see figures 3.1 and 3.3.

### 3.5. Calculation of the TC using photon differencing.

At low light levels we are dealing with a signal which consists of discrete photo-events. Each photo-event will have a certain time and position co-ordinate associated with it. In the following sections we will

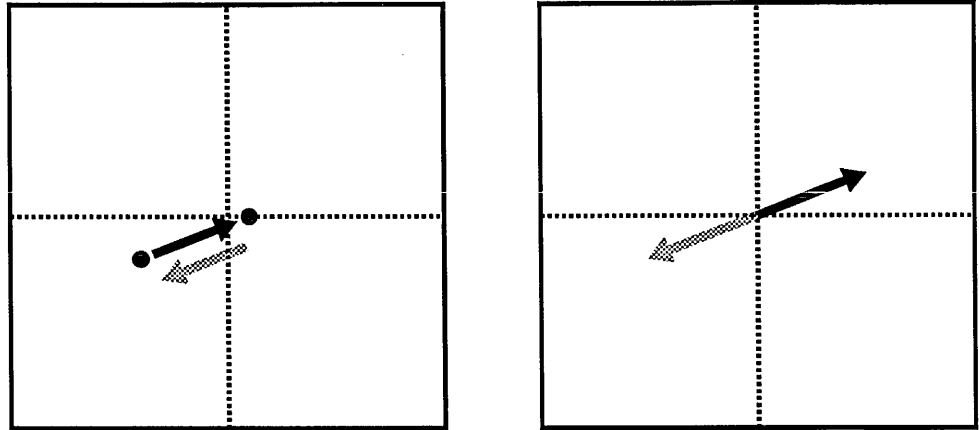
| Symmetry of the bispectrum.                    |                     |    |    |    |                    |                 |                                  |
|------------------------------------------------|---------------------|----|----|----|--------------------|-----------------|----------------------------------|
| Transform                                      | Co-ordinate Mapping |    |    |    | Sector on Fig 3.1. |                 |                                  |
|                                                | 2                   | 1  | 3  | 4  | <b>u</b>           | <b>v</b>        | A                                |
| <b>u</b> → <b>v</b><br><b>v</b> → <b>u</b>     | 1                   | 2  | 4  | 3  | <b>v</b>           | <b>u</b>        | A → B                            |
| <b>u</b> → - <b>u</b><br><b>v</b> → - <b>v</b> | -2                  | -1 | -3 | -4 | - <b>u</b>         | - <b>v</b>      | A → C<br>B → D                   |
| <b>u</b> → <b>u</b><br><b>v</b> → - <b>u-v</b> | 2                   | -3 | 3  | -7 | <b>u</b>           | -( <b>u+v</b> ) | A → E<br>B → F<br>C → G<br>D → H |
|                                                | 1                   | -3 | 4  | -7 | <b>v</b>           | -( <b>u+v</b> ) |                                  |
|                                                | -2                  | 3  | -3 | 7  | - <b>u</b>         | ( <b>u+v</b> )  |                                  |
|                                                | -1                  | 3  | -4 | 7  | - <b>v</b>         | ( <b>u+v</b> )  |                                  |
| <b>u</b> → <b>v</b><br><b>v</b> → <b>u</b>     | -3                  | 2  | -7 | 3  | -( <b>u+v</b> )    | <b>u</b>        | E → I<br>F → J<br>G → K<br>H → L |
|                                                | -3                  | 1  | -7 | 4  | -( <b>u+v</b> )    | <b>v</b>        |                                  |
|                                                | 3                   | -2 | 7  | -3 | ( <b>u+v</b> )     | - <b>u</b>      |                                  |
|                                                | 3                   | -1 | 7  | -4 | ( <b>u+v</b> )     | - <b>v</b>      |                                  |

*The first column in the table shows the co-ordinate transforms under which the bispectrum is invariant. The second and third columns trace a pair of example co-ordinates through the transformations. The fourth column traces a generalised co-ordinate through the transformations. The fifth column shows how the region labelled A on Fig. 3.1 is transformed into the other labelled regions.*

**Fig 3.3**

assume that the stream of photons is divided into a series of separate time-slices or frames.

It is well known that the auto-correlation of a series of such frames, can be calculated by the "photon-differencing" method<sup>16</sup>. Photon differencing refers to the process of constructing a histogram of all possible co-ordinate difference vectors, between all photons in a single data frame. We now show that apart from a trivial bias term, which results from self-correlation of photons, simple photon differencing yields the



*Two photons from a photon limited frame. Their difference vectors are indicated by the arrows.*

*Increment the auto correlation histogram at the points indicated by the difference vectors.*

*This illustration shows graphically what the difference vectors of two photo events represent. It shows how these difference vectors are used to determine an auto-correlation. In this illustration the object array and auto-correlation arrays are shown as being the same size. Of course in general the auto-correlation array must be twice the size of object array.*

Fig. 3.4

auto-correlation of the input frame.

$$i^{(2)}(\mathbf{x}_1) = \int_{-\infty}^{\infty} i^*(\mathbf{x})i(\mathbf{x}+\mathbf{x}_1)d\mathbf{x}. \quad 3.5.1$$

We write a single photon limited frame as  $\sum_{j=1}^{j=N} \delta(\mathbf{x}-\mathbf{x}_j)$  ( see Chapter 1 ).

Substituting this in eqn. 3.5.1 :

$$d^{(2)}(x_1) = \int_{-\infty}^{\infty} \sum_{j=1}^{j=N} \delta(x-x_j) \sum_{k=1}^{k=N} \delta(x+x_1-x_k) dx \quad 3.5.2$$

Note that in making this substitution we have made the assumption that

$\sum_{j=1}^{j=N} \delta(x-x_j)$  is real, which in the case of a photon limited image is certainly

justified. The integrand of eqn. 3.5.2 is unity when:

$$x = x_j. \quad 3.5.3$$

and

$$x = x_k - x_1. \quad 3.5.4$$

For all other values of  $x$  the integrand is zero.

Substitution of eqn 3.5.3 in eqn 3.5.4 gives as the non-zero condition:

$$x_1 = x_k - x_j \quad 3.5.5$$

Thus we conclude:

$$d^{(2)}(x_1) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \delta(x_1 - (x_k - x_j)). \quad 3.5.6$$

$d^{(2)}(x_1)$  in equation 3.5.6 consists of two types of terms.

$$N(N-1) \text{ terms where; } j \neq k \quad 3.5.7$$

$$N \text{ terms where; } j = k \quad 3.5.8$$

The terms in 3.5.7 give us an unbiased estimator of the auto-correlation. The terms in 3.5.8 are the so-called photon bias terms<sup>17</sup>. Thus the correct expression for the bias free, photon differencing auto-correlation is;

$$i^{(2)}(x_1) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \delta(x_1 - (x_k - x_j)) (1 - \delta_{jk}). \quad 3.5.9$$

where  $\delta_{jk}$  is the Kronecker delta.

The autocorrelation of a real function has two fold symmetry  $i^{(2)}(\mathbf{x}) = i^{(2)}(-\mathbf{x})$ . This symmetry can be interpreted in a rather interesting way in the light of the photon differencing algorithm. Notice from Fig. 3.4 and Eqn. 3.5.9 that each pair of photon events give rise to two vector differences, a negative and a positive going vector, where we arbitrarily define a positive going vector as a vector lying in the half plane defined by  $\frac{\pi}{4} < \arg(\mathbf{x}) \leq \frac{3\pi}{4}$ . Each vector obviously contains exactly the same information, hence the two fold symmetry of the auto-correlation.

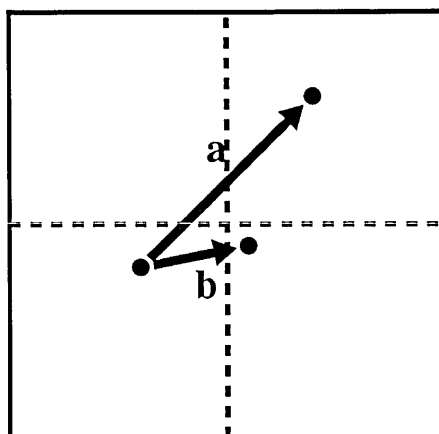
The number of operations required to compute the auto-correlation of a data frame scales as the square of the number of photons in that frame. Photon differencing is thus a very efficient way of calculating auto-correlations at low light level, but rapidly becomes less efficient as the number of photon events per frame increases. By similar reasoning it is straight-forward to show that the bias free TC of a photon-limited data frame, is given by:

$$i^{(3)}(\mathbf{x}_1, \mathbf{x}_2) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \sum_{l=1}^{l=N} (\delta(\mathbf{x}_1 - (\mathbf{x}_j - \mathbf{x}_k)) \delta(\mathbf{x}_2 - (\mathbf{x}_j - \mathbf{x}_l))) \times \quad 3.5.10$$

$$(1 - \delta_{jk}) (1 - \delta_{jl}) (1 - \delta_{kl}).$$

This equation represents the modified photon differencing procedure which is illustrated in Fig. 3.5.

Given three photon events, there are six different combinations of vector difference pairs (bivectors) which we may construct. A knowledge of any one of these bivectors would be sufficient to allow us to reconstruct the relative spatial arrangement of the original photon events, and thus



*For a two dimensional function the triple correlation difference vector is four dimensional; we call it a bivector because it is actually composed of 2, two dimensional difference vectors.*

*In this example of a bivector, vector **a** would be used as an offset into the first two dimensions of the TC histogram, vector **b** would then be used as the offset in the third and fourth dimensions of the histogram.*

**Fig. 3.5**

generate the other five bivectors. Thus by analogy with the the auto-correlation case, we can see that this implies that the TC of a real function must have six fold symmetry. This is exactly what we have already shown in Section 3.3.

The number of operations required to compute the TC by the photon differencing method increases as the cube of the number of photons in each frame. Thus the efficiency of TC determination by photon differencing falls off even more rapidly than it does for the autocorrelation. However at low light levels this method can be extremely fast by virtue of the fact that photon differencing needs only integer arithmetic operations, which are generally considerably faster than their floating point counterparts.

An unfortunate restriction in the use of this method is that the computational device used to compute the triple correlation must contain a sufficient number of storage elements to accumulate at least a non-redundant sixth of the TC. When working with one dimensional data, this is generally not a problem. However for a two dimensional signal we must in general accumulate at least a sixth of the four dimensional triple correlation. For example if we were to use a digital computer to calculate the TC of a two dimensional object sampled on a 128 square lattice, we would require  $(128 \times 2)^4$  or 4 Giga-words of memory; even using the six fold symmetry of the TC would reduce this only to 670 M-words, or typically 2.7Gbytes if using 32 bit integers. Note that just as in the case of the auto-correlation, if our starting function is sampled on a  $n \times n$  array, its TC must be placed in a  $2n \times 2n \times 2n \times 2n$  array in order to avoid aliasing problems.

At the expense of having to carry out considerably more calculation, it is possible to reduce the amount of memory required to a reasonable level; the way this is done is explained in Chapter five, in the sub-plane section 5.4.

Provided enough memory is available the photon differencing method should be the method of choice for calculating the TC at low light levels. In practice this means at present that simple photon differencing is only useful in one dimension at extremely low light levels.

### 3.6. Photon bias

We mentioned in the previous section that with present computing technology, it is usually impractical to compute the whole of the TC because of its sheer size. In practice therefore methods other than simple

photon differencing must be used (see Chap.5.). Most alternative methods rely on averaging the bispectrum rather than the TC. Because of the way in which the reconstruction algorithm works it is possible to use only a very small subsection of the bispectrum. Indeed because of signal-to-noise considerations most of the bispectrum is of no use for image reconstruction<sup>18,19</sup>. When the bispectrum is averaged, removal of photon bias is more problematical than when photon differencing is carried out. In practice this means that extra calculations are needed to enable us to remove explicitly the photon bias.

If all the photons in an input frame have unity weight, then the bias terms take on a particularly simple form. However practical imaging photon detectors invariably have uneven gain over their active area, necessitating a flat field correction before processing. The simplest way to apply a flat field correction is to weight each photon inversely to the detector sensitivity at its co-ordinates; under these circumstances the photon bias correction becomes a little more complex.

We could of course apply a flat field correction by accepting or rejecting a photon event at a particular co-ordinate, with a probability inversely dependent to the flat field response of the detector at that co-ordinate. Although this approach keeps dead-time correction simple, it reduces the signal to noise of any recovered spectra by a greater extent than the weighting method, because the information carried by the rejected photons is not used in any way.

First we will show the effect of the photon bias when we average the power spectra of photon limited frames. Since the Fourier transform is a linear operator, it follows that the Fourier transform of a photon limited frame is simply the sum of the Fourier transforms of the individual photon

events in that frame. In the most general case each photon event is represented by a weighted delta function and its Fourier transform will be a simple weighted exponential, ie;  $FT (w_j \delta(\mathbf{x}-\mathbf{x}_j)) = w_j e^{-i2\pi \mathbf{u} \mathbf{x}_j}$ . Thus the Fourier transform of a photon limited frame (see Chapter 1) can be written as;

$$I_w(\mathbf{u}) = \sum_{j=1}^{j=N} w_j e^{-i\mathbf{u} \mathbf{x}_j}, \quad 3.6.1$$

where  $w_j$  is the weight of the  $j^{\text{th}}$  photon event due to flat field correction and the subscript  $w$  on the Fourier transform of  $i(\mathbf{x})$ ,  $I_w(\mathbf{u})$ , is used to indicate that the Fourier transform is calculated from a frame of weighted photon events.

Thus the power spectrum of this weighted photon event frame will be;

$$D_w^{(2)}(\mathbf{u}) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j w_k e^{-i\mathbf{u} \mathbf{x}_j} e^{i\mathbf{u} \mathbf{x}_k} \quad 3.6.2$$

As in the analysis of photon differencing in the previous section, we can identify two types of terms in the above expression.

We have  $N(N-1)$  terms where  $j \neq k$ . These terms as, in the photon differencing analysis, give the unbiased estimator of the power-spectrum.

There are  $N$  terms where  $j=k$ , which constitute the photon bias. These are the same terms as those we do not compute when using a photon differencing algorithm.

From this identification of the terms it follows that;

$$D_w^{(2)}(\mathbf{u}) = A I_w^{(2)}(\mathbf{u}) + \sum_{j=1}^{j=N} w_j^2, \quad 3.6.3$$

where the constant  $A = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j (w_k) (1 - \delta_{jk})$ .

If we are compute the spectra of a set of photon limited frames directly

we have to explicitly subtract the bias term from each spectrum estimate<sup>20</sup>, making photon bias correction slightly more difficult in this case. Thus the unbiased estimator for the power-spectrum of a weighted photon limited frame is;

$$A I^{(2)}(\mathbf{u}) = D_w^{(2)}(\mathbf{u}) - \sum_{j=1}^{j=N} w_j^2. \quad 3.6.4$$

When all  $w_j \equiv 1$  the form of the unbiased estimator simplifies to;

$$N(N-1)I^{(2)}(\mathbf{u}) = D^{(2)}(\mathbf{u}) - N. \quad 3.6.5$$

We can follow an analogous process to calculate the unbiased estimator for the bispectrum. Rewriting Eqn. 3.6.2 for the case of the bispectrum we get ;

$$D^{(3)}(\mathbf{u}, \mathbf{v}) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \sum_{m=1}^{m=N} w_j w_k w_m e^{-i\mathbf{u}\mathbf{x}_j} e^{-i\mathbf{v}\mathbf{x}_k} e^{i(\mathbf{u}+\mathbf{v})\mathbf{x}_m}. \quad 3.6.6$$

This equation gives rise to three distinct type of terms, they are;

$N(N-1)(N-2)$  terms where  $j \neq k \neq m$  (a)

$N(N-1)$  terms where  $j \neq k = m$  (b)

$N$  terms where  $j = k = m$  (c)

As before terms of type (a) give the unbiased estimator of the bispectrum.

Terms of type (b) and (c) are both photon bias terms.

Terms of type (c) simply give rise to a bias contribution equal to  $\sum_{j=1}^{j=N} w_j^3$ .

Terms of type (b) give rise to three subsets of terms having the following form;

$$k = m \quad \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j e^{-i\mathbf{u}\mathbf{x}_j} w_k^2 e^{i\mathbf{u}\mathbf{x}_k} (1 - \delta_{jk}). \quad 3.6.7$$

$$j = m \quad \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_k e^{i\mathbf{v}\mathbf{x}_k} w_m^2 e^{-i\mathbf{v}\mathbf{x}_m} (1 - \delta_{km}). \quad 3.6.8$$

$$j = k \quad \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j^2 e^{-i(u+v)x_j} w_m e^{i(u+v)x_m} (1 - \delta_{jm}). \quad 3.6.9$$

The  $(1 - \delta)$  terms are included to exclude terms of type (c) from here.

By analogy with equation 3.6.1 which expresses the Fourier transform of a weighted photon limited frame, terms of the form  $\sum_{j=1}^{j=N} w_j^2 e^{-iux_j}$  can be seen

to represent the Fourier transform of the the squared weights of the photon events in a frame. To calculate the Fourier transform of a frame, for all photon events in that frame, add the weight of the photon event to the image buffer at the position of that photon event. When all photon events have been accumulated Fourier transform the image buffer. In order to calculate  $\sum_{j=1}^{j=N} w_j^2 e^{-iux_j}$ , we follow exactly the same procedure, but add the squared weights of each photon event to the image buffer.

We choose to denote this quantity using the following notation;

$$\sum_{j=1}^{j=N} w_j^2 e^{-iux_j} = I_{w^2}(\mathbf{u}). \quad 3.6.10$$

Using this notation the unbiased estimator of the bispectrum becomes;

$$\begin{aligned} B \ I^{(3)}(\mathbf{u}, \mathbf{v}) &= D^{(3)}(\mathbf{u}, \mathbf{v}) \\ &- C \left[ I_w(\mathbf{u}) I_{w^2}^*(\mathbf{u}) - I_w(\mathbf{v}) I_{w^2}^*(\mathbf{v}) - I_{w^2}^*(\mathbf{u}+\mathbf{v}) I_w(\mathbf{v}) \right] \\ &- \sum_{j=1}^{j=N} w_j^3, \end{aligned} \quad 3.6.11$$

where;

$$B = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \sum_{m=1}^{m=N} w_j w_k w_m (1-\delta_{jk})(1-\delta_{km})(1-\delta_{jm}) \quad 3.6.12$$

and

$$C = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j w_k^2 (1-\delta_{jk}). \quad 3.6.13$$

If we try to compute the bias terms of type (b), i.e. terms of the form  $I_w(\mathbf{u})I_{w^2}^*(\mathbf{u})$ , it immediately becomes apparent that these terms themselves suffer from photon bias. However it is easy to see that;

$$B I_w(\mathbf{u})I_{w^2}^*(\mathbf{u}) = D_w(\mathbf{u})D_{w^2}^*(\mathbf{u}) - \sum_{j=1}^{j=N} w_j^3, \quad 3.6.14$$

where,

$$D_w(\mathbf{u})D_{w^2}^*(\mathbf{u}) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} w_j e^{-i\mathbf{u}\cdot\mathbf{x}_j} w_k^2 e^{i\mathbf{u}\cdot\mathbf{x}_k}. \quad 3.6.15$$

We are finally able to write down the expression for the unbiased estimator for the bispectrum:

$$\begin{aligned} B I^{(3)}(\mathbf{u}, \mathbf{v}) &= D^{(3)}(\mathbf{u}, \mathbf{v}) - \\ &D_w(\mathbf{u})D_{w^2}^*(\mathbf{u}) - D_w(\mathbf{v})D_{w^2}^*(\mathbf{v}) - D_w^*(\mathbf{u}+\mathbf{v})D_{w^2}(\mathbf{v}) + \\ &2 \sum_{j=1}^{j=N} w_j^3. \end{aligned} \quad 3.6.16$$

In the case where no flat field correction is applied to the data, so that all photon events have unity weight, the unbiased bispectrum estimator simplifies to;

$$A I^{(3)}(\mathbf{u}, \mathbf{v}) = D^{(3)}(\mathbf{u}, \mathbf{v}) - D^{(2)}(\mathbf{u}) - D^{(2)}(\mathbf{v}) - D^{(2)}(\mathbf{u}+\mathbf{v}) + 2N \quad 3.6.17$$

It is important to realise that in this expression  $N$  is the number of photon-events in a the particular frame to be corrected, not the mean number of events per frame. In practice, because of efficiency considerations, the biased estimator and the correction terms are estimated over the whole data-set and subtracted at the end of the calculation.

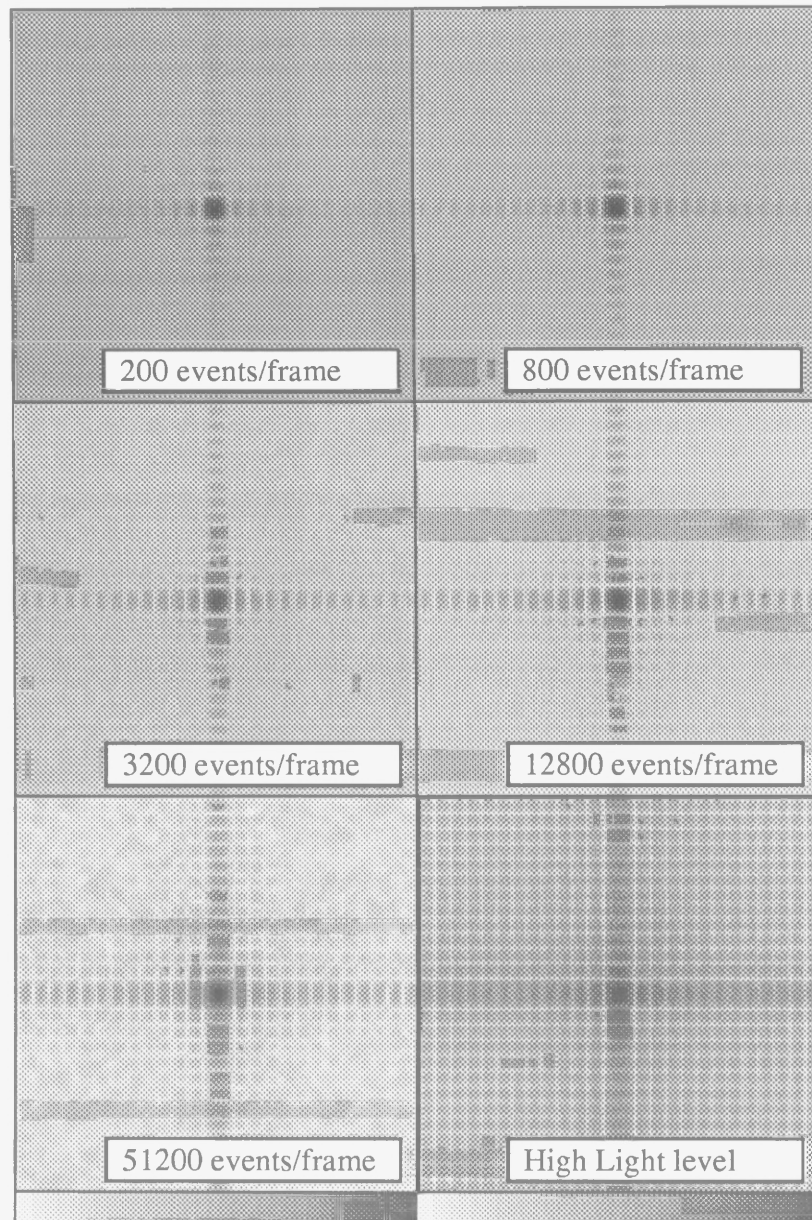
From the expression for the unbiased power spectrum estimator (Eqn. 3.6.5), we can see that the magnitude of the photon bias term in a normalised power-spectrum is of the order of  $\frac{1}{N}$  where  $N$  is the mean number of

photon events per frame. Away from the zero frequency peak, the values in a normalised power-spectrum are typically much smaller than unity. Thus the photon bias term at low light level, is in general much greater than the required power spectrum, and remains significant even at quite high photon event rates, i.e. rates of thousands of photons per frame, see Figure 3.6. We see then that it is extremely important to compensate for photon bias accurately, if meaningful results are to be obtained from correlation measurements. Obviously if the data we are analysing consists of weighted or unweighted photon event co-ordinates, compensating for the photon bias is a trivial problem, since the bias can be computed exactly.

If we try to use a simple intensified television system to record photon limited images, we will have considerable difficulty in correcting for photon bias when the data is processed. The problem is that photon events recorded with this type of camera have a range of intensities and sizes. At all but the lowest light levels, it is thus very difficult to identify and locate individual photon events. One must therefore compute the spectra of the images as if they were high light level. Unfortunately, without knowing the positions of each photon event it is impossible to compute accurately the weighted photon bias terms needed in Eqns. 3.6.3 and 3.6.16 without including unwanted cross terms due to evaluating the  $w_j^2$  terms.

### **3.7. Reconstruction of phase from the bispectrum.**

As indicated earlier it is possible to recover the Fourier phase of a signal from its bispectrum<sup>21,22</sup>. Equation 3.2.9, relating the Fourier phase of the signal to the phase of its bispectrum, is the basis for the following phase recovery technique. In order to simplify the notation we will



*These plots show the effect of ignoring the photon bias when computing a power spectrum (in this case a simple 2 dimensional sinc function). All intensities are plotted on a log scale in order to emphasise the lower values in the power spectrum. It is apparent that even with light levels as high as 50,000 events per frame, photon bias produces a significant uniform background.*

**Fig. 3.6**

consider the bispectrum of a one dimensional signal. It is straightforward to rewrite the following equations for a higher dimensional signal. Thus

in one dimension, for a **real** signal, equation 3.2.9 becomes;

$$\beta(u+v) = \beta(u) + \beta(v) - \beta^{(3)}(u,v). \quad 3.7.1$$

Where  $u$  and  $v$  are now scalar quantities, and  $\beta$  is taken to mean the phase of the corresponding spectrum, i.e.;

$\beta(u)$  is the phase of  $I(u)$ .

$\beta^{(3)}(u,v)$  is the phase of  $I^{(3)}(u,v)$ .

Using 3.7.1 we can write down the following set of relationships;

$$\beta(2) = \beta(1) + \beta(1) - \beta^{(3)}(1,1). \quad 3.7.2$$

$$\beta(3) = \beta(2) + \beta(1) - \beta^{(3)}(2,1). \quad 3.7.3$$

$$\beta(4) = \beta(3) + \beta(1) - \beta^{(3)}(3,1) \dots \beta(4) = \beta(2) + \beta(2) - \beta^{(3)}(2,2). \quad 3.7.4$$

.....

.....

$$\beta(u+1) = \beta(u) + \beta(1) - \beta^{(3)}(u,1) \dots \beta(u+1) = \beta(u-1) + \beta(2) - \beta^{(3)}(u-1,2). \quad 3.7.5$$

We know that for a real signal the phase at zero frequency in the Fourier domain,  $\beta(0)$  must be equal to zero. If we know the value of the signal phase at the first frequency point,  $\beta(1)$  we can use equation 3.7.2 to obtain the phase at the second frequency point  $\beta(2)$ . Using this value  $\beta(2)$  in equation 3.4.3 we can obtain the phase at the third frequency point  $\beta(3)$ . Similarly using the equations implied between equation 3.7.4 and 3.7.5 we can recursively reconstruct the rest of the Fourier phase values of the signal. In order to carry out this process we usually assume the phase at the first frequency point,  $\beta(1)$  is zero. In order to justify this assumption, let us look at the effect of this phase on the phase of the reconstructed signal. Rewriting equations 3.7.2 to 3.7.5 using the recursive

substitution just mentioned;

$$\beta(2) = 2 \beta(1) - \beta^{(3)}(1,1), \quad 3.7.6$$

$$\beta(3) = 3 \beta(1) - [\beta^{(3)}(1,1) + \beta^{(3)}(2,1)], \quad 3.7.7$$

$$\beta(4) = 4 \beta(1) - [\beta^{(3)}(1,1) + \beta^{(3)}(2,1) + \beta^{(3)}(3,1)], \quad 3.7.8$$

$$\beta(n) = n \beta(1) - \sum_{j=0}^{j=n} \beta^{(3)}(j,1). \quad 3.7.9$$

We can easily see now that the phase of the  $n^{\text{th}}$  Fourier frequency of the reconstruction contains a term,  $n \times \beta(1)$ , where  $\beta(1)$ , a constant, is simply the first Fourier phase of the reconstruction. This phase feature, known as a linear additive phase can easily be shown to be equivalent to a linear translation in the signal domain. If we translate a function  $i(x)$  by an amount  $x_j$  then its Fourier transform becomes;

$$\begin{aligned} \text{FT}(i(x+x_j)) &= \int_{-\infty}^{\infty} i(x+x_j) e^{-ixu} dx, \\ &= e^{ix_j u} \int_{-\infty}^{\infty} i(x+x_j) e^{-i(x+x_j)u} d(x+x_j), \\ &= e^{ix_j u} I(u), \end{aligned} \quad 3.7.10$$

which is simply the original Fourier transform, with the phase of each term increased by a quantity which is linearly dependent on the Fourier frequency.

More generally, equation 3.7.1 tells us that if we know  $\beta(u)$  and  $\beta(v)$  we can find  $\beta(u+v)$ . Thus each signal Fourier frequency can be found from many separate equations, allowing us the potential to improve

the SNR of our reconstruction by averaging. Fig 3.7 illustrates the complete usage of a non-redundant twelfth of the bispectrum in phase retrieval.

| Phase reconstruction from the bispectrum. |                                   |                                   |                                   |                                   |                                   |
|-------------------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|-----------------------------------|
|                                           |                                   | $\beta(6)$<br>$\beta(3)+\beta(3)$ |                                   |                                   |                                   |
|                                           |                                   | $\beta(4)$<br>$\beta(2)+\beta(2)$ | $\beta(5)$<br>$\beta(3)+\beta(2)$ | $\beta(6)$<br>$\beta(4)+\beta(2)$ |                                   |
|                                           | $\beta(2)$<br>$\beta(1)+\beta(1)$ | $\beta(3)$<br>$\beta(2)+\beta(1)$ | $\beta(4)$<br>$\beta(3)+\beta(1)$ | $\beta(5)$<br>$\beta(4)+\beta(1)$ | $\beta(6)$<br>$\beta(5)+\beta(1)$ |
| $\beta(0)$<br>$\beta(0)+\beta(0)$         | $\beta(1)+\beta(0)$               | $\beta(2)+\beta(0)$               | $\beta(3)+\beta(0)$               | $\beta(4)+\beta(0)$               | $\beta(5)+\beta(0)$               |

*In the above diagram each square represents a sampling point of the discrete bispectrum of a one dimensional signal. The large  $\beta$  value in each square represents the object Fourier phase that can be computed using equation 3.7.1, from the bispectrum value represented by that square and the Fourier phase estimates indicated by the two smaller  $\beta$  terms.*

Fig 3.7

As mentioned before, the generalisation of this reconstruction technique to bispectra of multi-dimensional signals is relatively straightforward but very tedious, so is not discussed here; it has however been implemented to reconstruct from data generated by the algorithms explained in Chapter 5.

It has been suggested that a least squares technique should be used to reconstruct optimally a function's phase from its bispectrum<sup>23</sup>. If we

define the principal phase to be in the range  $-\pi \rightarrow +\pi$  then in Eqn. 3.7.1  $\beta(u+n)$  lies in the range  $-3\pi \rightarrow +3\pi$ . We therefore see that estimates of  $\beta$  from separate equations, may differ by  $0, \pi, 2\pi$  or  $4\pi$ . Unfortunately the multi-valued nature of expressions represented by Eqn. 3.7.1 precludes the use of a straightforward least squares algorithm. Recent work has demonstrated that it is in general possible to phase unwrap the bispectrum whilst carrying out a least squares fit. It would appear that at least for one dimensional signals, this method overcomes the multivaluedness problem and allows good reconstructions to be generated with the least squares method. A further problem with using a least squares approach is that the solution it finds has the error spread uniformly across all the solutions variables. It is not apparent that this is the best way to partition the phase error in the reconstruction, if the ultimate aim is to reconstruct the signal rather than its Fourier transform. For instance one could argue that it was better to reconstruct low frequency phases with as high a SNR as possible, at the expense of SNR on the high frequency phases. This would enable us to eliminate much of the noise by low pass filtering and thus obtain a well-characterised signal reconstruction. By well-characterised we could say that the signal was reconstructed with a given SNR and a given frequency bandwidth.

Even when using the recursive reconstruction technique, the multivaluedness of the phase can be a problem. Obviously we need to reduce all estimates of a given phase to the principle range, before averaging, otherwise we will end up with phase estimates containing spurious factors of the form  $\frac{n\pi}{m}$ . The problem arises with noisy data, when for instance, a particular signal phase has a principle value near  $\pi$ . When estimating this phase from the bispectrum, the presence of noise would mean that the

principle value of the phase estimate could equally well be  $+\pi$  or  $-\pi$ . An average of many such estimates would clearly give us a net phase value near zero. In order to circumvent this problem, we average normalised or weighted complex amplitudes rather than the corresponding phase values. Since we are averaging complex amplitudes, phase manipulations are conveniently performed by multiplication and division of the corresponding normalised complex vectors. This is computationally more intensive, but it turns out that the time needed to reconstruct is negligible in comparison to the time needed to estimate the average bispectrum of many frames.

### 3.8. Photon noise.

At low photon levels the dominant contribution to power-spectrum or bispectrum noise is usually found to be photon, or shot-noise. We will start by considering the effect of photon noise on the power-spectrum, for which the calculation of signal-to-noise is well known<sup>20,24,25</sup>. As in the case of the photon differencing analysis, the photon noise analysis for the power-spectrum and bispectrum are closely related. In the following text we will use angle brackets,  $\langle \dots \rangle$  to indicate the operation of averaging over Poisson statistics. In Chapter six we will carry this analysis further when we consider imaging through turbulence; in this case the image is non deterministic necessitating a further averaging operation over the statistics of the image.

In order to calculate the SNR for the power-spectrum we must calculate its, variance, that is we must calculate the quantity;

$$\text{VAR} = \langle |I^{(2)}(\mathbf{u})|^2 \rangle - \langle I^{(2)}(\mathbf{u}) \rangle^2, \quad 3.8.1$$

Where  $I^{(2)}(\mathbf{u})$  denotes the power-spectrum of the signal  $I(\mathbf{u}) = I(\mathbf{u})I^*(\mathbf{u})$ .

Assume that the data we are using consists of a series of time frames of photon event co-ordinates, with a mean occupancy of  $\bar{N}$  photon events per frame. In order to simplify the calculation we will assume that we are not using a flat field correction, which implies that all photon events will have the same (unit) weight. The Fourier transform of a frame of  $N$  unweighted photons can be written as;

$$I(\mathbf{u}) = \sum_{k=1}^{k=N} e^{-i\mathbf{u}\mathbf{x}_k}. \quad 3.8.2$$

The only non trivial term which we need to compute is the first term in the power-spectrum variance, equation 3.8.1. Making use of equation 3.6.5, the expression for the unbiased power-spectrum estimator, we can rewrite this term as follows;

$$\begin{aligned} \langle |I^{(2)}(\mathbf{u})|^2 \rangle &= \langle (D^{(2)}(\mathbf{u}) - N)^2 \rangle, \\ &= \langle (D^{(2)}(\mathbf{u}))^2 - 2ND^{(2)}(\mathbf{u}) + N^2 \rangle, \end{aligned} \quad 3.8.3$$

where  $D^{(2)}(\mathbf{u}) = D_w^{(2)}(\mathbf{u}) = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} e^{-i\mathbf{u}\mathbf{x}_j} e^{i\mathbf{u}\mathbf{x}_k}$  is the biased power-spectrum estimator. We must now average these terms over the Poisson statistics of the photon events. In order to do this we must make use of the factorial moment theorem;

$$\left\langle \frac{N!}{(N-r)!} \right\rangle = \bar{N}^r, \quad 3.8.4$$

where  $\bar{N}$  is the mean rate of a Poisson process.

Using this result we can immediately write down that;

$$\langle N^2 \rangle = \bar{N}^2 + \bar{N}. \quad 3.8.5$$

Using equation 3.6.3, the expression for  $D^{(2)}(\mathbf{u})$  we may write;

$$2ND^{(2)}(\mathbf{u}) = N \left[ N(N-1)I^{(2)}(\mathbf{u}) + N \right]. \quad 3.8.6$$

thus;

$$\langle 2ND^{(2)}(\mathbf{u}) \rangle = 2N^2(\bar{N}+2) I^{(2)}(\mathbf{u}) + \bar{N}^2 + 2\bar{N} \quad 3.8.7$$

The quantity  $(D^{(2)}(\mathbf{u}))^2$  is slightly more complicated to deal with. In order to average over Poisson statistics we must first expand it in terms of the photo-events in an individual frame.

$$|I^{(2)}(\mathbf{u})|^2 = \sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \sum_{l=1}^{l=N} \sum_{m=1}^{m=N} e^{-i\mathbf{u}\cdot\mathbf{x}_j} e^{i\mathbf{u}\cdot\mathbf{x}_k} e^{-i\mathbf{u}\cdot\mathbf{x}_l} e^{i\mathbf{u}\cdot\mathbf{x}_m} \quad 3.8.8$$

Following a similar procedure to that adopted for the photon bias calculation, we can split this quadruple summation into four distinct classes of terms, they are classified according to the number of equal indices as follows;

|                                                                                                                                                                                                  |                                                         |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------|
| $j \neq k \neq l \neq m$                                                                                                                                                                         | $N(N-1)(N-2)(N-3)$ terms of type a).                    |
| $  \left. \begin{array}{l}  j = k \neq l \neq m \\  j = l \neq k \neq m \\  j = m \neq k \neq l \\  k = l \neq j \neq m \\  k = m \neq j \neq l \\  l = m \neq j \neq k  \end{array} \right\}  $ | $N(N-1)(N-2)$ terms of type b).                         |
| $  \left. \begin{array}{l}  j = k = l \neq m \\  j = k = m \neq l \\  j = l = m \neq k \\  k = l = m \neq j \\  j = l \neq m = k \\  j = k \neq m = l \\  j = k = l = m  \end{array} \right\}  $ | $N(N-1)$ terms of type c).<br><br>$N$ terms of type d). |

Each class of terms is dealt with separately.

Terms of type a) are represented by the following summation;

$$\sum_{j=1}^{j=N} \sum_{k=1}^{k=N} \sum_{l=1}^{l=N} \sum_{m=1}^{m=N} e^{-iux_j} e^{iux_k} e^{-iux_l} e^{iux_m} \times \quad 3.8.9$$

$$(1-\delta_{jk})(1-\delta_{jl})(1-\delta_{jm})(1-\delta_{kl})(1-\delta_{km})(1-\delta_{lm})$$

$$= N(N-1)(N-2)(N-3) \left[ I^{(2)}(\mathbf{u}) \right]^2.$$

We can write down immediately;

$$\langle N(N-1)(N-2)(N-3) \left[ I^{(2)}(\mathbf{u}) \right]^2 \rangle = \bar{N}^4 \left[ I^{(2)}(\mathbf{u}) \right]^2. \quad 3.8.10$$

The terms of type b) can be further split into two distinct sub-classes. The first sub-class of which the term with  $j = k \neq l \neq m$  is a member, is characterised by the two equal indices occurring in exponentials of opposite sign. There are four terms in this sub-class, which finally yield a term of the form;

$$4\bar{N}^3 I^{(2)}(\mathbf{u}).$$

The two remaining terms are characterised by having their equal indices occurring in exponentials of the same sign, they yield the following two terms;

$$\bar{N}^3 I(2\mathbf{u}) I^*(\mathbf{u})^2$$

and

$$\bar{N}^3 I^*(2\mathbf{u}) I(\mathbf{u})^2$$

These two terms are particularly interesting because they indicate that, perhaps rather surprisingly, the variance of the power-spectrum at one frequency has a dependence upon the value of the power-spectrum at the first harmonic of that frequency.

Terms of type c) can again be split into various sub-classes. There are four terms of the form  $j = k = l \neq m$ , where the equalities link two exponentials of one sign with a single exponential of opposite sign, these yield;

$$4\overline{N}^2 I^{(2)}(\mathbf{u})$$

We also find a single term where the equalities link exponentials of the same sign, this is the term  $j = l \neq m = k$  which yields;

$$\overline{N}^2 I^{(2)}(2\mathbf{u}) ,$$

Again this term shows the power-spectrum variance depends upon the power-spectrum at the first harmonic frequency.

Finally we have the terms  $j = k \neq m = l$  where all the exponential factors cancel yielding;

$$2\overline{N}^2 .$$

Terms of type d) yield;

$$\overline{N}$$

Collecting all of these terms together ;

$$\langle |I^{(2)}(\mathbf{u})|^2 \rangle = \tag{3.8.11}$$

$$\overline{N}^4 (I^{(2)}(\mathbf{u}))^2 + 2\overline{N}^3 I^{(2)}(\mathbf{u}) + \overline{N}^2 +$$

$$\overline{N}^2 I^{(2)}(2\mathbf{u}) + \overline{N}^3 I(2\mathbf{u}) I^*(\mathbf{u})^2 + \overline{N}^3 I^*(2\mathbf{u}) I(\mathbf{u})^2$$

The second term in the variance expression is trivially evaluated;

$$\begin{aligned} \langle I^{(2)}(\mathbf{u}) \rangle^2 &= \left[ \overline{N}^2 I^{(2)}(\mathbf{u}) \right]^2 \\ &= \overline{N}^4 (I^{(2)}(\mathbf{u}))^2 \end{aligned} \tag{3.8.12}$$

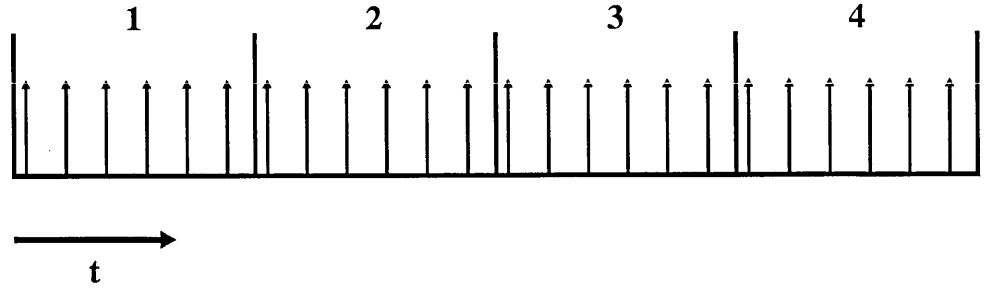
Thus the SNR of the photon differencing power-spectrum is;

$$\text{SNR} = \frac{\bar{N} I^{(2)}(\mathbf{u})}{\left[ (1 + \bar{N} I^{(2)}(\mathbf{u}))^2 + I^{(2)}(2\mathbf{u}) + \bar{N} (I(2\mathbf{u}) I^*(\mathbf{u})^2 + I^*(2\mathbf{u}) I(\mathbf{u})^2) \right]^{\frac{1}{2}}} \quad 3.8.13$$

This is the general expression for the signal-to-noise of a photon-limited autocorrelation, when the classical image intensity  $i(\mathbf{x})$  is a deterministic function.

### 3.9. Photon differencing with a continuous time window

In Chapter 2, one of the advantages cited for using a photon-counting and time-tagging detector was the possibility of using the temporal information provided by such a detector to obtain a higher SNR in signal reconstructions. We will use a rather simple argument to show that this is the case for all orders of correlation processing. As usual we will start by considering a second order correlation, namely by computing the SNR gain to be expected when evaluating the auto-correlation from time tagged data. It is apparent from equation 3.8.13 the expression for the SNR of the power-spectrum in the presence of photon noise, that for  $\bar{N} \gg 1$  the SNR scales roughly as  $\sqrt{\bar{N}}$ , where  $\bar{N}$  is the mean number of photon events per frame. When  $\bar{N} \leq 1$  we can assume in general that  $I^{(2)}(\mathbf{u}) \ll 1$  for frequencies other than those near zero, which implies that the SNR scales roughly as  $\bar{N}$ . We will call the regime where  $\bar{N} \leq 1$  the low light level regime, when  $\bar{N} \gg 1$  we will call it the medium light level regime. In speckle interferometry where  $\bar{N}$  is measured in photons per speckle, we will almost always be in the low light level regime. In most other applications we will usually be in the medium light level regime.



*A series of events occurring at uniformly spaced time intervals.*

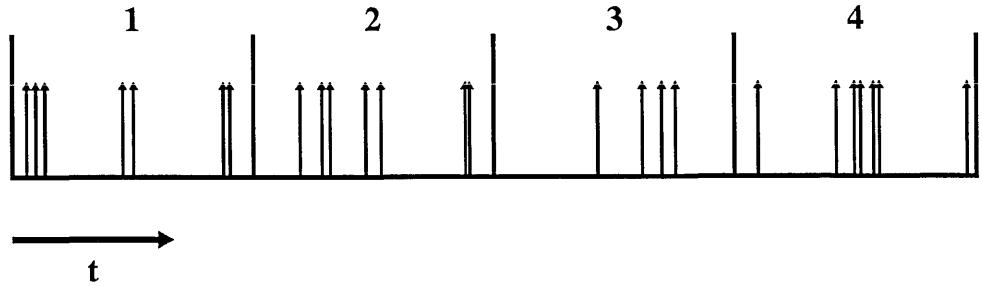
**Fig. 3.8**

---

When computing a photon-differencing auto-correlation, the number of non-redundant vectors computable from one frame of time length  $T$  containing  $N$  photon events is  $N(N-1)$ . Exactly half of these vectors are redundant, in the sense that they can be derived from the other half using the symmetry of the auto-correlation. Thus the number of non-redundant difference vectors computable from a frame of  $N$  photon events is  $\frac{1}{2}N(N-1)$ . Averaging this number over Poisson statistics we see that we expect to compute  $\frac{1}{2}\bar{N}^2$  difference vectors for each section of data of time length  $T$ . Thus in the low light level regime we assert that the SNR scales as  $\sqrt{\bar{N}_d}$  and in the medium light regime as  $\sqrt[4]{\bar{N}_d}$ , where  $\bar{N}_d$  is the mean number of difference vectors generated by a section of data of time length  $T$ .

Now let us consider a data stream consisting of events spaced regularly in time, see Figure 3.8, this stream is divided up into adjacent time frames of length  $T$  each containing  $N$  photons. Consider the second frame in this time series. If we were to calculate its contribution to the auto-correlation using a framed data set, it would produce  $\frac{1}{2}N(N-1)$  non redundant difference vectors. If instead of moving in one jump to the next data frame, we were to use a moving time window of length  $T$ , moving it along by one event between each calculation, until it was coincident with frame 2, we would obtain extra difference vectors due to differences between events in the latter part of frame one, and the first part of frame 2. Indeed, for each event in frame 2 we would generate  $N-1$  non redundant difference vectors by following this procedure. Thus using a continuous time window doubles the number of difference vectors calculable by photon-differencing. With this particular distribution of events we can only carry out photon differencing in the medium light regime, where the number of events per time length  $T$  is greater than 2. In this regime we will thus see a SNR improvement by a factor of about  $\sqrt[4]{2}$  or 19 percent.

Figure 3.9 illustrates a stream of photons obeying constant rate Poisson statistics. Referring back to Chapter 2, we see that one of the basic assumptions underlying the semi-classical theory of light detection, is that photon event times are statistically independent of when we start observing. Thus even if we observe time windows of length  $T$  which happen to partially overlap, we should retain Poisson statistics, provided a sufficiently long average is taken. As above, if we divide a photon event stream into adjacent frames of length  $T$ , we would be able to calculate  $N(N-1)$  vector differences from a frame of  $N$  photon events. Now imagine that instead of having non-overlapping windows we divide the data stream



*A time series of events with Poisson statistics.*

**Fig. 3.9**

---

into a large number of overlapping time windows whose start times are sufficiently close together as to ensure that the probability of a new frame having more than one new photon event over the previous frame is negligibly small. We will only obtain a new non-overlapping photon differences from such a window when a new photon event is included in it. When this happens we will on average expect  $\bar{N}$  photons to occur in the rest of the frame. That is the occupancy of the rest of the time window is completely unaffected by the fact that the time window happens to terminate on a photon event. We therefore expect on average that each new photon will contribute  $\bar{N}$  new photon difference vectors to the auto-correlation histogram. If we take a long time series of such frames we expect that on average a section of data of length  $T$  will contribute  $\bar{N}^2$  photon difference vectors. Therefore by comparison with the result for discrete frame processing we get the following SNR improvements: for the

low light regime, an improvement of  $\sqrt{2}$  or 41 percent and for the medium light regime an improvement of  $\sqrt[4]{2}$  or 19 percent.

This argument is easily extended to the calculation of higher order correlations by photon-differencing. In the case of the TC, for a discrete time window containing  $N$  photon events we expect to add  $\frac{1}{6}N(N-1)(N-2)$  unique bivectors to the TC histogram. On average we therefore expect each discrete time frame to add  $\frac{1}{6}\bar{N}^3$  bivectors to the TC histogram. In the continuous case we expect each new photon event should contribute  $\frac{1}{2}N(N-1)$  unique bivectors, or  $\bar{N}^2$  when averaged over Poisson statistics. Thus we would expect section of data of time length  $T$  to contribute  $\frac{1}{2}\bar{N}^3$  bivectors. The SNR expression for the TC is very much more complicated than that for the auto-correlation, but it still scales roughly as  $\bar{N}$  in the low light level regime and  $\sqrt{\bar{N}}$  in the medium light level regime. Thus in the low light level regime we expect to see an SNR improvement by a factor of about  $\sqrt[3]{3}$  or 44 percent and an improvement in the medium light regime of about  $\sqrt[6]{3}$  or 20 percent.

In general we can observe that for discrete time windows we contribute on average  $\frac{1}{k!}\bar{N}^k$   $k^{\text{th}}$  order difference vectors to the  $k^{\text{th}}$  order correlation for each time window of length  $T$ . In the case of a continuous time window, we will contribute  $\frac{1}{(k-1)!}\bar{N}^k$   $k^{\text{th}}$  order difference vectors for the same amount of data. This gives an over-all SNR gain of  $\sqrt[k]{k}$  in the low light level regime and  $2k\sqrt{k}$  in the medium light level regime.

One of the assumptions we have made in this calculation is that the temporal correlation of the signal we are measuring is box car shaped. Of

course in general this is not the case. It should therefore be possible to produce even greater gains in SNR by matching the response of our moving time window to the temporal correlation structure of the measured signal.

### 3.10. TC of an object on a uniform background

Uniform background illumination is a very common form of additive noise often seen in real data; its two main sources are intrinsic background illumination of the object through scatter etc, and possibly a dark current or leakage current from the detector itself.

Imagine we are imaging an object which consists of two distinct sub-objects undergoing relative motion <sup>26</sup>.

$$o(x) = a(x) + b(x+x_1) . \quad 3.10.1$$

The bispectrum of such an object will be;

$$O^{(3)}(u,v) = (A(u)+B(u)e^{iux_1})(A(v)+B(v)e^{ivx_1})(A^*(u+v)+B^*(u+v)e^{-i(v+u)x_1}), \quad 3.10.2$$

where the exponential term is the linear additive phase factor which results from the relative displacement of objects a and b. Thus;

$$O^{(3)}(u,v) = A^{(3)}(u,v) + B^{(3)}(u,v) + \text{cross terms}. \quad 3.10.3$$

In particular when b is a uniform distribution then both  $B(u)$  and  $B^{(3)}(u,v)$  will be weighted delta functions at the origin. Thus all of the cross terms in equation 3.10.3 will similarly consist of delta functions at the origin. In this case equation 3.10.3 becomes;

$$O^{(3)}(u,v) = A^{(3)}(u,v) + W\delta(0). \quad 3.10.3$$

This leaves the phase of  $A^{(3)}(u,v)$  unchanged and thus allows us to

reconstruct the phase of the object spectrum  $A(\mathbf{u})$  reconstructing  $a$  by simple background subtraction.

If object  $b$  is not a uniform background we will have to integrate the cross terms of equation 3.10.3 over the probability distribution of the displacement vector  $\mathbf{x}_j$ . This operation will make extraction of the phase of  $A^{(3)}(\mathbf{u}, \mathbf{v})$  or  $B^{(3)}(\mathbf{u}, \mathbf{v})$  difficult or impossible.

## **Imaging A Randomly Translating Object.**

### **4.1. Introduction**

One of the simplest problems in which one can imagine using the triple correlation (TC), is that of imaging a randomly translating (but otherwise undistorted) object at extremely low light level. It is clear that if it is not possible to reconstruct an object under these fairly undemanding circumstances, that imaging under more severe conditions (such as in the presence of turbulence) will fail. When doing this experiment, we have to consider the non ideal behaviour of the detector. In particular it is important with IPD that we minimise the effect of local dead time.

At high light levels if we wished to increase the SNR of an image of a translating object, we would probably do this by averaging a series of centroided images. That is, we would take a set of short exposure images of the object and add them together, after having first made the centres of gravity of each separate image coincident. In the following text we will refer to this procedure as the centroiding technique or shift and add algorithm. As a basis for comparison with the TC reconstruction technique, we have performed this centroiding algorithm at low light level, on the same data sets as used for the TC reconstructions.

### **4.2. The shift and add algorithm**

At high light levels the shift and add or centroiding technique works well, but at low light levels it becomes non-linear, giving a distorted

image. To see this let us consider the effect of centroiding a series of image frames with each image containing only  $N$  photon events per recorded frame. Suppose we have some function  $p(\mathbf{x})$ , which represents a normalised high light level version of our image; imaging  $p(\mathbf{x})$  with  $N$  photon events, a particular arrangement of photon events can be represented as follows;

$$\mathbf{i} = \sum_{k=1}^{k=N} \delta(\mathbf{x}-\mathbf{x}_k) . \quad 4.2.1$$

This photon event distribution will occur with the following probability;

$$q(\mathbf{i}) = \prod_{k=1}^{k=N} p(\mathbf{x}_k), \quad 4.2.2$$

where we make the assumption that the normalised high light level image  $p(\mathbf{x})$  is the probability density function for photon event detection and that all photon events are independent of each other. This is the first axiom which we assumed when deriving the semi-classical model in Chapter 2. The centre of gravity of such a frame  $\mathbf{c}$  is merely its first spatial moment, scaled by the inverse of the number of photon events in the frame;

$$\mathbf{c} = \sum_{k=1}^{k=N} \frac{\mathbf{x}_k}{N}. \quad 4.2.3$$

Thus after recentering using the above expression for the centroid, the co-ordinate of each photon event in the frame becomes;

$$\begin{aligned} \chi_j &= \mathbf{x}_j - \sum_{k=1}^{k=N} \frac{\mathbf{x}_k}{N}. \\ &= \frac{(N-1)}{N} \mathbf{x}_j - \sum_{\substack{k=1 \\ k \neq j}}^{k=N} \frac{\mathbf{x}_k}{N}. \end{aligned} \quad 4.2.4$$

The probability density for the position of a recentred photon event  $\chi_j$ , is

found by integrating the probability of obtaining  $\chi_j$  over all possible combinations of  $x_j$  and  $x_k$  that are consistent with  $\chi_j$ ;

$$p_c(\chi_j) = \int dx_1 \cdots \int dx_N \prod_{k=1}^{k=N} (p(x_k)) \delta(\chi_j - \frac{(N-1)}{N} x_j + \sum_{\substack{k=1 \\ k \neq j}}^{k=N} \frac{x_k}{N}), \quad 4.2.5$$

where the integrations run over the non zero support of the image.

In order to progress further with this calculation we move to the Fourier domain and write the characteristic function of  $\chi_j$  as;

$$P_c(\Omega) = \int dx_1 \cdots \int dx_N \prod_{k=1}^{k=N} (p(x_k)) \cdot \int d\chi_j \delta(\chi_j - \frac{(N-1)}{N} x_j + \sum_{\substack{k=1 \\ k \neq j}}^{k=N} \frac{x_k}{N}) e^{i\chi_j \Omega} d\chi_j. \quad 4.2.6$$

First we proceed with the integration over  $\chi_j$  ;

$$P_c(\Omega) = \int dx_1 \cdots \int dx_N \prod_{k=1}^{k=N} (p(x_k)) e^{i(\frac{(N-1)}{N} x_j + \sum_{\substack{k=1 \\ k \neq j}}^{k=N} x_k) \Omega}. \quad 4.2.7$$

Now we complete the remaining integrations;

$$P_c(\Omega) = P(\frac{N-1}{N} \Omega) \cdot (P^*(\frac{\Omega}{N}))^{N-1}. \quad 4.2.8$$

From this final equation we can see that for reasonably large values of  $N$ , photon limited centroiding returns an image which is a narrowed version of the correct image convolved with an object dependent point-spread function, which tends towards a delta-function as the number of observed photon events increases. Looking at the two terms of Eqn. 4.2.8 separately: The term  $P(\frac{N-1}{N} \Omega)$  approaches the image transform  $P_c(\Omega)$  as  $N$  becomes large and even for small  $N$  differs from the image transform only by virtue of a co-ordinate scaling factor  $\frac{N-1}{N}$ . In image space the

term  $(P^*(\frac{\Omega}{N}))^{N-1}$  represents a multiple convolution of the image on a co-ordinate scale reduced by a factor of  $\frac{1}{N}$ ; this term can never have a larger support than the original image, and for all finite support images will tend to a delta function as  $N$  becomes large.

Of special interest is the case of centroiding with only two photon events per frame, whence Eqn. 4.2.8 reduces to;

$$P_c(\Omega) = P(\Omega/2)P^*(\Omega/2). \quad 4.2.9$$

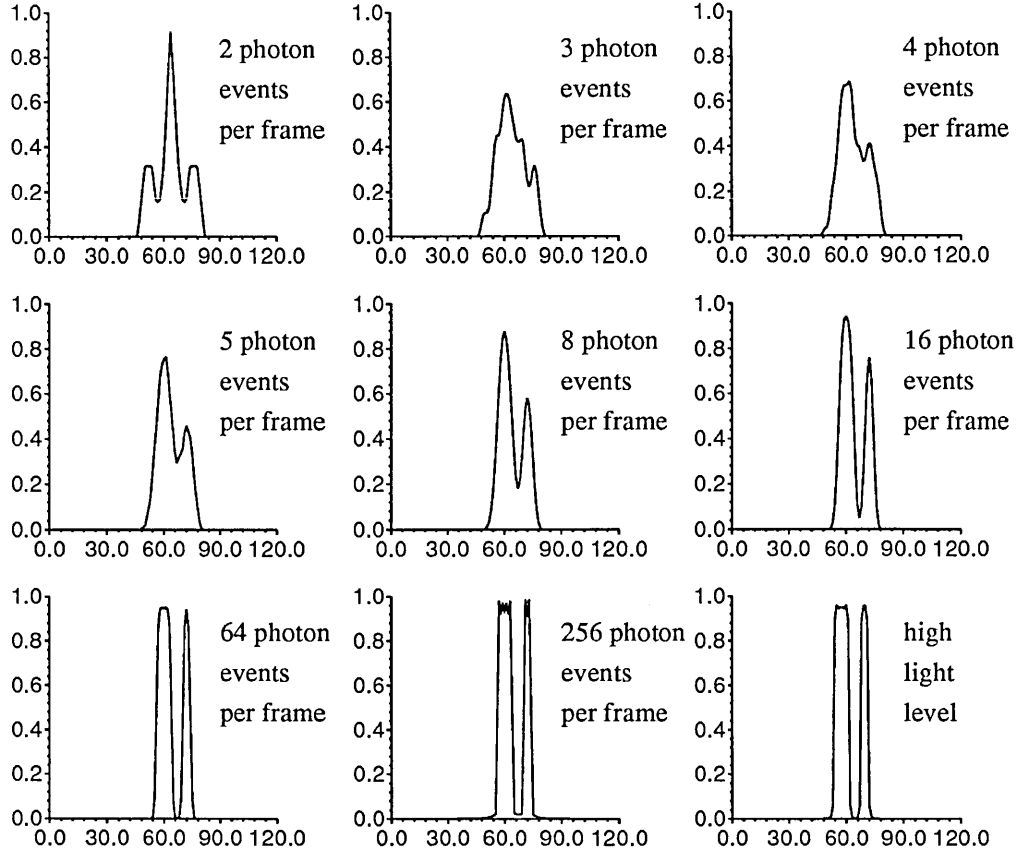
This is merely a scaled version of the power-spectrum of the correct image. In general if we have only two photons per frame, no method will be able to recover more information than that contained in the auto-correlation of the correct image.

With three photons per frame Eqn 4.2.8 reduces to;

$$P_c(\Omega) = P(2\Omega/3)P^*(\Omega/3)P^*(\Omega/3), \quad 4.2.10$$

which for a two dimensional image represents a 2 dimensional sub-plane of its four dimensional bispectrum. For larger  $N$  the result will be a 2 dimensional sub-plane of the corresponding  $2N$  dimensional  $N^{\text{th}}$  order spectrum. Figure 4.1 shows an example of what one would expect to see after imaging a simple object at low light using the centroiding technique for various numbers of photons per frame.

We could consider that the photon limited centroiding algorithm just described suffers from a photon bias, in a similar way to the photon bias which the auto-correlation and TC can contain. The rule for avoiding photon bias in a correlation, specifies that no photon should correlate with itself. Using this rule in the centroiding algorithm, we must specify that when we centroid a photon event, the centroid vector that we use should



*The above set of graphs illustrate the distortion inherent in using simple centroiding at low light level. Each graph shows what we would expect to recover by centroiding and averaging a set of frames containing a fixed number of photon events.*

**Fig 4.1**

be calculated without including the photon event that we are currently centroiding. Applying this rule to the expression for the centroided coordinates (equation 4.2.4), we get;

$$\chi_j = x_j - \sum_{\substack{k=1 \\ k \neq j}}^{k=N} \frac{x_k}{(N-1)}. \quad 4.2.9$$

It is easy to see that this alters expression for the characteristic function of the centroided image PDF to;

$$P_c(\Omega) = P(\Omega) \cdot (P^*(\frac{\Omega}{(N-1)}))^{N-1}. \quad 4.2.10$$

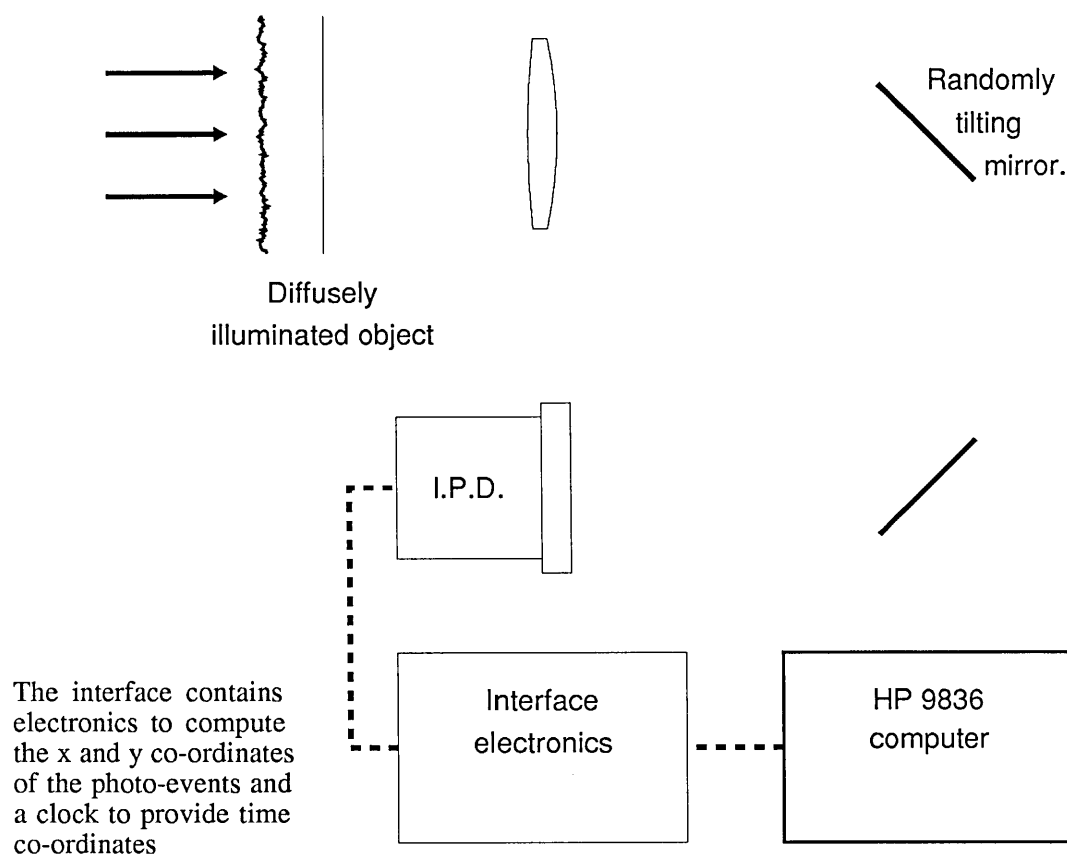
This expression gives a very similar result <sup>to</sup> that expressed by equation 4.2.8, the only difference being a different scale of the co-ordinate systems.

In the following measurements the centroid result consists of the summation of all centroided frames containing 3 or more photon events. To predict what we would expect to see under these circumstances it is necessary to average the resultant given by equations 4.2.8 or 4.2.10 over the Poisson statistics of N, the number of photons in each frame.

### 4.3. The experimental setup

Figure 4.2 shows the arrangement of the experimental apparatus, that was used to image a randomly translating object.

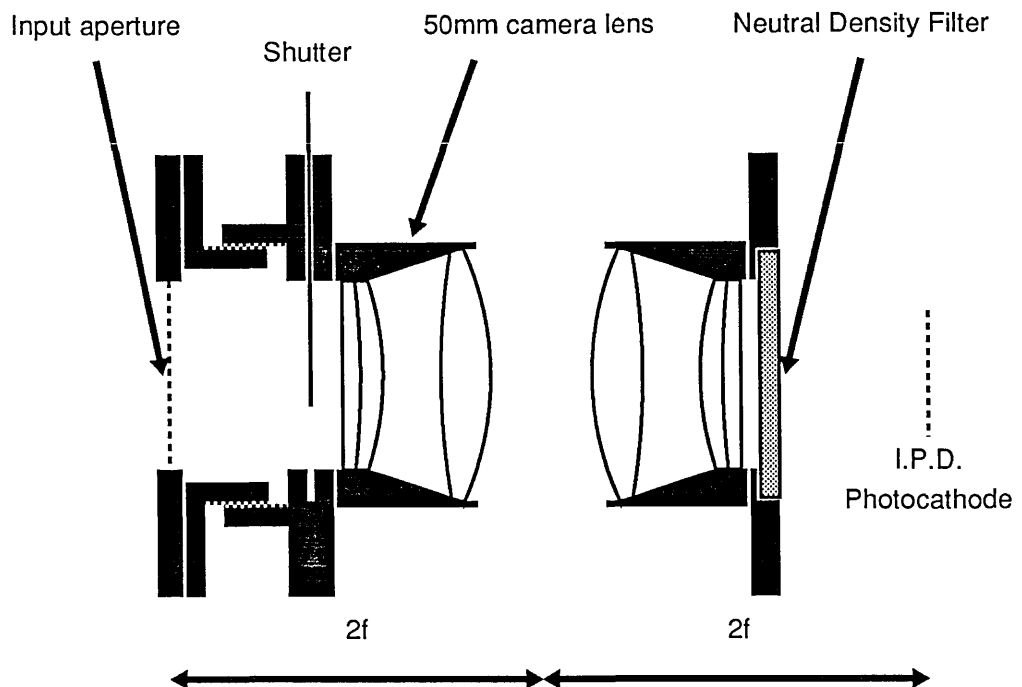
An object mask was back lit by a light box containing a tungsten lamp and a collimating lens. In order to improve the uniformity of the mask illumination two sheets of ground glass were placed between the light box and the mask. The mask was imaged onto the entrance pupil of an IPD. detector by a 600 mm achromatic doublet. The magnification of the image on the IPD photocathode was changed by moving this doublet. An iris diaphragm just in front of the lens cut its effective diameter to between 1 and 2 cm, giving the imaging system an f number of between 30 and 60 and thus a high depth of field. We expected the IPD to provide a resolution of at best 128 pixels across its 18mm photocathode (a resolution of 140 $\mu$ m) and thus an f/60 imaging system giving a point spread



*The arrangement of apparatus for acquiring photon event data for a translating object, using an IPD detector.*

**Fig 4.2**

function of approximately  $40\mu\text{m}$  width, should not seriously degrade the spatial resolution. In order to be able to place a variable entrance pupil in front of the IPD a 4 blade adjustable rectangular aperture was imaged onto the photocathode using two 50 mm camera lenses placed back to back. These lenses were arranged such that the nominal magnification of this system was unity (see Figure 4.3). A light tight shutter was placed just behind the IPD entrance pupil, allowing the IPD to remain powered



*The lens system used to re-image a variable sized entrance window on to the IPD photocathode*

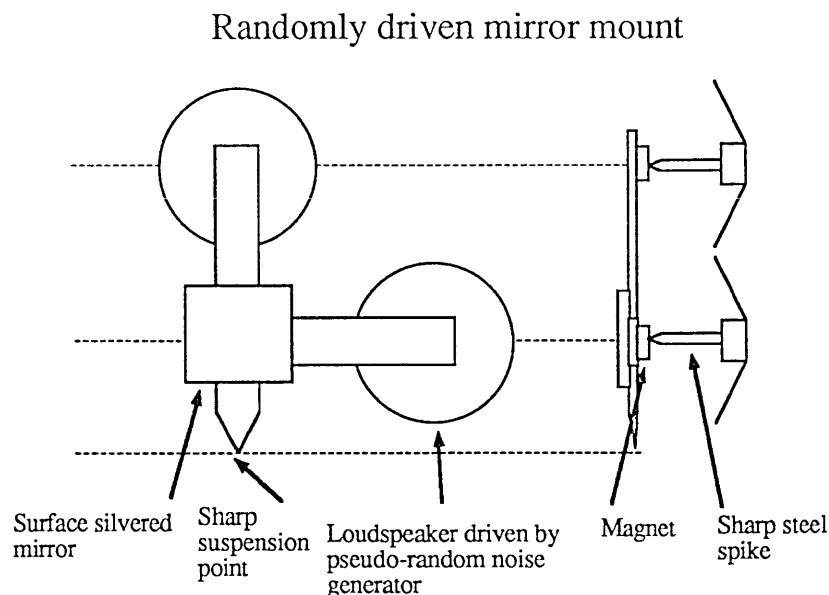
**Fig 4.3.**

up while the room lights were on. This was an important feature, since the dark count of the detector takes a long time to stabilise. With our IPD detector at room temperature, the dark count rate started at about 500-600 counts per second just after power up, reducing to approximately 200 counts per second after 5 or 6 hours of power-up and declining slowly thereafter. We placed a factor 4 ( $10^4$ ) neutral density filter between the pupil re-imaging lenses and the IPD entrance window and if necessary further filters could be inserted just in behind the light tight shutter. This neutral density allowed us to work with a reasonably bright object which

was convenient both from the point of view of aligning the system and for reducing the effect of background illumination. We also found it possible to run the detector with the shutter open with low level room illumination, which again eased the task of setting up the system. The light path from the object mask was folded using two mirrors, such that the entrance pupil of the IPD was facing away from the light source and computer equipment, thus reducing the amount of background illumination seen by the IPD.

#### 4.4. Image translation

One of the mirrors in the light path was constructed so as to allow it to be tilted on two axes. The mirror was mounted on two levers which were driven by loudspeaker coils, such that driving the loud speaker coils tilted the mirror. A digital noise source was used to drive the loud speakers so that a varying translation of known statistics could be produced<sup>27</sup>. The maximum data rate achievable was essentially governed by the local dead-time of our IPD detector. In order to reduce its effects to an acceptable level it was necessary to keep the local count rate of the detector to a low level. The frame rate of the detector had then to be correspondingly long in order to collect enough signal in each frames, this in turn required that the motion of the image across the detector had to be slow. In order to spread the light flux over as large an area as possible on the detector, the IPD was rotated until the y axes of the digitised coordinates coincided with the long axes of the object mask. The signal could then be integrated over a large range of y co-ordinates to give a one dimensional signal. A further advantage to this integration was that it reduced the problems due to the highly non-flat response of the IPD.



*The construction of the randomly tiltable mirror.*

**Fig 4.4.**

Typically a frame length of 6.5 ms was employed with the clock rate of the noise generator set to about 70 Hz. After digital filtering through a sync filter the resulting noise spectrum should have been flat, with a sharp cut-off at 4.4 Hz. A simple RC filter was used to suppress the fundamental band-pass aliased by the 70 Hz clock signal and its higher harmonics.

As pointed out earlier the local dead-time of the detector was potentially a serious problem when collecting the data, particularly because we were using such a compact object. According to the manufacturer, a circular region of the detector of radius about  $40\mu$  is capable of registering about 1 photo-event every 100ms. Since the local dead time is polarisable

it is important to keep the count rate at any point on the detector well below this rate, otherwise serious distortion of the the photo-event statistics will result. When stationary our object covered an area of about  $80 \times 12$  8 bit pixels, each of these pixels is equivalent to a square detector area of  $18/256$  millimeters on a side. Thus the absolute maximum count-rate for the whole object is of the order of 10000 events per second. In order to keep dead-time effects to less than 10 percent of the object intensity we must therefore observe at below 1000 events per second.

When the object was moving it effectively covered more of the detector surface than when stationary (the motion speed was significant when compared with the 100ms local dead-time), thus the maximum count rate could be slightly increased. The maximum photon rate was therefore kept well below 2000 events per second for all of the data acquisition runs. An interesting corollary to this is that at light levels of the order of 2000 events per second we might expect the TC image of the moving object to be sharper than a long exposure image of a stationary object taken at the same photon rate. This is because the detector is given time to recover from local charge depletion as the image moves over its surface.

#### 4.5. Data storage

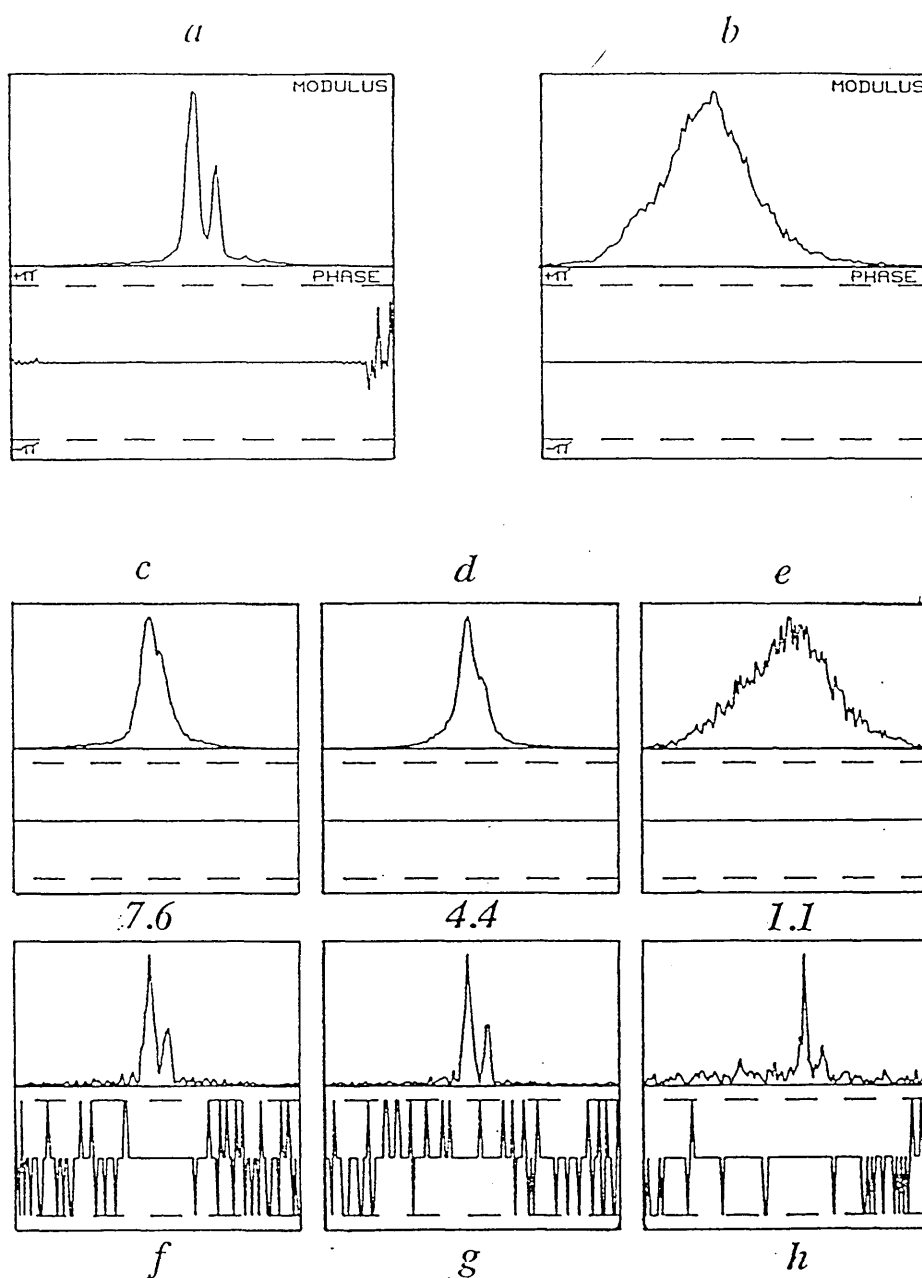
Data from the IPD photon position calculator was sent via a specially built time tagging interface to a HP9836 computer, using the HP GPIO interface. Each time-tagged photo-event consisted of a 32 bit long word comprising two 8 bit spatial co-ordinates and a 16 bit time co-ordinate. Software running on the HP9836 allowed one to integrate the incoming data into an image for setup purposes, or collect up to

approximately two hundred thousand time-tagged photo events for later analysis. The length of the time window into which the data was divided was chosen to be short enough as to effectively freeze the motion of the object. Since the resolution of the IPD is only of the order of 50 line pairs across the whole photocathode (5 line pairs/mm) the data processing algorithms used only the most significant seven bits of the 8 bit spatial co-ordinates. Processing was then carried out using at most 128 element arrays.

The photon events destined for later processing were read from the computer's input buffer and transferred to a Pascal record containing fields for the photon events  $x, y$  and  $t$  coordinates. The photon event co-ordinates were passed through the windowing and time framing filters described below before being written to disk.

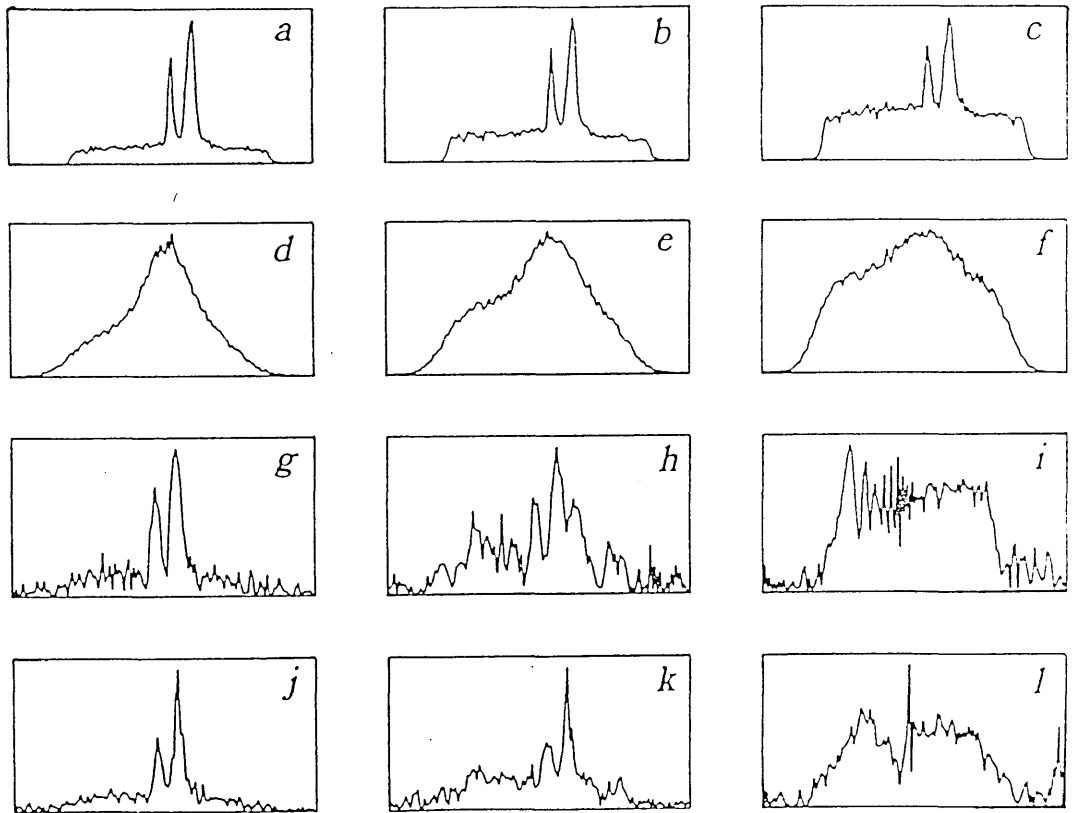
In the imaging mode of the data logging program it was possible to position a window over any rectangular sub-section of the image. Subsequently in the data collection mode of the program only photo-events occurring inside this window were stored on disk. After passing through the windowing filter the photo-events were collected into individual time frames. At this point a consistency check on the time co-ordinate of the data was performed, the purpose of this being to detect bad data which occasionally occurred due to a problem with the timing interface. This problem was eventually traced to an incorrectly set timing capacitor.

Triple correlation of the data was carried out using a machine-coded photon differencing algorithm which was capable of processing about one hundred 10 photon frames per second (see Appendix C). The theory behind triple correlation imaging is given in Chapter 3.



Results of applying the TC and centroiding techniques to photon limited data from an randomly translating object imaged by an IPD. Figure a is a long exposure image of the object when static; figure b is the long exposure image of the same object when randomly translating. Figures f, g, h and c, d, e are the reconstructions of this translating object using, respectively, the TC and centroiding techniques. The mean number of photon events per frame of data is 7.6 for c and f, 4.4 for d and g, and 1.1 for e and h.

Fig 4.5



Results obtained when TC imaging a moving object in the presence of a noisy uniform background. Starting at the top each of the four horizontal rows represent, a,b,c the long exposure image<sup>of the static object</sup>, d,e,f the image by centroiding, g,h,i the image by TC using 1 row of the bispectrum, j,k,l the image by TC using the whole of the bispectrum. In every case the object signal was on average 3.1 photon events per frame. For the vertical column a,d,g,j the background rate averaged 4.3 photon events per frame, for column b,e,h,k 8.7 events per frame, and for column c,f,i,l 16.7 events per frame.

Fig 4.6

#### 4.6. Experimental results

Figure 4.5 shows some results obtained by centroiding and by triple correlation using frames with different mean photon event numbers. The

detector was run at room temperature and thus the dark current was quite high. For these measurements it was of the order of 0.6 events per frame. At the very lowest rate of 1.1 photon events per frame there were thus only on average 0.5 signal photon events per frame and this means that out of the 32767 frames collected only about 470 contained 3 or more object photon events, giving a total of about 1500 useful object photon events in the reconstruction.

For a one dimensional signal it is possible to reconstruct the Fourier phase of the signal from a single line of the two dimensional bispectrum. However when the bispectrum is noisy, it would seem to be advantageous to use the whole non-redundant twelfth of the bispectrum in order to bolster the SNR of the reconstruction. To show the effect of this increase in SNR the input data was made deliberately noisy by applying a significant uniform background illumination. The results obtained when reconstructing this object from the noisy data for varying amounts of background noise are shown in Figure 4.6. It is evident that both forms of the triple correlation perform much better than the centroiding technique in these circumstances. When reconstructing from the complete bispectrum a weighted phase average was used, each reconstructed phase being weighted in proportion to the magnitude of the bispectrum element from which it was derived. This approach is much simpler than using the correct expression for the bispectrum noise, but is accurate enough to produce a significant improvement in the quality of the reconstruction.

In this experiment, due to the maximum size of the entrance slit, the uniform background did not extend over the whole of the array. This will have given rise to distortion by corrupting the cross-terms between the object and background (see Chap. 3 ). In this particular case the problem

could have been ameliorated to some extent by only considering the central 64 elements of the data array and processing a 64 square triple correlation.

## 5

### Calculation of the TC or bispectrum for large arrays.

#### 5.1. Introduction

In Chapter six we will look at the application of the TC/bispectrum to the problem of speckle interferometry, or high resolution astronomical imaging. The application of the TC to this problem requires us to be able to manipulate the bispectrum of large two dimensional arrays. Since we need to evaluate the bispectrum in order to use the reconstruction algorithm described in Chapter 3, if we choose to calculate the TC first it is necessary that we calculate the whole of the non-redundant TC in order to enable us to compute its Fourier transform. As the TC of a one dimensional signal is two dimensional, and the TC of a two dimensional signal is four dimensional it is clear that calculation of the TC becomes inappropriate due to its impractically large size, when we are dealing with large two dimensional arrays. A further point to bear in mind is that as mentioned in Chapter 3, calculating the TC by the photon differencing technique rapidly becomes inefficient as the number of photon events in each data frame becomes large. Thus when we are dealing with relatively high photon event rates, the calculation of the TC again becomes inappropriate. For these two reasons, when we wish to process large arrays of data we will in general need to calculate bispectrum elements directly, rather than compute the TC first. In this chapter we will describe three different techniques for making use of the TC/bispectrum technique for large arrays; these are, the radon method, the direct method, and a modified photon differencing scheme which we call the sub-plane method.

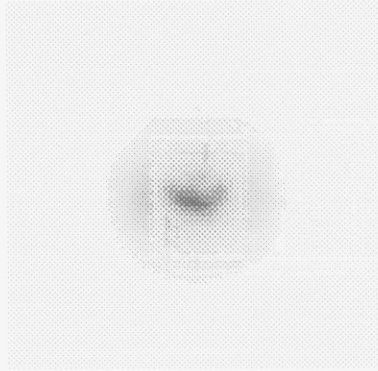
Two of these methods, namely the direct method and the radon method, can also be useful when the data photon event rate becomes too high for photon differencing to be efficient. For each of the three techniques described, two reconstructions of an object from two different simulated photon event data sets are presented. The first data set is a simple-minded simulation of imaging a randomly translating object at extremely low photon event rates. Each reconstruction from this data-set is derived by processing 2000 frames of data with an average photon event rate of 10 events per frame. The second data set consists of simulated speckle interferometry data. The speckle data-set is intrinsically much more noisy than the first data set, so even with larger numbers of frames and with more photon events per frame, the reconstructions tend to be more noisy. The parameters used in this particular simulation are;

2 metre telescope.

20 centimetre seeing (  $r_0 = 20\text{cm}$ , where  $r_0$  is the Fried parameter<sup>28,29</sup>).

4 centimetre sampling when projected to the aperture plane of the telescope.

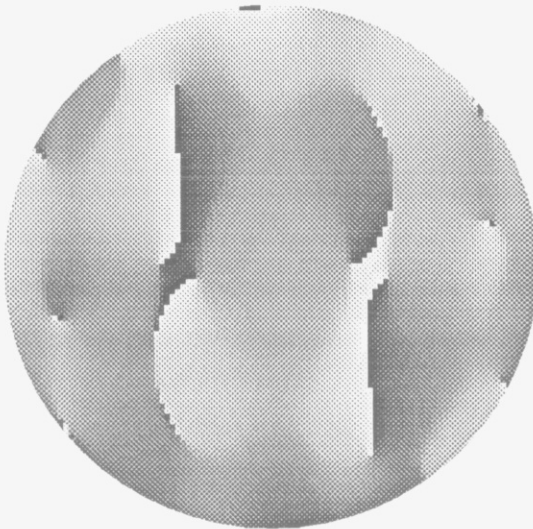
Since we are primarily concerned with the quality of the reconstructed phase, each reconstructed object is generated by combining the reconstructed phase with the true object Fourier modulus. Thus these reconstructions are better than could be achieved in practice, where the Fourier modulus of the object would have to be determined as well its phase. The test object that we use, and its corresponding phase is shown in Figure 5.1. In these simulations we have not imposed an optical transfer function upon the the Fourier modulus. At sufficiently high light level to give good SNR this would be the situation in speckle



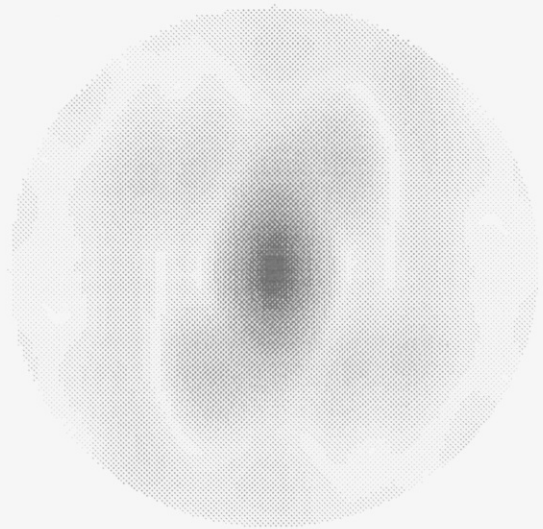
Diffraction Limited Image  
Uncompensated For Telescope OTF



Diffraction Limited Image  
Compensated For Telescope OTF



Fourier Phase



Fourier Modulus

*The test object used to exercise the following triple correlation reconstruction methods. The object has been low passed filtered to simulate being imaged by a diffraction limited imaging system. It is shown as it would be seen through an un-aberrated telescope, and as it would look if we compensated the image for the optical OTF out to the diffraction limit. Also shown is the objects Fourier phase and modulus; both show a sharp circular cut off at the diffraction limit of the telescope.*

**Fig 5.1**

---

interferometry, where the transfer function is calibrated by observing an unresolvable source. Also shown in Figure 5.1 is a simulation of what the object would look like if it were imaged through an unaberrated 2 metre telescope, without the OTF being compensated.

## 5.2. The Direct method

This method is the most obvious way to tackle the problem of computing the bispectrum for a large array. Instead of attempting to calculate the TC and then Fourier transforming to get the bispectrum, we compute a small segment <sup>of</sup> the bispectrum directly for each individual frame of data. These bispectrum estimates are then averaged over the whole data set.

We use the following processing steps.

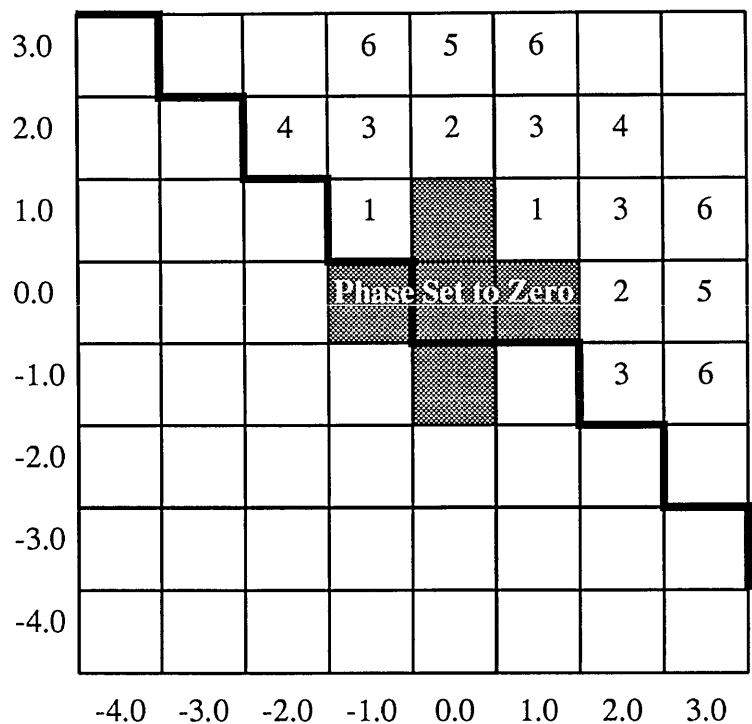
- 1): Decide which bispectrum values we are going to compute, and construct a lookup table mapping each bispectrum value into three Fourier phase values (see Equation 3.2.4 of Chapter 3).
- 2): Calculate the Fourier transform of each frame of data. Using bispectrum lookup table, compute the average bispectrum over all of the data frames.
- 3): Correct the bispectrum for the effects of photon bias.
- 4): Reconstruct Fourier phase values from the computed bispectrum, using a variant of the recursive reconstruction algorithm described in Chapter 3.

Due to insufficient SNR in the estimated bispectrum, it is often not possible to reconstruct the phase of an object to the Nyquist frequency of the sampling array. Indeed, for speckle data, it will often be the case that the Nyquist frequency of the recorded speckle pattern will be considerably

greater than the telescope cut-off frequency. In this algorithm we make use of this fact to reduce the computational load of object reconstruction. We do this by estimating the frequency to which we will be able to reconstruct, and then, only calculating sufficient bispectrum elements to reconstruct to this frequency. The number of bispectrum values to be computed is controlled by two parameters; the maximum reconstruction frequency, and the maximum number of bispectrum estimates which will be used to reconstruct each Fourier phase.

In order to speed up the calculation of the bispectrum the Fourier transform to bispectrum mapping is placed in a lookup table. This consists of a one dimensional array containing a list of four address offsets. The first offset is that of the relevant bispectrum element in the list of bispectrum values being computed. The remaining three offsets are the offsets in the Fourier transform array of three values which must be multiplied together in order to calculate the associated bispectrum value. In this way the task of the bispectrum averaging program is reduced to that of calculating the Fourier transform of each frame, progressing down the bispectrum lookup table performing the complex multiplications specified therein and then incrementing the bispectrum vector. One of the strengths of this method is that simply by changing the subroutine which calculates the bispectrum lookup table it is possible to sample different parts of the bispectrum.

The way in which the bispectrum values were chosen in the program used to generate Figures 5.3 and 5.4, is illustrated in Figure 5.2. This array represents the Fourier spectrum that we are trying to reconstruct. The order in which we reconstruct various spectrum components is indicated by the numbers in each array element. Due to hermitian symmetry it is



*The above grid represents the Fourier Phase of an object we wish to reconstruct. The Numbers in the grid cells indicate the order in which these phase values would be reconstructed, when using the direct method to calculate the bispectrum. The heavy line divides the grid according to hermitian symmetry, the dark shaded cells have their phase set to zero, before the recursive reconstruction is started.*

**Fig 5.2**

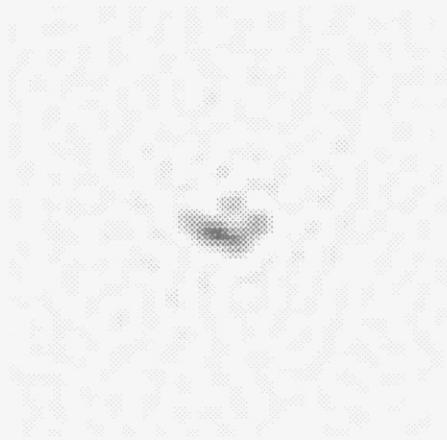
obviously only necessary to compute one half of the required spectrum. The idea behind this ordering is to compute spectrum values with small offsets from the origin first. When choosing which bispectrum values we should use in order to derive the  $n^{\text{th}}$  phase value in the reconstruction sequence, we must obviously obey the condition that we can only use spectrum values which we have already computed. We must also make sure that we do not choose two apparently different bispectrum values which are in fact the same due to symmetry. When we are processing

speckle interferometry data, we have two further constraints on the choice of bispectrum values.

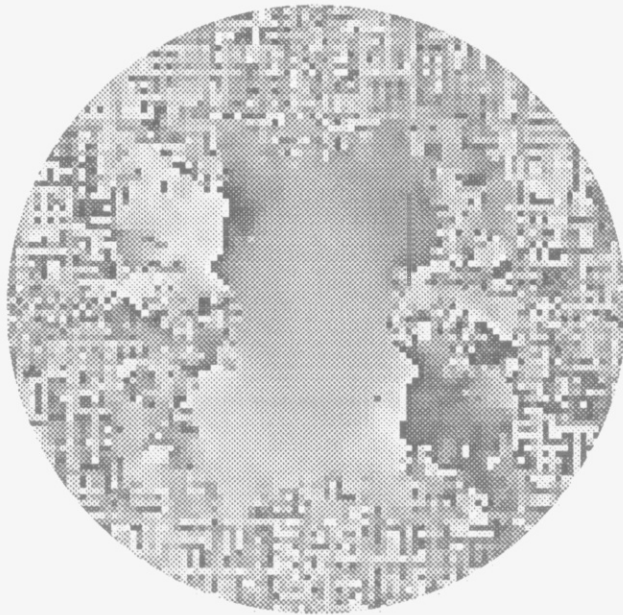
- 1): It has been shown that bispectrum values which lie close to the axis have significantly higher SNR than values which lie far from the axes. This implies that at least one of the spectrum values used to form each bispectrum value should be for as small a frequency as possible.
- 2): The noise on adjacent bispectrum values tends to be correlated, which means that if possible we should avoid using adjacent bispectrum values.

Obviously these two conditions tend to act against each other, making the choice of an optimal set of bispectrum values difficult.

A potential problem with the above scheme is the build up of rounding errors and loss of dynamic range as the calculation proceeds. For example if we were to accumulate the bispectrum as single precision complex variable which has about 6.5 significant figures or 22 bits of dynamic range then after accumulating only 1000 frames of data we would only be considering the most significant 3.5 figures (8 bits) of any further data frames. The rounding noise by this point would have grown to  $\sqrt{1000}$  times the mean rounding noise on a single addition, or approximately 1.5 decimal places, rendering any further processing of little value. In this case the simplest remedy is to use double precision variables. A slightly more complicated, but more efficient solution is to use two single precision arrays, and accumulate for example 1000 frames worth data in the first. This accumulated data would then be added to the second array, then a further 1000 frames of data accumulated in the first array. Using this procedure considerably extends the number of data frames that we can



Reconstructed  
Image



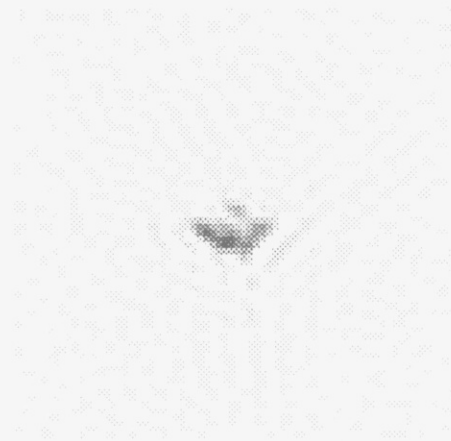
Reconstructed  
Fourier Phase

*Reconstruction by direct method using 20000 frames of simulated speckle data with a mean photon event count of 200 per frame.*

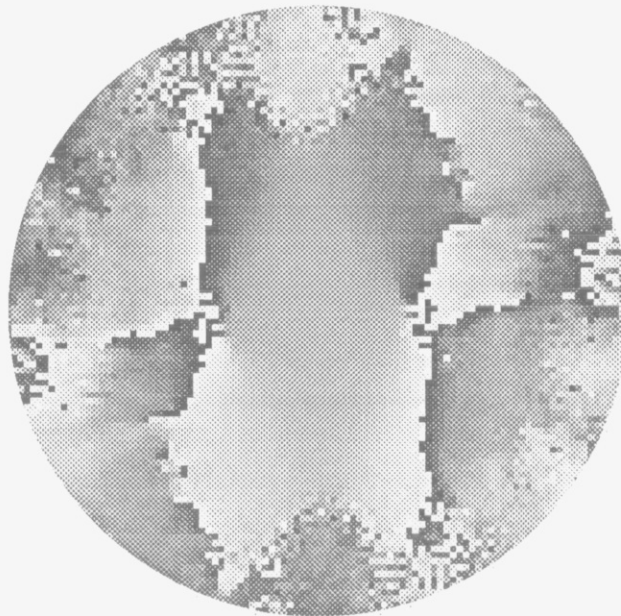
**Fig 5.3**

---

process before lack of dynamic range and rounding error become a significant problem. This solution is obviously only more efficient provided single precision arithmetic is faster than double precision arithmetic.



Reconstructed  
Image



Reconstructed  
Fourier Phase

*Reconstruction by direct method from simulated data for a translating object. The data contains 2000 frames, of on average 10 photon events per frame.*

**Fig 5.4**

---

When dealing with photon event data there are two ways in which the Fourier transform of a data frame may be computed. The most obvious way is to simply place the photon events in an array and then use a fast Fourier transform (FFT) algorithm to compute the Fourier transform.

A second way is to realise that the Fourier transform of a single (delta function) photon event is a simple exponential term. Since the Fourier transform is a linear operator, the Fourier transform of a number of photon events, is simply the sum of these exponential factors.

$$I(\mathbf{u}) = \sum_{k=1}^{k=N} e^{-i\mathbf{u}\mathbf{x}_k} . \quad 5.2.1$$

Since we do not in general attempt to reconstruct the whole of an objects' Fourier phase, it is not necessary to compute the whole Fourier transform of each frame, in order to compute the required bispectrum values.

For example let us make a comparison of the two methods assuming we are processing a data stream representing an image which has been sampled by a 128 square array of sampling points, assuming that we hope to be able to reconstruct the central 64 by 64 square of Fourier phase. For both algorithms we will also assume that the relevant sine and cosine factors needed for the calculation, can be found in a lookup table. Using a conventional real FFT we must carry out approximately  $\frac{n}{4}\log(n)$  butterfly operations each comprising 10 floating point calculations; about  $6 \times 10^5$  floating point operations in all. Using the exponential sum transform, including hermitian symmetry, each photon event in a frame requires  $32 \times 64 \times 2 = 4096$  floating point operations to add its exponential phase factor to the Fourier transform. From these figures we see that under these circumstances, with less than 140 photon events per frame it is more efficient to compute the Fourier transform of a frame using direct sine and cosine addition, rather than an FFT.

### 5.3. The radon method

This method relies upon reducing a two dimensional (or possibly more than two dimensional) reconstruction problem, into a number of more manageable one dimensional problems. In order to reduce the problem we make use of the projection-slice theorem<sup>30</sup>. The central-slice or projection-slice theorem states that if we form a one dimensional projection of a two dimensional function at an angle  $\theta$ , the Fourier transform of this projection will be the same as the Fourier transform of the original two dimensional function on a line through its origin at an angle  $90^\circ + \theta$ .

The Fourier transform of a two dimensional function  $i(x)$  is written as;

$$I(u,v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i(x,y) e^{-iux} e^{-ivy} dx dy. \quad 5.3.1$$

Let us consider the special case where we project the function  $i(x,y)$  onto the  $x$  axis of the coordinate system, an operation carried out by the following integration;

$$i_p(x) = \int_{-\infty}^{\infty} i(x,y) dy. \quad 5.3.2$$

Where we have used the subscript  $p$  to denote that the function is a one dimensional projection of a higher dimensional function. We see that the Fourier transform of this projected function is simply;

$$I_p(u) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i(x,y) dy e^{-iux} dx. \quad 5.3.3$$

But this is simply equation 5.3.1 with the Fourier coordinate  $v$  set to zero.

That is, we can write;

$$I_p(x) = I(x,0). \quad 5.3.4$$

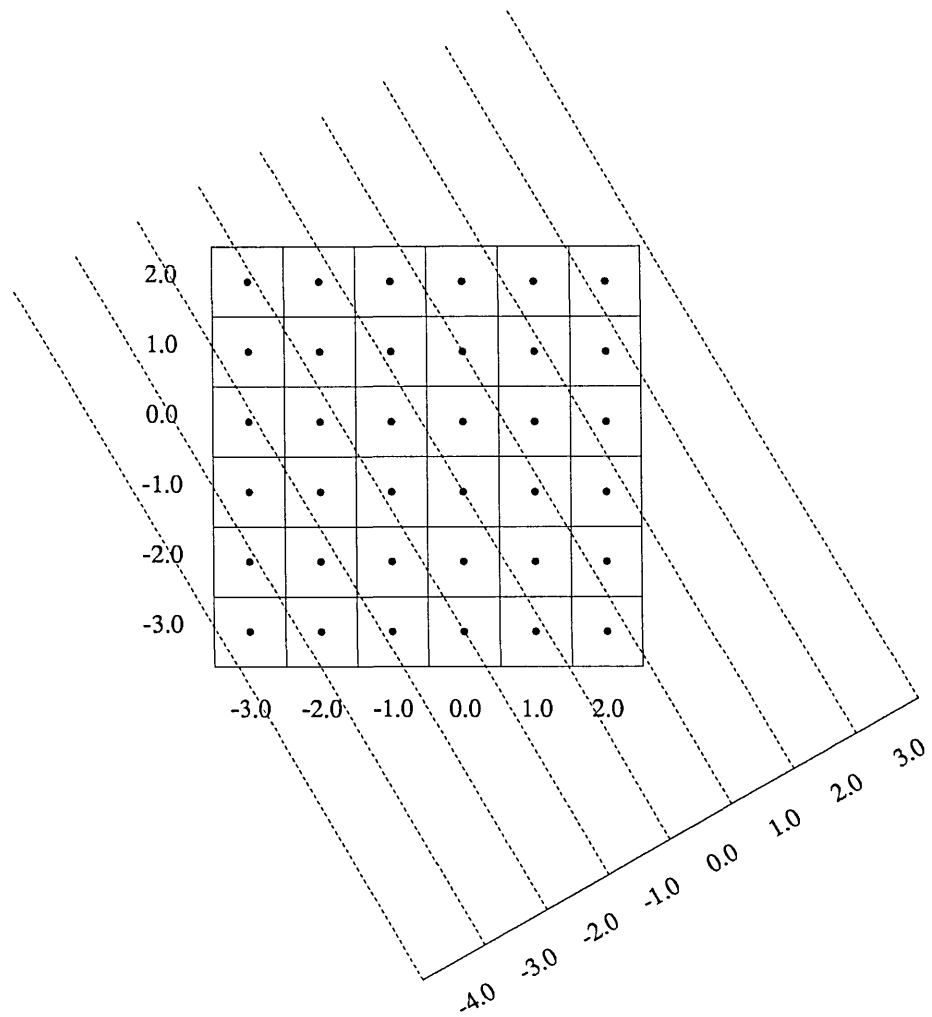
In other words Equation 5.3.3 represents the Fourier transform of the original transform, taken along the line  $v = 0$ , which proves the central slice theorem for the particular case of projection along the coordinate axes. For projection at an arbitrary angle  $\theta$  we must first perform a rotation of the coordinate system such that the new  $x$  axis lies at an angle  $\theta$  to the original  $x$  axis. This is accomplished by applying the two dimensional rotation matrix;

$$R = \begin{bmatrix} \cos(\theta) & + \sin(\theta) \\ - \sin(\theta) & \cos(\theta) \end{bmatrix} \quad 5.3.5$$

This theorem is readily generalised to  $n$  dimensions, which would in theory allow us to use the radon method for treating signals of any dimensionality. In two dimensions this technique is closely related to the computer tomography techniques which are extensively used in medical imaging. We find that with the relatively simple objects which we have studied, a fairly small number of projection angles (e.g. 36 or every 5 degrees) is sufficient for a reasonable reconstruction.

The steps in processing are:-

- 1): Project the two dimensional coordinate data onto a number of one dimensional lines at various angles. This is done by rotating the coordinate system by an angle  $\theta$  using the rotation matrix <sup>and</sup> ignoring the  $y$  coordinate effects the projection.
- 2): For each angle of projection, compute the average triple correlation or bispectrum over the data set. At this step we may use a photon differencing algorithm to compute the triple correlation, or compute



*This figure shows the correct way to project a sampled two dimensional array onto a sampled one dimensional array. The pixel boundaries of the one dimensional array have been extended to cover the two dimensional array. To effect the projection we must integrate to find the total energy in each of the extended one dimensional pixels.*

**Fig 5.5**

the bispectrum directly.

- 3): From the mean bispectrum at each angle, reconstruct the corresponding central slice of the object Fourier transform.

- 4): Combine the various central slices into an estimate of the object Fourier transform, thus arrive at an estimate for the object. If possible object domain constraints such as positivity, should be included in this step.

In practice the object modulus is computed separately using the standard speckle interferometry techniques. The various projection triple correlations can be computed numerically, possibly dividing the calculations among several machines. Alternatively since we are working with one dimensional functions, a hybrid opto-electronic processor<sup>31</sup> could be used to compute the correlations.

The problem of projecting a multi-dimensioned sampled array onto a one dimensional sampled vector is more problematical than the basic mathematics for continuous functions would suggest<sup>32</sup>. For discrete arrays of data, the problem is that if we just project the the coordinates of the centre of each array element onto a one dimensional vector, the projected coordinates do not lie uniformly along the projected vector. As a consequence of this, merely projecting an array by rotating its element coordinates and rounding them into the projection line coordinate will give a systematic bias to the projected function. The correct way to project such an array is to project the area of each pixel on to the projected vector, as shown in Figure 5.4.

When using photon event coordinate data it may not be appropriate to project the coordinates using an area projection, since this would result in the "intensity" due to a single photon event being divided into 2 or 3 pixels in the projected vector. Even if we do not wish to perform a photon differencing calculation with the data, such a projection would make correction of the photon bias terms difficult. If the photon event

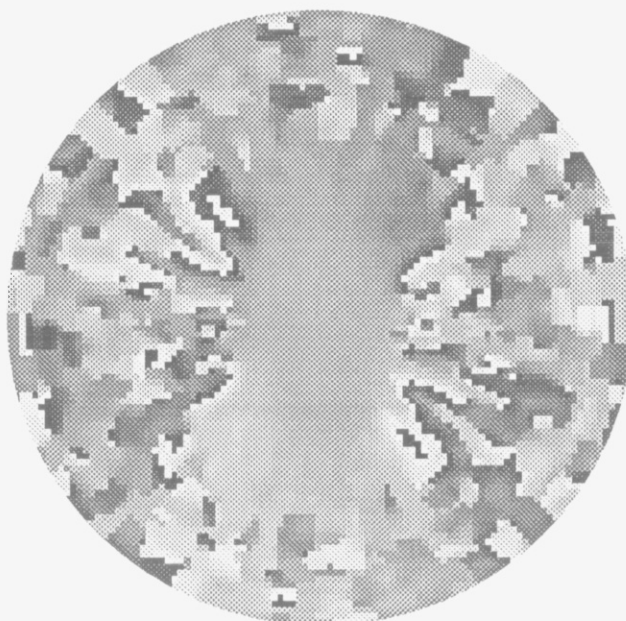
coordinates had very high spatial accuracy there would be no resampling problem since they would effectively define a continuous function on the scale of the projected array. The approach we have used is to arbitrarily extend the accuracy of the photon event coordinates by assigning each event a random position within its pixel. In this way we can project the coordinates at any angle without introducing any bias. The problem with this approach is that it introduces extra noise into the coordinates. Effectively what we have done is to extend the ‘accuracy’ of the coordinates beyond the sampling arrays Nyquist frequency. However this extra ‘accuracy’ contains no information, only noise which will appear below the Nyquist frequency due to aliasing.

It should be noted that when reconstructing from the bispectrum one is free to choose the values of the transform phase at the zeroth and first frequencies. To first order, provided the object is small, this is equivalent to being free to specify the phase gradient at the origin. Since the Fourier transform of the object intensity distribution is analytic, it follows that the phase gradient at any point must be independent of direction. Thus provided we make appropriate choices for the first two frequencies of each reconstruction (eg both zero), we can guarantee there will be no phase errors between the slices. In the image domain this is equivalent to making the centres of gravity of each projection coincident, a condition which must obviously be satisfied.

Although it would be possible to implement the radon transform algorithm using photon differencing, this would in general be inefficient. The number of operations required for a photon-differencing TC is proportional to  $\bar{N}^3$  where  $\bar{N}$  is the mean number of photon events per frame and this is independent of the size of the correlation array. If we calculate the



Reconstructed  
Image



Reconstructed  
Fourier Phase

*Reconstruction by radon method using 25000 frames of simulated speckle data containing on average 200 photon events per frame.*

**Fig 5.6**

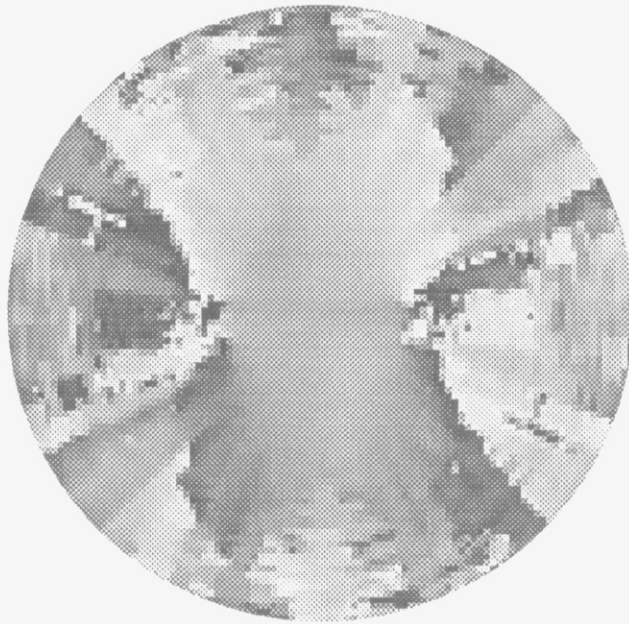
---

bispectrum directly method the number of operations scales roughly as  $M^2$  where  $M$  is side length of our detector array in pixels.

Thus in the case of the radon transform where  $M$  is small, direct calculation of the bispectrum is likely to be most efficient, unless the number



Reconstructed  
Image



Reconstructed  
Fourier Phase

*Reconstruction by radon method from simulated data for a translating object. The data contains 2000 frames, of on average 10 photon events per frame.*

**Fig 5.7**

---

of photon events per frame is very low.

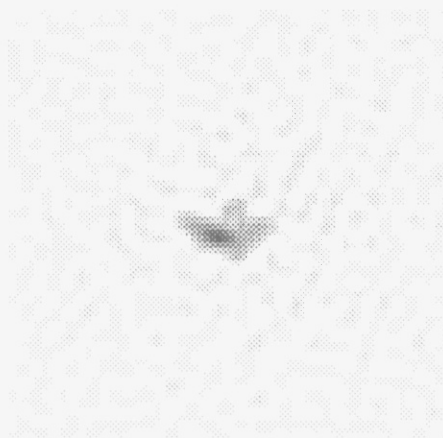
#### 5.4. Sub plane calculation.

In the direct method we integrate or average the bispectrum directly, and for the radon method direct calculation of the bispectrum is in general much more efficient than using a photon-differencing algorithm. We saw in Chapter three that because we can more easily make use of the temporal information in a photon event stream, there is a SNR advantage to using a photon-differencing algorithm to compute the average TC, compared with computing the average bispectrum, which is most conveniently done on a framed basis. We also indicated that due to the enormous size of a four dimensional TC array the calculation of the TC using straightforward photon differencing algorithm was impractical. It is however possible to formulate a modified photon-differencing algorithm which ultimately produces estimates for two dimensional sub-planes of the four dimensional bispectrum. Using this method it is possible, at the expense of an increased calculation burden to reduce vastly the memory requirement of a TC photon differencing algorithm.

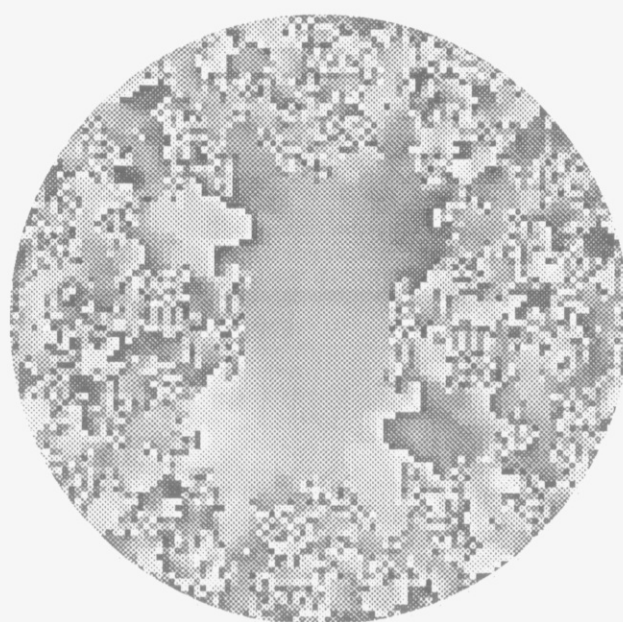
We will start by writing down the expression for a photon limited TC;

$$i^{(3)}(x_1, x_2) = \int_{-\infty}^{\infty} dx \sum_{k=1}^{k=N} \delta(x-x_k) \sum_{\substack{l=1 \\ l \neq k}}^l \delta(x+x_1-x_l) \sum_{\substack{m=1 \\ m \neq k \\ m \neq l}}^m \delta(x+x_2-x_m) . \quad 5.4.1$$

The conditions  $l \neq k$  and  $m \neq k, m \neq l$  are necessary to eliminate the photon bias terms. The bias terms are described in detail by Wirnitzer <sup>33</sup>.



Reconstructed  
Image



Reconstructed  
Fourier Phase

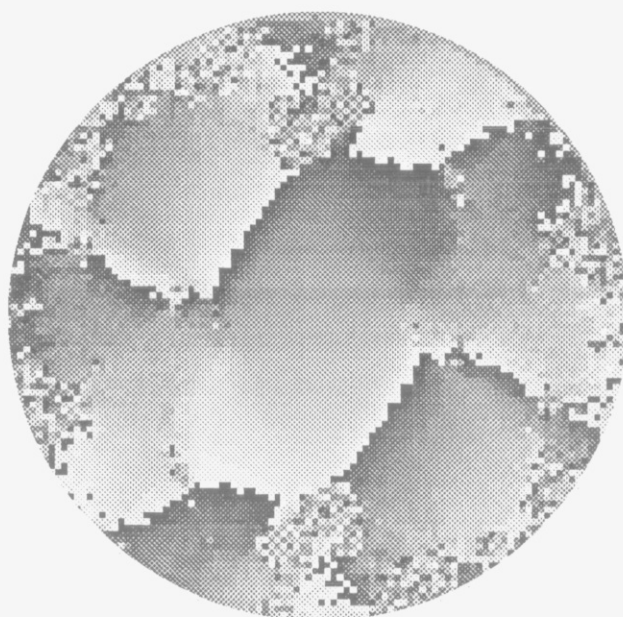
*Reconstruction by sub plane method from simulated speckle data using 12500 frames comprising on average of 200 photon events.*

**Fig 5.8**

---



Reconstructed  
Image



Reconstructed  
Fourier Phase

*Reconstruction by sub-plane method from simulated data for a translating object. The data contains 2000 frames, of on average 10 photon events per frame.*

**Fig 5.9**

---

The integrand is non-zero if and only if:

$$\mathbf{x}_k = \mathbf{x} , \quad 5.4.2$$

$$\mathbf{x}_l = \mathbf{x} + \mathbf{x}_1 \quad \Rightarrow \quad \mathbf{x}_l = \mathbf{x}_k + \mathbf{x}_1 , \quad 5.4.3$$

and

$$\mathbf{x}_m = \mathbf{x} + \mathbf{x}_2 \quad \Rightarrow \quad \mathbf{x}_m = \mathbf{x}_k + \mathbf{x}_2 . \quad 5.4.4$$

Using the conditions 5.4.3 and 5.4.4, equation 5.4.1 can be rewritten as follows.

$$i^{(3)}(\mathbf{x}_1, \mathbf{x}_2) = \sum_{k=1}^{k=N} \sum_{\substack{l=1 \\ l \neq k}}^N \sum_{\substack{m=1 \\ m \neq k \\ m \neq l}}^N \delta(\mathbf{x}_1 - (\mathbf{x}_l - \mathbf{x}_k)) \delta(\mathbf{x}_2 - (\mathbf{x}_m - \mathbf{x}_k)). \quad 5.4.5$$

Equation 5.4.5 clearly represents a photon-differencing process.

By direct application of this equation a 128 square image would normally generate a  $256 \times 10^6$  element TC, which would require about 1Gb of memory, or 170Mb even if we made full use of the symmetry of the TC.

Fortunately, at the cost of increasing the computational burden, (3 integer ops.  $\rightarrow$  8 floating point ops) we can compute a single two dimensional sub-plane of the bispectrum, using a modified photon-differencing technique. Essentially we do this by pre-computing two dimensions, of the four dimensional Fourier transform, which transforms the TC to the bispectrum. The result of this operation is to associate a complex weighting factor with each photon event. This is done for some constant Fourier vector  $\mathbf{V}$  producing the bispectrum sub-plane  $I^{(3)}(\mathbf{u}, \mathbf{V})$ . As already pointed out, the regions closest to the axes of the bispectrum have in

general the highest SNR. If we compute these high SNR regions by the photon differencing sub-plane method, the algorithm turns out, to be similar to the photon-differencing method of generating the Knox-Thompson transform<sup>19,34</sup>. It becomes clear from this formulation that the Knox-Thompson spectrum converges toward the bispectrum when the vector  $\mathbf{V}$  becomes small.

Using Eq. 5.4.5 the bispectrum may be written as:

$$I^{(3)}(\mathbf{u}, \mathbf{v}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} d\mathbf{x}_1 d\mathbf{x}_2 e^{-i\mathbf{u} \cdot \mathbf{x}_1} e^{-i\mathbf{v} \cdot \mathbf{x}_2} \times \sum_{k=1}^N \sum_{l=1}^N \sum_{m=1}^N \delta(\mathbf{x}_1 - (\mathbf{x}_l - \mathbf{x}_k)) \delta(\mathbf{x}_2 - (\mathbf{x}_m - \mathbf{x}_k)). \quad 5.4.6$$

where to aid clarity, we have omitted the conditions that eliminate the photon bias.

If we now put  $\mathbf{v}$  equal to a constant, say  $\mathbf{v} = \mathbf{V}$ , we may carry out the  $\mathbf{x}_2$  integration forthwith.

Since  $\mathbf{x}_2$  is non zero only at  $\mathbf{x}_2 = (\mathbf{x}_m - \mathbf{x}_k) :-$

$$I^{(3)}(\mathbf{u}, \mathbf{V}) = \int_{-\infty}^{\infty} d\mathbf{x}_1 e^{-i\mathbf{u} \cdot \mathbf{x}_1} \left[ \sum_{k=1}^N \sum_{l=1}^N \sum_{m=1}^N \delta(\mathbf{x}_1 - (\mathbf{x}_l - \mathbf{x}_k)) e^{-i\mathbf{V} \cdot (\mathbf{x}_m - \mathbf{x}_k)} \right]. \quad 5.4.7$$

Separating out the summation over  $m :-$

$$I^{(3)}(\mathbf{u}, \mathbf{V}) = \int_{-\infty}^{\infty} d\mathbf{x}_1 e^{-i\mathbf{u} \cdot \mathbf{x}_1} \left[ \sum_{m=1}^N e^{-i\mathbf{V} \cdot \mathbf{x}_m} \right] \times \left[ \sum_{k=1}^N \sum_{l=1}^N \delta(\mathbf{x}_1 - (\mathbf{x}_l - \mathbf{x}_k)) e^{i\mathbf{V} \cdot \mathbf{x}_k} \right]. \quad 5.4.8$$

Excluding the photon noise bias terms this becomes:

$$I^{(3)}(\mathbf{u}, \mathbf{V}) = \int_{-\infty}^{\infty} d\mathbf{x}_1 e^{-i\mathbf{u} \cdot \mathbf{x}_1} \left[ \sum_{m=1}^N (1 - \delta_{mk})(1 - \delta_{ml}) e^{(-i\mathbf{V} \cdot \mathbf{x}_m)} \right] \quad 5.4.9$$

$$\times \left[ \sum_{k=1}^N \sum_{l=1}^N \delta(\mathbf{x}_1 - (\mathbf{x}_l - \mathbf{x}_k)) - (1 - \delta_{kl}) e^{(i\mathbf{V} \cdot \mathbf{x}_k)} \right],$$

where  $\delta_{kl}$  is the kronecker delta.

The quantity inside the second set of large brackets can be calculated using a photon-differencing technique, where each vector is given a complex weighting  $e^{i(\mathbf{V} \cdot \mathbf{x}_k)} \sum e^{-i(\mathbf{V} \cdot \mathbf{x}_m)}$ . After finding the mean value of this quantity over the data set, the integration over  $\mathbf{x}_1$  (Fourier transform) is carried out, yielding the bispectrum subplane  $I^{(3)}(\mathbf{u}, \mathbf{V})$ .

We could of course put  $\mathbf{u}$  and  $\mathbf{v}$  equal to a constant, this would generate a single bispectrum value (this would be an inefficient version of the direct method). It is however more efficient to generate a whole plane at a time, thus replacing two complex multiplications per bivector per frame with a final Fourier transform after the averaging is complete.

For a two dimensional signal the minimum requirement for reconstructing an image is to have calculated the two bispectrum sub-planes  $\mathbf{V} = (0,1)$  and  $\mathbf{V} = (1,0)$ . In this minimal configuration the method is very closely related to the Knox-Thompson method. In general, in order to bolster signal-to-noise, we would calculate more than the minimum 2 sub-planes of the bispectrum.

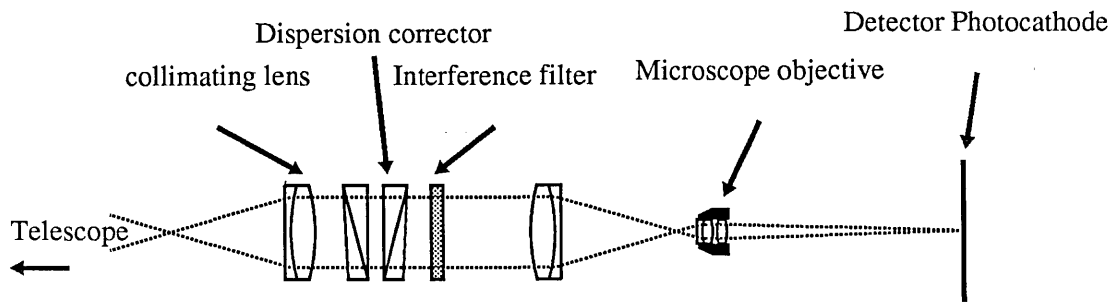
## **Imaging Through Atmospheric Turbulence.**

### **6.1. Introduction**

This chapter is concerned mainly with the use of the TC and bispectrum in the context of speckle interferometry. The angular resolution of large ground-based telescopes is seriously degraded by atmospheric turbulence. Reasonable seeing at a good astronomical site produces long exposure images of roughly equivalent size to those which would be formed by a 10-20cm telescope. The size of the telescope whose diffraction limit is roughly the same size as a long exposure image, is called the Fried's parameter, usually denoted by the symbol  $r_0$ . Thus for a telescope of 4m aperture, we could hope for a twenty to forty times improvement in resolution, if in spite of atmospheric turbulence, we could form diffraction limited images. We will describe in this chapter how the TC and bispectrum can at least go some way towards solving this problem.

### **6.2. Speckle interferometry**

Speckle interferometry was first suggested by A.Labeyrie<sup>4</sup>, as a means of obtaining diffraction limited information from large telescopes, in spite of the deleterious effects of atmospheric seeing. The first step in this interferometric technique is to record a large number of short exposure images of stellar speckle patterns. Light from the object under study is passed through a narrow bandwidth filter (typically 10-25nm) to increase its temporal coherence, and then a high magnification image is



*Typical optical system for carrying out optical speckle interferometry. Notice the atmospheric dispersion corrector. At about 60 degrees from the zenith, using a 10nm bandwidth, dispersion would produce an elongation of approximately 100 milli arc-seconds, in a stellar image. Thus if we are trying to image to the 35 milli arc-second resolution of a 4m telescope, correction of atmospheric dispersion is critical.*

**Fig 6.1**

formed on a two dimensional imaging detector, for example a film, CCD or IPD. Fig 6.1 shows the sort of optical system which would be used to observe stellar speckle. Under these circumstances the image of an unresolved object can be seen to be composed of a rapidly evolving speckle pattern. The envelope of the speckle pattern is of approximately the same size as a long exposure image of the object. For the case of an unresolved object the speckle size is of the order of the theoretical point spread function of the unaberrated telescope. Figure 6.2 shows some typical speckle patterns for both resolved and unresolved sources. Typically

the coherence time of such a speckle pattern is between 1 and 20 milliseconds<sup>35</sup>, Fig 6.3 shows some typical speckle temporal coherence curves measured at several observing sites. For a framed detector it has been shown that for maximum SNR in speckle interferometry an effective exposure time of about twice the speckle coherence time is optimal<sup>36,37</sup>.

The intensity distribution of a speckle pattern can be written in the following form;

$$i(x) = o(x) \otimes p(x), \quad 6.2.1$$

where;

$i(x)$  is the observed speckle pattern.

$o(x)$  is the intensity distribution of the object.

$p(x)$  is the instantaneous point spread function of the atmosphere/telescope combination.

In speckle interferometry many short exposure images are recorded and the mean of their auto-correlations, or equivalently, power-spectra is computed. The resulting function can be shown to contain diffraction limited information.

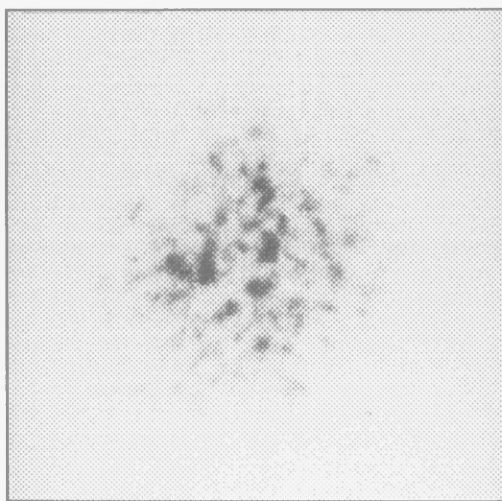
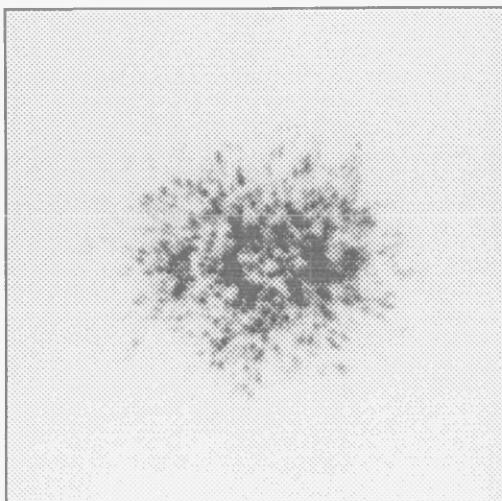
If as previously, we use capital letters to represent the Fourier transform of a function denoted by the corresponding lower-case letters, eq 6.2.1 gives;

$$I(u) = O(u) \times P(u), \quad 6.2.2$$

which implies;

$$|I(u)|^2 = |O(u)|^2 \times |P(u)|^2. \quad 6.2.3$$

In speckle interferometry the quantity  $|I(u)|^2$  of Eq. 6.2.3 is averaged over the atmospheric statistics;



*This figure shows typical astronomical speckle patterns. The first speckle pattern is produced by observing an unresolvable star. The second speckle pattern is due to observing a resolvable object under similar conditions (the object in this case is  $\alpha$ -Orionis). Note that the speckle pattern of the resolvable object is 'blurred' in comparison with that from the unresolvable object. This blurring is caused because to first order we can think of each "speckle" as being a distorted diffraction limited image of the object.*

**Fig 6.2**

---

$$\langle |I(\mathbf{u})|^2 \rangle = \langle |O(\mathbf{u})|^2 \times |P(\mathbf{u})|^2 \rangle. \quad 6.2.4$$

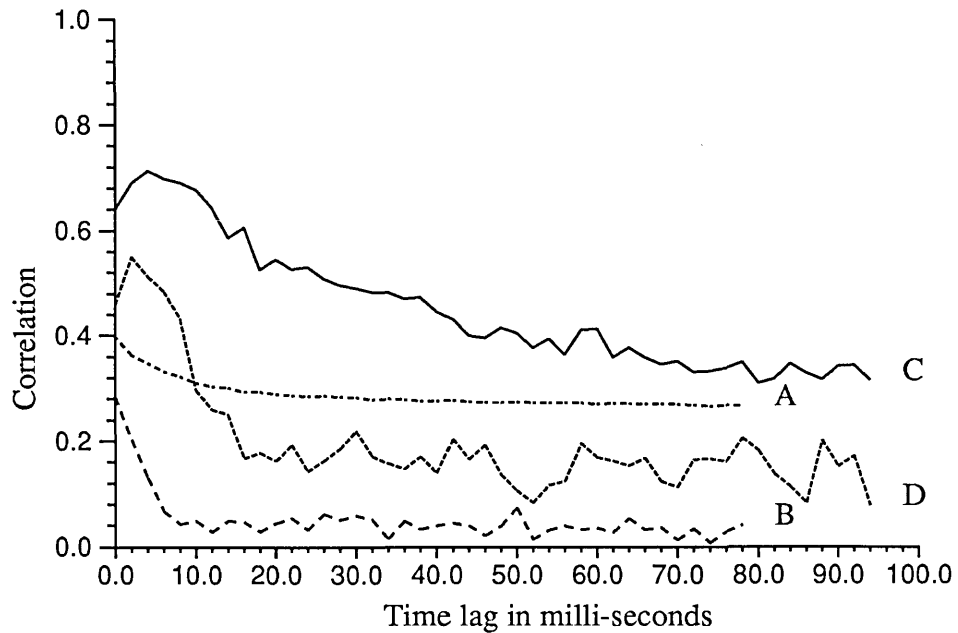
We now note that  $O(\mathbf{u})$ , the Fourier transform of the object intensity profile, is constant through all frames and can therefore be taken outside the averaging operator;

$$\langle |I(\mathbf{u})|^2 \rangle = |O(\mathbf{u})|^2 \times \langle |P(\mathbf{u})|^2 \rangle. \quad 6.2.5$$

The quantity  $\langle |P(\mathbf{u})|^2 \rangle$  can be shown to be non zero (but small) out to the diffraction limit of the telescope. It is known as the speckle transfer function. Under stable seeing conditions  $\langle |P(\mathbf{u})|^2 \rangle$  may be determined by observing a point source, i.e.  $O(\mathbf{u}) = \delta(\mathbf{x})$  close to the object of interest. In general  $\langle |P(\mathbf{u})|^2 \rangle$  is a non-zero positive function and hence by simple division ( or more complex processing ) we can extract  $|O(\mathbf{u})|^2$  from  $\langle |I(\mathbf{u})|^2 \rangle$ , see Eq. 6.2.5.

The major drawback with this technique is that it furnishes only the Fourier modulus, or equivalently the auto-correlation of the object under observation. In practice this means that it can only be usefully applied to the observation of relatively simple objects, for instance binary stars where the modulus can be used to determine the binary separation and angle, or for measuring stellar diameters.

If we are given an auto-correlation, the problem of finding the generator function for that auto-correlation is known as the phase problem. This is because in Fourier space, the equivalent problem is to find the Fourier phase of a function, given only its Fourier modulus, and possibly the knowledge that the function is positive and real. It is known that for a discrete two dimensional function, provided that we have complete information, there is usually a unique solution to this problem. Thus theoretically it is possible to reconstruct an image by the use of speckle



*Some typical temporal correlation curves for speckle interferometry data. Curves A and B come from PAPA camera data taken at La Silla, Chile (courtesy of P. Nisenson) and curves C and D from IPD data taken at the WHT on La Palma. Notice that the IPD data curves have a dip at the origin which is not present in the PAPA data curves. This dip is due to the effects of local dead-time. Theoretically, the correlation curves should asymptotically approach zero at large lags; the primary reason that these curves do not all approach zero is due to extremely long time scale intensity fluctuations due to cloud and/or dust partially occluding the star. These curves should also approach 1 at the origin; many effects contribute to reduce this value, the most significant is probably spatial averaging of the speckle pattern.*

**Fig 6.3**

interferometry. However, in practice it appears to be extremely difficult to solve the phase problem, especially in the presence of noise. Thus in practice, in order to obtain images from speckle interferometry, we have to obtain phase, as well as modulus information.

### 6.3. Recovery of phase information

Many different techniques that allow the diffraction limited Fourier phase of an image to be measured, in spite of the deleterious effects of atmospheric turbulence have been proposed. These methods, can in general be placed in one of two categories, depending upon whether measurements are recorded in the pupil plane, or the image plane.

Examples of image plane techniques are:

Shift and add <sup>38,39</sup>.

Exponential filter <sup>40,41</sup>.

Phase averaging <sup>42,43,44</sup>.

Phase differential <sup>45</sup>.

Knox-Thompson (Cross power spectrum) <sup>46</sup>.

Triple correlation <sup>20</sup>.

Examples of pupil plane interferometric techniques are:

Non redundant apertures using phase closure <sup>5,47,48,49</sup>.

Pupil plane shearing interferometry <sup>50,51</sup>.

### 6.4. Triple correlation

Of the image plane methods the Knox-Thompson and triple correlation techniques have been used most extensively for astronomical imaging. The Knox-Thompson in particular is used quite routinely in the analysis of one dimensional infrared data and by the Harvard group for analysis of two dimensional visible wavelength data from a PAPA detector <sup>52</sup>. The Knox-Thompson and triple correlation techniques are very closely related, the Knox-Thompson transform being very similar to a sub-plane of the

bispectrum <sup>53</sup>.

As in the case of the power-spectrum analysis, we write the bispectrum of a single speckle pattern as;

$$I^{(3)}(\mathbf{u}, \mathbf{v}) = O^{(3)}(\mathbf{u}, \mathbf{v}) \times P^{(3)}(\mathbf{u}, \mathbf{v}), \quad 6.4.1$$

where the superscript <sup>(3)</sup> indicates a bispectrum term (Fourier transform of a third order correlation) and where the object bispectrum  $O^{(3)}(\mathbf{u}, \mathbf{v})$  is given by the expression;

$$O^{(3)}(\mathbf{u}, \mathbf{v}) = O(\mathbf{u}) \times O(\mathbf{v}) \times O^*(\mathbf{u} + \mathbf{v}), \quad 6.4.2$$

where  $O(\mathbf{u})$  is the Fourier transform of the object surface intensity distribution. Averaging over the atmospheric statistics, gives;

$$\langle I^{(3)}(\mathbf{u}, \mathbf{v}) \rangle = \langle O^{(3)}(\mathbf{u}, \mathbf{v}) \times P^{(3)}(\mathbf{u}, \mathbf{v}) \rangle. \quad 6.4.3$$

Since  $O(\mathbf{u})$  is constant over all speckle patterns we may take its bispectrum outside the averaging operator.

$$\langle I^{(3)}(\mathbf{u}, \mathbf{v}) \rangle = O^{(3)}(\mathbf{u}, \mathbf{v}) \times \langle P^{(3)}(\mathbf{u}, \mathbf{v}) \rangle. \quad 6.4.4$$

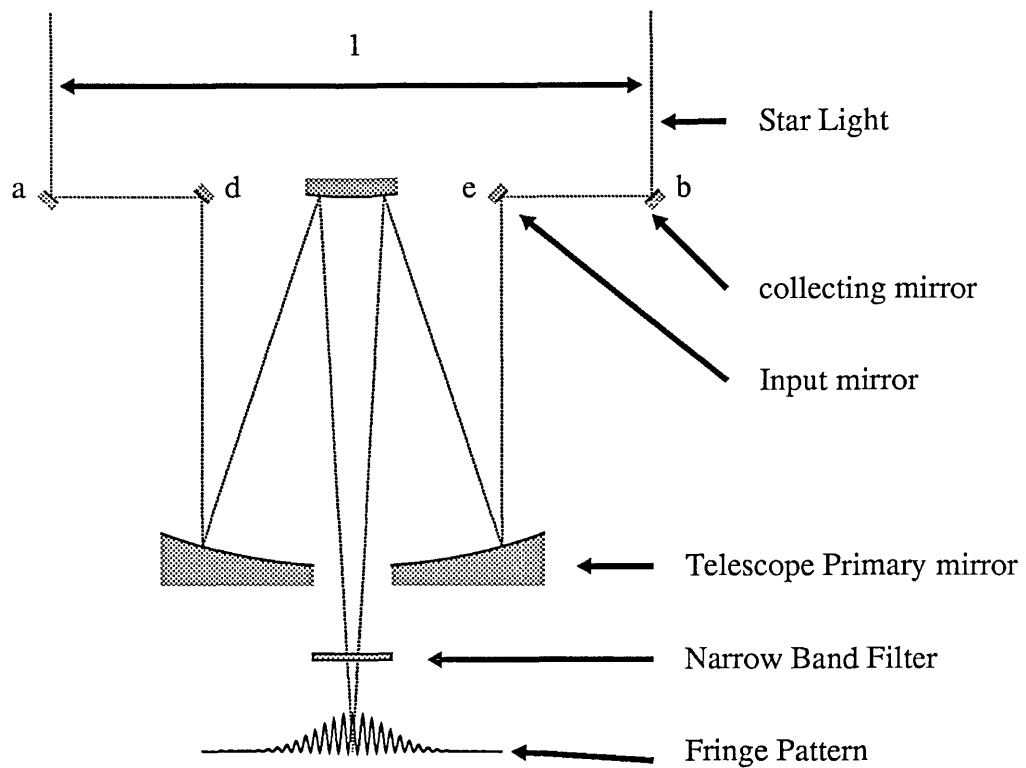
In this case it can be shown that the bispectral transfer function  $P^{(3)}(\mathbf{u}, \mathbf{v})$  is non-zero out to the diffraction limit of the telescope. Furthermore, provided that the telescope aberrations are small compared with those introduced by the atmosphere, the bispectral transfer function is real and positive. Under these conditions the value of bispectral speckle transfer function does not affect the phase of the estimated bispectrum, so for the purpose of phase reconstruction, need not be compensated. Theoretically then, it is not necessary to compensate for the bispectral transfer function in order to reconstruct the object phase from a speckle derived bispectrum. In practice it would appear that this condition is usually fulfilled by most telescopes, so no bispectrum correction is needed. As in the

previous case, if necessary, the transfer function  $\langle P^{(3)}(\mathbf{u}, \mathbf{v}) \rangle$  of this process can be calculated by observing a point source.

### 6.5. Phase closure and triple correlation

The phase closure technique has been used to great effect in radio-astronomy for a number of years. There has recently been much interest in implementing this technique at optical wavelengths<sup>48</sup> and this has led to a realisation that most of the speckle imaging methods can be interpreted in terms of phase closure relationships. In particular this technique makes it easy to see why the bispectral speckle transfer function should be non-zero out to the diffraction limit of the telescope.

To understand phase closure let us first look at the Michelson stellar interferometer (Fig 6.4). Light from an astronomical object is collected by two small mirrors, 'a' and 'b' separated by a large distance  $l$ . The light from these mirrors is passed through a narrow bandwidth filter and then brought to a common focus. Under these circumstances an image of the object will be formed. This image will be of extremely low resolution since its point-spread function will be that of the small collecting mirrors. However this image will be crossed by fine fringes, whose frequency is proportional to the separation of the two input mirrors 'd', 'e', provided the magnification of the imaging system is kept constant. According to the Van Cittert-Zernike theorem, the visibility of these fringes will be proportional to the magnitude of the Fourier transform of the brightness distribution of the object under observation. Similarly the phase of these fringes will be the Fourier phase of the object brightness distribution. Unfortunately due to atmospheric phase fluctuations and unknown telescope aberrations the phase of the fringes will be essentially random.

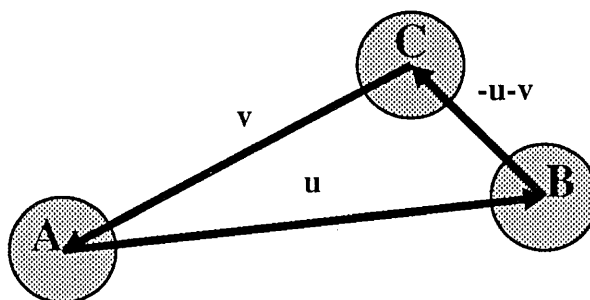


*Diagrammatic representation of a Michelson or Fizeau stellar interferometer.*

**Fig 6.4**

Furthermore the atmospheric phase will vary rapidly over time causing the fringe pattern to move erratically, necessitating the use of a fast exposure camera if accurate visibility measurements are to be made. We could add here that even in the absence of atmospheric aberrations, it would be very difficult to measure the object Fourier phase over a wide range of spatial frequencies by this method. Such measurements would require that the optical path lengths within the instrument be measurable within a fraction

of a wavelength over distances of tens of meters. When measuring visibility, the metrology requirement is reduced to being well within the temporal coherence length of the light.



*Three sub apertures in a mask used to make phase closure measurements. The phases of the resulting measurements must be computed in the direction of the arrows (or the opposite direction), in order to generate a phase closure.*

Fig 6.5

---

If we add a third mirror 'c' to a Michelson interferometer, the resulting image will have three separate fringe patterns super-imposed upon it. The visibilities and phases of these fringe patterns can be determined, by Fourier transforming the image. In this way it is possible to determine the phase and magnitude of the Fourier transform of the object intensity distribution at three frequencies simultaneously.

If we look at a single short exposure recording of such a fringe pattern, we can write down the following three equations:

$$\phi_{ab}^{\text{measured}} = \phi_{ab}^{\text{obj}} + \phi_a^{\text{atmos}} - \phi_b^{\text{atmos}}, \quad 6.5.1$$

$$\phi_{bc}^{\text{measured}} = \phi_{bc}^{\text{obj}} + \phi_b^{\text{atmos}} - \phi_c^{\text{atmos}}, \quad 6.5.2$$

$$\phi_{ca}^{\text{measured}} = \phi_{ca}^{\text{obj}} + \phi_c^{\text{atmos}} - \phi_a^{\text{atmos}}, \quad 6.5.3$$

where for instance,  $\phi_{ca}$  is the measured phase of the fringe pattern due to the two apertures 'c' and 'a',  $\phi_{ca}^{\text{obj}}$  is the phase of this fringe pattern that we would measure in the absence of any atmospheric or system aberrations and  $\phi_a^{\text{atmos}}, \phi_c^{\text{atmos}}$ , are the unknown atmospheric/geometric aberrations of each collecting mirror. Adding the above three equations together, we see that the terms due to atmospheric and other aberrations cancel out leaving us with;

$$(\phi_{ab} + \phi_{bc} + \phi_{ca})^{\text{measured}} = (\phi_{ab} + \phi_{bc} + \phi_{ca})^{\text{obj}}. \quad 6.5.4$$

Thus we can measure the quantity  $(\phi_{ab} + \phi_{bc} + \phi_{ca})^{\text{obj}}$  even in the presence of system aberrations— this measurement is called a phase closure measurement.

Moving to the Fourier plane of the object brightness distribution, the Fourier component associated with, for example the baseline  $\overline{ab}$  between the mirrors 'a' and 'b' can be represented by the vector  $k\overline{ab}$  in the Fourier plane,  $k$  is a scaling constant. Since mirrors 'a', 'b', 'c' must form a triangle (we consider lines and points as collapsed triangles) the following relationship must be true:

$$k\overline{ab} + k\overline{bc} + k\overline{ca} = 0. \quad 6.5.5$$

or;

$$k\overline{ab} + k\overline{bc} = -k\overline{ca}. \quad 6.5.6$$

Writing  $\mathbf{u} = k\overline{ab}$  and  $\mathbf{v} = k\overline{bc}$  then we can re-write equation 6.5.4 as follows:

$$(\phi_{ab} + \phi_{bc} + \phi_{ca})^{\text{measured}} = \text{Phase of } \left[ I(u)I(v)I^*(u+v) \right] \quad 6.5.7$$

Where  $I()$  is the Fourier spectrum of the object brightness distribution.

The quantity on the right hand side of Eqn. 6.5.7 is readily identifiable as an element of the bispectrum of the object brightness distribution. Thus the methods of phase closure and triple correlation although at first sight radically different are in fact very closely related, by virtue of the fact that the same primary quantity is being measured in both methods.

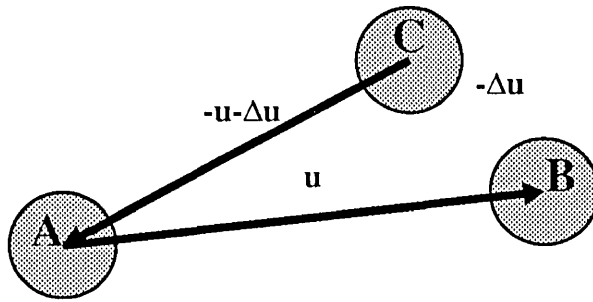
In radio-astronomy, phase closure techniques are used between widely separated telescopes, allowing baselines approaching the diameter of the Earth to be used, and thus very high resolution maps to be produced.

In the optical regime phase closure measurements have been produced by placing masks over the apertures of single large telescopes. This limits the resolution obtainable to about 0.025 arc seconds. However several groups are at present working on phase closure systems using several small independent telescopes to collect light, allowing potential baselines of tens or even hundreds of meters to be reached.

For the above explanation to hold, each of the telescopes or light collecting mirrors (siderostats) must be small enough that atmospherically induced aberrations over the aperture are much less than a wavelength. In the case of radio astronomy this limit is much greater than the size of any present radio telescope. In the optical regime if we assume that tilt aberrations are removed this condition still imposes a maximum aperture of the order of 20cm even for the very best observing sites.

Furthermore it is not necessary for the path to be perfectly closed. For an optical telescope, the light field in the aperture at points separated

by the order of Frieds parameter  $r_0$  or less, will be strongly correlated, thus any path which closes to within this tolerance will have a phase closure component.



*The Knox Thompson Transform makes use of only two of the fringe phase values which can be measured from three apertures. Hence the Knox Thompson transfer function is non-zero only if the length of vector  $\overline{cb}$  is less than Frieds parameter  $r_0$ .*

**Fig 6.6**

---

If we think of placing two small apertures in the aperture plane of a large telescope, we will have formed a Michelson interferometer. Again, provided that the apertures are small enough that atmospheric aberrations over each aperture are approximated well by a tilt term, the van Cittert-Zernike theorem states that the visibility of the fringe pattern produced by such an interferometer is proportional to the modulus of the Fourier transform of the target object intensity distribution, at the fringe frequency of the interferometer. The phase of the fringe pattern is of course random, being almost totally dominated by atmospheric contributions. Each

Fourier component of a speckle pattern is derived from a summation of all such fringe patterns over all aperture pairs of the corresponding separation and orientation. Thus we can write;

$$I(\mathbf{u}) = \sum_{j=1}^{j=N} V_j(\mathbf{u})e^{i\phi_j}. \quad 6.5.8$$

where  $V_j(\mathbf{u})$  is the visibility of a Michelson interferometer fringe pattern at frequency  $\mathbf{u}$  and  $\phi_j$  its associated phase. A bispectrum value is formed from three such Fourier values, thus:

$$\begin{aligned} I^{(3)}(\mathbf{u}, \mathbf{v}) &= I(\mathbf{u})I(\mathbf{v})I^*(\mathbf{u}+\mathbf{v}). \\ &= \sum_{j=1}^{j=N} V_j(\mathbf{u})e^{i\phi_j} \sum_{k=1}^{k=N} V_k(\mathbf{v})e^{i\phi_k} \sum_{l=1}^{l=N} V_l(\mathbf{u}+\mathbf{v})e^{-i\phi_l} \end{aligned} \quad 6.5.9$$

We define  $j = k = l$  to be the condition that the sub-apertures leading to estimates of  $V_j(\mathbf{u})e^{i\phi_j}$  etc, form a closed loop in the aperture plane, as shown in Figure 6.5. That is; the statement  $j = k$ , means that measurements of  $V_j(\mathbf{u})e^{i\phi_j}$  and  $V_k(\mathbf{u})e^{i\phi_k}$  have at least one sub-aperture in common. When  $j = k = l$  Eqn. 6.5.9 gives a phase closure term as expressed in Eqn. 6.5.4. All other cross terms from Eqn. 6.5.9 have random phase and so should average to zero if enough bispectrum estimates are averaged. We will return to this equation in section 6.6 where we discuss the SNR of the power-spectrum and bispectrum.

We see therefore that, although phase closure measurements and triple correlation measurements are made in very different ways, they are essentially measuring the same quantity. Fig. 6.6 shows the phase closure diagram for the Knox Thompson cross correlation

$$I^{KT}(\mathbf{u}, \Delta\mathbf{u}) = I(\mathbf{u})I^*(\mathbf{u}+\Delta\mathbf{u}). \quad 6.5.10$$

From Fig. 6.6 and 6.5 it is apparent that provided  $\Delta\mathbf{u} < r_0$  the Knox

Thompson correlation will have a phase closure component, since the phase of the fringes due to  $\overline{bc}$  will be nearly constant, provided  $|\overline{bc}| < r_0$ .

## 6.6. Signal to noise

The signal to noise calculation for triple correlation imaging through atmospheric turbulence follows closely that given by Goodman and Belsher<sup>24</sup> for the power-spectrum. A complete and rigorous derivation of the SNR of both the power spectrum and the bispectrum is given in the Thesis of G.R. Ayers<sup>19</sup>.

Looking back to Chapter three, we find that the SNR of the photon limited power-spectrum is given by the expression;

$$\text{SNR} = \frac{\overline{N} I^{(2)}(\mathbf{u})}{\left[ \left( 1 + \overline{N} I^{(2)}(\mathbf{u}) \right)^2 + I^{(2)}(2\mathbf{u}) + \overline{N} \left[ I(2\mathbf{u}) I^*(\mathbf{u})^2 + I^*(2\mathbf{u}) I(\mathbf{u})^2 \right] \right]^{\frac{1}{2}}} \quad 6.6.1$$

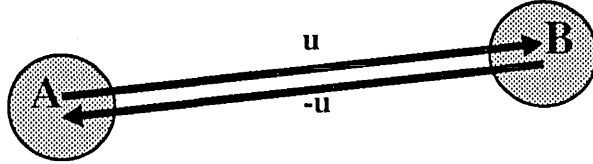
To find the SNR of the power-spectrum in the case of the photon limited speckle interferometry, we must average this expression over the atmospheric statistics. If the instantaneous amplitude transfer function of the atmosphere/telescope combination is written as  $P(\mathbf{u})$  then we must make the following substitution in equation 6.6.1;

$$I(\mathbf{u}) = O(\mathbf{u}) \times P(\mathbf{u}), \quad 6.6.2$$

where  $O(\mathbf{u})$  is the spectrum of the object we are observing.

Providing  $\mathbf{u}$  is much greater than the atmospheric correlation length  $r_0$ , which is commonly called Fried's seeing parameter<sup>50,51</sup>, then  $I(\mathbf{u})$  and  $I^*(2\mathbf{u})$  will be uncorrelated. Here we are also making an implicit assumption that size of the atmospheric aberrations is much greater than the telescope aberrations. When these conditions are fulfilled only power spectrum like terms will remain after the variance is averaged over the

atmospheric transfer function statistics.



*This arrangement of sub apertures gives us a power spectrum term.*

**Fig 6.7**

That is, the complex terms in equation 6.6.1, whose phase is dependent upon the instantaneous atmospheric transfer function, will average to zero. Thus, for  $u > r_0$ , equation 6.6.1 now becomes;

$$\text{SNR} = \frac{\bar{N} I^{(2)}(u)}{\left[ \left( 1 + \bar{N} I^{(2)}(u) \right)^2 + I^{(2)}(2u) \right]^{\frac{1}{2}}} \quad 6.6.3$$

where;

$$I^{(2)}(u) = O^{(2)}(u) \times S^{(2)}(u). \quad 6.6.4$$

The speckle transfer function  $S^{(2)}(u)$  is simply the averaged modulus of the instantaneous atmosphere/telescope transfer function, taken over the atmospheric statistics.

$$S^{(2)}(u) = \langle P(u)P^*(u) \rangle. \quad 6.6.5$$

One can show that to first approximation, provided that  $u$  is much greater

than  $r_0$ , the value of speckle transfer function is approximately the inverse of the mean number of speckles per short exposure image times the theoretical telescope modulation transfer function  $T_D(\mathbf{u})$ . Making this approximation allows us to rewrite equation 6.6.3

$$\text{SNR} = \frac{\bar{M} T_D(\mathbf{u}) O^{(2)}(\mathbf{u})}{\left[ \left[ 1 + \bar{M} T_D(\mathbf{u}) O^{(2)}(\mathbf{u}) \right]^2 + O^{(2)}(2\mathbf{u}) \right]^{\frac{1}{2}}}, \quad 6.6.6$$

where  $\bar{M}$  is the average number of photon events per speckle.

An intuitive way of seeing why the transfer function is reduced proportionately to the number of speckles is provided by a phase closure type argument. Figure 6.7 indicates the arrangement of subapertures needed to measure a power spectrum term, namely a Michelson interferometer. We see that even this operation can be viewed as a special case of phase closure measurement, where sub-apertures B and C happen to be coincident. When we use the whole telescope aperture to find the power-spectrum at vector  $\mathbf{u}$ , we sum the contributions of all aperture pairs of the required separation. That is, as in equation 6.5.5, the Fourier transform of a speckle image at frequency  $\mathbf{u}$  is given by.

$$I(\mathbf{u}) = \sum_{j=1}^{j=N} V_j(\mathbf{u}) e^{i\phi_j}, \quad 6.6.7$$

where  $N$  is the number of distinct (independent) sub-aperture pairs of separation  $\mathbf{u}$  which can exist within the telescope pupil, and  $V_j(\mathbf{u}) e^{i\phi_j}$  is the fringe pattern due to the  $j^{\text{th}}$  sub-aperture pair. Sub-apertures which are separated by less than the Fried's parameter will give strongly correlated signals. Thus to be 'distinct', aperture pairs must be separated by a distance greater than the Fried's parameter. The number of aperture pairs  $N$

$\propto \left[ \frac{D}{r_0} \right]^2 \times T_D(\mathbf{u})$ , where  $T_D(\mathbf{u})$  is the transfer function of the unaberrated

telescope (aperture overlap integral). We may identify the term  $\left[ \frac{D}{r_0} \right]^2$  as being roughly equal to the number of speckles in a short exposure image. The power spectrum is defined as;

$$O^{(2)}(\mathbf{u}) = I(\mathbf{u}) \times I^*(\mathbf{u}). \quad 6.6.8$$

Substituting in equation 6.6.8 using the definition of  $I(\mathbf{u})$  from equation 6.6.7, we find;

$$O^{(2)}(\mathbf{u}) = \sum_{j=1}^{j=N} V_j(\mathbf{u}) e^{i\phi_j} \times \sum_{k=1}^{k=N} V_k(\mathbf{u}) e^{-i\phi_k}. \quad 6.6.9$$

This gives  $N$  term where  $j = k$  (as shown in Figure 6.7), these are true power spectrum terms which must have zero phase and thus add coherently to give us our signal. We also have  $N(N-1)$  cross terms between disparate pairs of sub-apertures where  $j \neq k$ . The  $N(N-1)$  cross terms have random phase due to the effects of the atmosphere and thus do not add up coherently. They form a random walk in the complex plane and thus on average produce a contribution of energy proportional to  $\sqrt{N(N-1)}$ , or approximately proportional to  $N$ . Of course, without the effects of the atmosphere, all terms from equation 6.6.9 would have zero phase and add coherently and contribute to the signal. We can thus write down the following results:

1): For an unaberrated telescope;

Signal at frequency  $\mathbf{u}$  in the power spectrum, proportional to  $N^2$ .

2): For a telescope in the presence of atmospheric seeing;

Signal at frequency  $\mathbf{u}$  proportional to  $N$ .

Noise (due to cross terms) at frequency  $\mathbf{u}$  proportional to  $N$ .

We can draw two conclusions from this result. The first is that the energy in the power-spectrum at frequency  $\mathbf{u}$  in the case of speckle

interferometry, is a factor of  $\frac{1}{N}$  of the energy for a single unaberrated telescope of similar size. From this we conclude that the speckle transfer function for such a telescope is a factor of  $\frac{1}{N}$  less than the theoretical modulation transfer function of that telescope. The second point is that, since the signal and noise terms for speckle interferometry are both proportional to  $N$ , then the SNR of speckle interferometry can never exceed 1. This is in agreement with the high light level limit of Equation 6.6.6.

As mentioned previously we can use a similar set of arguments to derive the effect of photon noise on the bispectrum estimator. However, in this case the standard deviation estimator contains a term involving a six fold summation over photon exponentials; this summation results in  $6!$  or 720 terms. Much more detail of this calculation, in particular its averaging over speckle statistics in speckle interferometry can be found in the Thesis of G.R Ayers <sup>19</sup>. In the case of the bispectrum, both the signal and its associated noise are complex. Referring back to equation 6.5.6, which is the bispectrum analogue of the expression given for the power spectrum in equation 6.6.9, we can identify two sets of terms. We have  $N$  terms with  $j = k = 1$ , which add coherently to give us our bispectrum signal. We also have  $N(N-1)^2$  terms where either  $j \neq k$  or  $k \neq 1$  or both. Hence for the case of the bispectrum we can write down the following results:

1): For an unaberrated telescope;

Signal at frequency  $\mathbf{u}$  in the bispectrum, proportional to  $N^3$ .

2): For a telescope in the presence of atmospheric seeing;

Signal at frequency  $\mathbf{u}$  proportional to  $N$ .

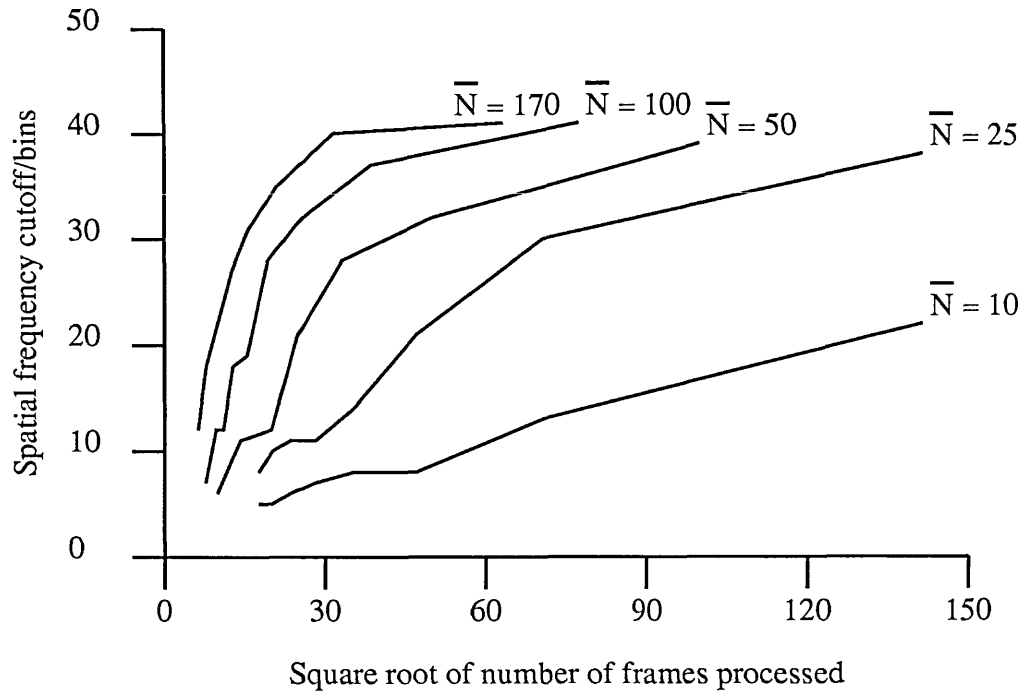
Noise (due to cross terms) at frequency  $\mathbf{u}$  approximately proportional

to  $N^2$ .

Thus the bispectrum transfer function is roughly a factor of  $\frac{1}{N^2}$  less than the transfer function one would get with the unaberrated telescope, in the absence of the atmosphere. Also in this case the maximum SNR obtainable would be  $\frac{1}{N}$ . Fortunately if one of the bispectrum vectors, for instance the vector  $(-\mathbf{u}-\mathbf{v})$  in Figure 6.5, is shorter than Fried's parameter, then cross terms in Equation 6.6.6, where  $j \neq k$  and/or  $j \neq 1$  will add coherently. In other words, these two sub-apertures will be sufficiently close together that the phase of the fringe pattern they produce will be largely unaffected by the atmosphere, and thus independent of the location of the aperture pairs. In this case, the bispectrum transfer function will only be approximately  $\frac{1}{N}$  down, i.e. it will be reduced by a similar amount as in the case of a power-spectrum measurement. Similarly the maximum SNR obtainable in a single measurement would increase to 1. In other words, close to the axes of the bispectrum, we see the same SNR and transfer characteristics as we see in the power spectrum.

## 6.7. Reconstruction of a simulated point source

Using the atmospheric simulation program written by G.R. Ayers<sup>19</sup> a sequence of simulated speckle patterns for various objects were generated. Some reconstructions from this data have already been shown in Chapter 5. This program was used to generate a set of speckle patterns for a point source viewed through a 2m telescope with 20cm seeing. The direct method of reconstruction, as described in Chapter 5, has been used to obtain these results. In order to determine the minimum number of frames of data needed to reconstruct an object using the bispectrum method,



*Cutoff frequency of a point source reconstruction for simulated speckle interferometry data. The simulation is for a 2m telescope with 20 cm seeing. Each of the curves represents a different mean number of photon events per frame.*

**Fig 6.8**

reconstruction of the object phase from this simulated data was attempted. This was considered to be the easiest possible phase recovery that was likely to be encountered in practice. In other words, when making real observations, the absolute minimum amount of data needed to recover useful information would be at least that needed to reconstruct the simulated point source. The results of this numerical experiment are presented in Figure 6.8. It can be seen that even for this extremely easy problem, at least 1000 frames of data at 170 photon events per frame are needed for

reconstruction.

## **6.8. Reconstruction of more complex objects**

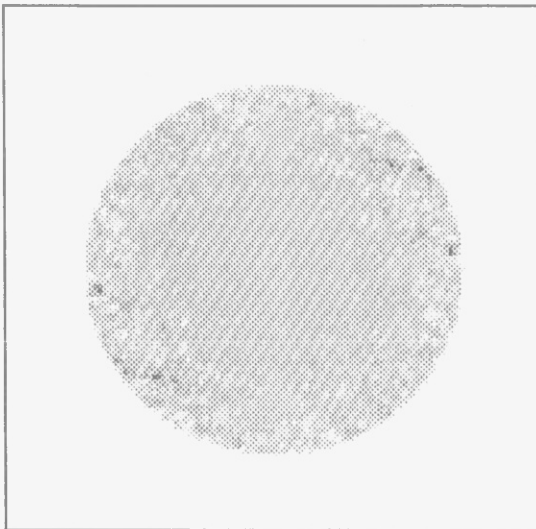
The following reconstruction was obtained from real speckle data provided by the Harvard Group. Although an extremely easy object to reconstruct, it demonstrates that the triple correlation algorithm can be used to image astronomical objects with high resolution.

The next page shows reconstructions of <sup>the</sup>  $\lambda$  wide spaced binary star  <sup>$\alpha$  PSC,  $M_v 4.3$ ,  
Sep 2"</sup> The data from which this reconstruction was made was provided by Clive Standley, of Harvard. The field of view in this case is approximately 9 arc seconds, meaning that the binaries are separated by approximately 1.6 seconds of arc, and are easily resolved on the long exposure picture. The data was taken using a 4m telescope, the speckle life time was approximately 10 ms, and the data rate 55,000 photon events per second. A 512 by 512 resolution PAPA detector was used as the detector, only the most significant 8 bits of the data being recorded for analysis. The frames on the following page are;

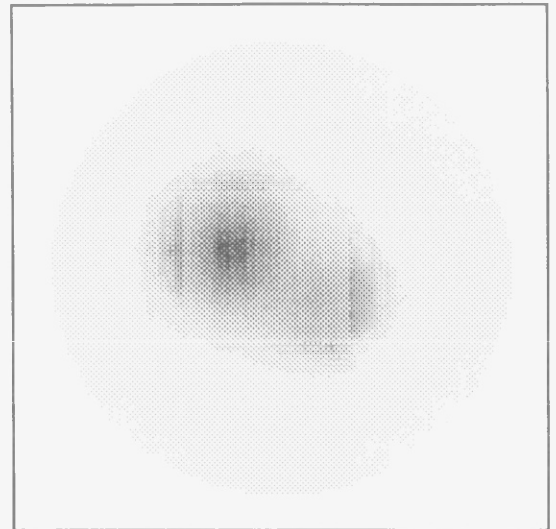
- A): Modulus of the binary star, recovered using traditional speckle interferometry methods.
- B): The long exposure image of the binary star. Note that under the prevailing seeing conditions the binary is resolved.
- C): Reconstructed phase using the radon technique.
- D): Reconstructed phase using the direct technique.
- E): Reconstructed object intensity map, using the radon technique. Note the radial structure in the background noise. This is due to errors recombination of the projected phase reconstructions, when the two dimensional phase is constructed.
- F): Reconstructed object using the direct technique.

Fig 6.9

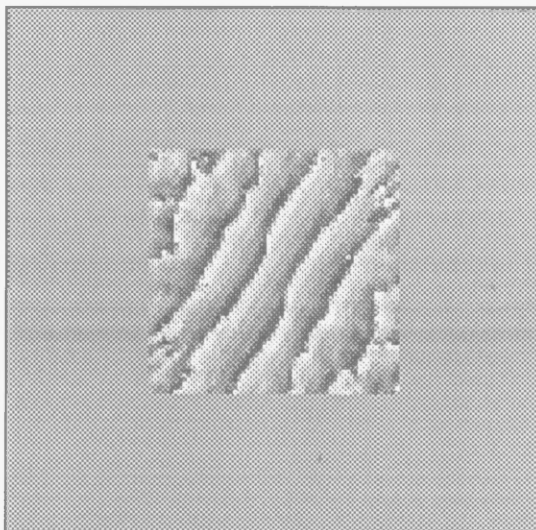
---



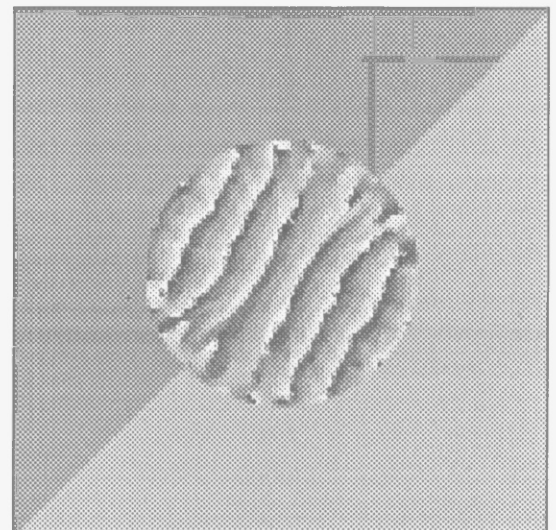
A



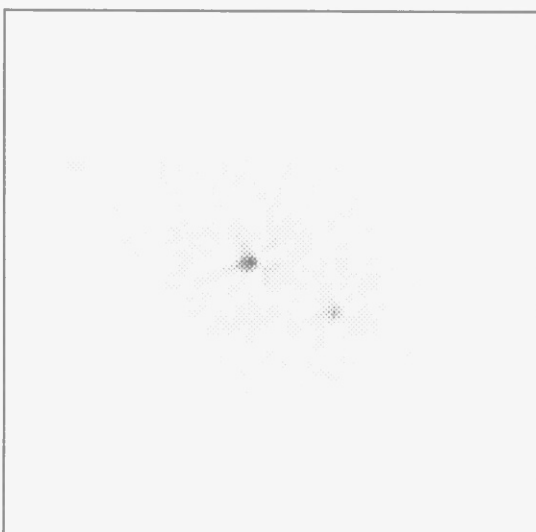
B



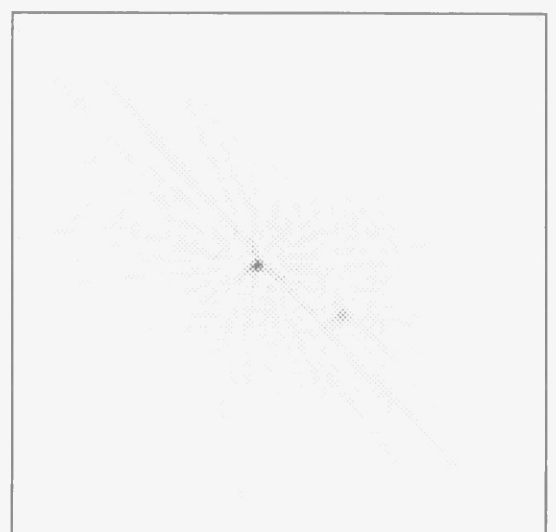
C



D



E



F

## 7

### **Conclusion.**

We have shown that the triple correlation can be used to produce images at extremely low light level, in the presence of image translation and/or the degrading influence of atmospheric turbulence. It has been demonstrated that, at least for the case of a randomly translating object, present imaging photon detectors are good enough to produce photon event data of sufficient quality to enable triple correlation type algorithms to be applied. The reconstructions obtained by the triple correlation route under these circumstances are much superior those obtained with a simple centroiding algorithm. One advantage of this type of detector over CCD detectors is that it is possible to give each detected photon event a time co-ordinate, as well as spatial coordinates. If full use is made of this time co-ordinate, it is shown that it is possible to increase the signal obtainable from a given data set. At first glance it would appear that application of the triple correlation to large two dimensional arrays would in practice be precluded by the large size of the four dimensional triple correlation. However we have shown that it is possible to process large arrays with the triple correlation algorithm. Three different algorithms have been investigated. Each has been coded and shown to work for imaging both in the case of a randomly translating object, and an object viewed through atmospheric turbulence. The field of astronomical speckle interferometry appears to be the most promising area for application of these techniques in the near future.

In the case of photon limited speckle interferometry we have briefly investigated the properties of the triple correlation with respect to signal-

to-noise ratio. This has been done with the help of an analogy based on the radio astronomy technique of phase closure. Finally a reconstruction of a binary star from speckle interferometry data is given.

In spite of the imminent launch of the space telescope, there is still motivation for pursuing the goal of ground based astronomical high resolution imaging. The space telescope only has a diameter of 2.5m, and hence its resolution is limited compared with the theoretical resolution of the largest ground based telescope. This gap will become wider still with as the next generation of 8 and 10m ground-based telescopes become available in the 1990's. ~~Furthermore, observation time on the space telescope is likely to very over-subscribed, leaving little time for high resolution imaging.~~ On the detector side, advances in avalanche photo-diodes promise to enable us to photon count at much higher quantum efficiencies and at longer wavelengths.

It would appear probable that the large baseline, multiple telescope interferometers, now under construction, or under study, by many groups, will provide the best means for very high resolution imaging. However there is still much scope for developing speckle interferometry techniques for large single telescopes, since these will in general be both quicker and easier to use than multiple telescope interferometers.

## APPENDIX A.

### A proof that the inter event statistics of a Poisson source are negative exponential.

For a Poisson source of mean  $\bar{M}$  events per second the probability of counting  $m$  events in time  $T$  is:

$$P_m(T) = \frac{(\bar{M}T)^m}{m!} e^{(-\bar{M}T)} . \quad A.1$$

If we start observing this process at some  $t$ , then the probability of our first observed event occurring at time  $T+t$  is the differential of  $P_0(t)$  wrt  $t$  at  $t = T$  (rate of change of  $m=0$  to  $m \neq 0$ ). From Equation A.1;

$$P_0(t) = e^{(-\bar{M}T)}, \quad A.2$$

thus

$$\frac{dP_0(T)}{dT} = \bar{M} e^{(-\bar{M}T)}, \quad A.3$$

giving us negative exponential inter-event statistics.

A more intuitive approach is to explicitly calculate the probability per unit time that a photon will arrive in the time interval  $T$  to  $T+\delta t$ , in the limit of  $\delta t$  going to zero:-

$$P_{arr}(T) = \lim_{\delta t \rightarrow 0} \frac{P_0(T) (P_1(T+\delta t) - P_1(T))}{\delta t} . \quad A.4$$

$$P_{arr}(T) = \lim_{\delta t \rightarrow 0} \frac{e^{(-\bar{M}T)} (\bar{M}(T+\delta t) e^{(-\bar{M}T+\delta t)} - \bar{M}T e^{(-\bar{M}T)})}{\delta t} . \quad A.5$$

$$P_{\text{arr}}(T) = \lim_{\delta t \rightarrow 0} \frac{e^{(-\overline{M}T)} \overline{M} \delta t e^{(-\overline{M}T)}}{\delta t} . \quad \text{A.6}$$

$$P_{\text{arr}}(T) = \overline{M} e^{(-\overline{M}T)}, \quad \text{A.7}$$

giving the same answer as above in Eqn. A.2.

## APPENDIX B.

### Derivation of the form of the bispectrum from the triple correlation.

In Chapter three the triple correlation is defined as;

$$i^{(3)}(x_1, x_2) = \int_{-\infty}^{\infty} i^*(x) i(x+x_1) i(x+x_2) dx . \quad B.1$$

The bispectrum  $I^{(3)}(u, v)$  is the Fourier transform of the triple correlation, thus;

$$I^{(3)}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^{(3)}(x_1, x_2) e^{-ix_1 u} e^{-ix_2 v} dx_1 dx_2. \quad B.2$$

Substituting Eqn. B.1 into B.2;

$$I^{(3)}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^*(x) i(x+x_1) i(x+x_2) dx e^{-ix_1 u} e^{-ix_2 v} dx_1 dx_2. \quad B.3$$

Now let us make the substitution  $y_1 = x+x_1$  since  $x$  and  $x_1$  are independent variables we can write  $dy_1 = dx_1$ . Proceeding with this substitution in Eqn. B.3 we write;

$$I^{(3)}(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^*(x) i(x+x_2) e^{-ix_2 v} e^{ixu} dx dx_2 \int_{-\infty}^{\infty} i(y_1) e^{-iy_1 u} dy_1 . \quad B.4$$

Recognising the integral over  $y_1$  as simple Fourier transform;

$$I^{(3)}(u, v) = I(u) \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} i^*(x) i(x+x_2) e^{-ix_2 v} e^{ixu} dx dx_2 . \quad B.5$$

If we make a similar substitution for  $x+x_2$  and proceed as before we arrive at;

$$I^{(3)}(u, v) = I(u) I(v) \times \int_{-\infty}^{\infty} i^*(x) e^{ix(u+v)} dx . \quad B.6$$

Now we must notice that;

$$\int_{-\infty}^{\infty} i^*(x)e^{ix(u+v)}dx = \left[ \int_{-\infty}^{\infty} i(x)e^{-ix(u+v)}dx \right]^* . \quad \text{B.7}$$

Using Eqn. B.7 in Eqn. B.6 we finally obtain;

$$I^{(3)}(u,v) = I(u)I(v)I^*(u+v). \quad \text{B.7}$$

This result is closely analogous with Parseval's relation for the auto-correlation and cross correlation.

## APPENDIX C.

### 68010 machine code routine for calculating the triple correlation by photon differencing.

#### 2-DIMENSIONAL TRIPLE CORRELATION.

##### Register usage.

- \* D0 Base address of triple correlation array.
- \* D1 Start address of time frame.
- \* D2 End address of time frame.
- \* D3 First photon event x1.
- \* D4 Second photon event y1.
- \* D5 Second photon event x2, vect(x1-x2), 3rd photon event x3, vect(x1-x3).
- \* D6 Second photon event y2, vect(y1-y2), 3rd photon event y3, vect(y1,-y3).
- \* address offset due to vect(x1,y1) -vect(x3,y3)
- \* D7 Unused.
- \*
- \* A0 Start address of triple correlation address look up table.
- \* A1 Loop1 counter.
- \* A2 Loop2 counter.
- \* A3 Loop3 counter.
- \* A4 Address pointed to by vect(x1,y1)-vect(x2,y2).
- \* A5 Global base.
- \* A6 Stack frame pointer.
- \* A7 Stack pointer.
- \*
- \* Program Flow
- \* SETUP
- \* a) Get start address of triple correlation array.
- \* b) Get address of (0,0,0,0) in triple correlation address table.
- \* c) Get start address of current time frame.
- \* d) Get end address of current time frame.
- \* e) Set up loop 1.
- \* LOOP 1 \*\*\*\*\*
- \* f) Get coordinates of first photon event.
- \* g) Set up loop 2.

\* LOOP 2 \*\*\*\*\*

- \* h) if not(loop1 = loop2) jump to statement j
- \* i) increment loop2, jump to end of this loop.
- \* j) Get coordinates of second photon event.
- \* k) Find difference  $x1 - x2$ , and store in  $x2$ .
- \* l) Check  $(x1-x2)$  against bounds, if outside jump to end of loop.
- \* m) Repeat operations k and l with  $y1$  and  $y2$ .
- \* n) Add most significant offset due to  $x1-x2$  to A4.
- \* o) Add less significant offset due to  $y1-y2$  to A4.
- \* p) Set up LOOP 3.

\* LOOP 3 \*\*\*\*\*

- \* q) If not(loop1 = loop3) jump to u.
- \* r) Increment loop3 and jump to end of this loop.
- \* s) If not(loop2 = loop3) jump to u.
- \* t) Increment loop3 and jump to end of this loop.
- \* u) Get coordinates of third photon event.
- \* v) Repeat statements k,l,m with  $x1,x3,y1,y3$ .
- \* w) Change  $(y1-y3)$  into address offset (multiply by four).
- \* x) Use address table to add address increment due to  $(x1-x3)$  to A4.
- \* y) Increment memory location in triple correlation histogram.
- \* z) Test for end of loop 3 condition if false continue loop 3.
- \* aa) Test for end of loop 2 condition if false continue loop 2.
- \* ab) Test for end of loop 1 condition if false continue loop 1.
- \* ac) Return to calling routine.

\*\*

\* The following is HP Pascal specific linker information.

\*

```
MNAME  TA2D
SRC    MODULE TA2D
SRC
SRC    EXPORT
SRC
SRC    CONST    LEN=16;
SRC    CONST    LOW_BOUND=-8;
SRC    CONST    HIGH_BOUND=+7;
SRC    CONST    MAX_PHOT_PER_FRAME=64;
SRC
SRC    TYPE      SHORT_INT=-32768..32767;
SRC
SRC    TYPE      TC2D_TYPE=ARRAY[LOW_BOUND..HIGH_BOUND,
SRC              LOW_BOUND..HIGH_BOUND,
```

```

SRC          LOW_BOUND..HIGH_BOUND,
SRC          LOW_BOUND..HIGH_BOUND]
SRC          OF INTEGER;
SRC TC2D_POINT=^TC2D_TYPE;
SRC
SRC TYPE     ADRS4D_TYPE=ARRAY[1..3*LEN] OF INTEGER;
SRC ADRS4D_POINT=^ADRS4D_TYPE;
SRC
SRC TYPE     FRAME_TYPE=ARRAY[1..128] OF SHORT_INT;
SRC
SRC PROCEDURE T_C_2D(TC2D:TC2D_POINT;
SRC          ADRS4D:ADRS4D_POINT;
SRC          FRAME:FRAME_TYPE;
SRC          FRAME_LEN:INTEGER);
SRC
SRC END;
COM TA2D,-O
DEF TA2D_TA2D
DEF TA2D_TC_2D
*
*   Code Starts Here
*
LEN      EQU 16
LOW_BOUND EQU -8
HIGH_BOUND EQU 7
TABLE_MIDDLE EQU 96  4*(LEN-LOW_BOUND)
ZERO_OFFSET EQU 32  4*(-LOW_BOUND)
DIM_OFSTEQU 64  4*LEN
*
TA2D_TC_2D  TRAP      #1          (Carry out stack check)
           DC.W      0          (No local space needed)
           MOVE.L    20(SP),D0    (Get start address of trip cor)
           ADDI.L    #ZERO_OFST,D0 (Point to zero element of 1st dimension)
           MOVEA.L   16(SP),A0    (Get start of address table)
           ADDA.L    #TABLE_MIDDLE,A0 (Calc middle of Address table)
           MOVE.L    12(SP),D1    (Get start of time frame)
           MOVE.L    8(SP),D2     (Find last element in time frame)
           ADD.L     D2,D2        (Find last element in time frame)
           ADD.L     D2,D2        (Find last element in time frame)
           ADD.L     D1,D2        (Find last element in time frame)

```

```

        MOVEA.L D1,A1                (Setup loop 1)
*****
LOOP1   MOVE.W  (A1)+,D3              (Get photon event coordinate x1)
        MOVE.W  (A1)+,D4              (Get photon event coordinate y1)

        MOVEA.L D1,A2                (Setup loop 2)
*****
LOOP2   MOVE.W  (A2)+,D5              (Get x2)
        MOVE.W  (A2)+,D6              (Get y2)
        CMPA.L  A1,A2                (Photon Bias correction)
        BEQ  ENDLP2                  (If Phot1 = Phot2 jump to end of loop)

        SUB.W   D3,D5                (Find x1-x2)
        CMPI.W  #LOW_BOUND,D5
        BLT  ENDLP2                  (Lower range check)
        CMPI.W  #HIGH_BOUND,D5
        BGT  ENDLP2                  (Higher range check)

        SUB.W   D4,D6                (Find y1-y2)
        CMPI.W  #LOW_BOUND,D6
        BLT  ENDLP2                  (Lower range check)
        CMPI.W  #HIGH_BOUND,D6
        BGT  ENDLP2                  (Higher range check)

        MOVEA.L D0,A4                (Get start if TC array)
        ADD.W   D5,D5
        ADD.W   D5,D5
        ADDA.L  DIM_OFST(A0,D5.W),A4 (Add address increment due to x1-x1)
        ADD.W   D6,D6
        ADD.W   D6,D6
        ADDA.L  $0(A0,D6.W),A4 (Add address increment due to y1-y2)

        MOVEA.L D1,A3                (Setup loop 3)
*****
LOOP3   MOVE.W  (A3)+,D5              (Get x3)
        MOVE.W  (A3)+,D6              (Get y3)
        CMPA.L  A3,A1                (Check for photon bias term)
        BEQ  ENDLP3
        CMPA.L  A3,A2                (Check for photon bias term)
        BEQ  ENDLP3

```

```
SUB.W    D3,D5          (Calculate x1-x3)
CMPI.W   #LOW_BOUND,D5
BLT ENDLP3              (Lower range check)
CMPI.W   #HIGH_BOUND,D5
BGT ENDLP3              (Higher range check)

SUB.W    D4,D6          (Calculate y1-y3)
CMPI.W   #LOW_BOUND,D6
BLT ENDLP3              (Lower range check)
CMPI.W   #HIGH_BOUND,D6
BGT ENDLP3              (Higher range check)

ADD.W    D6,D6          (Change from integer to address offset)
ADD.W    D6,D6
EXT.L    D6              (Addresses are long words)
ADD.W    D5,D5
ADD.W    D5,D5
ADD.L    -DIM_OFST(A0,D5.W),D6
(Add address offset from table due to x1 -x3 to offset from y1-y3)
ADDQ.L   #1,$0(A4,D6.L) (Increment TC histogram)

ENDLP3   CMP.L    A3,D2
        BGT LOOP3
ENDLP2   CMP.L    A2,D2
        BGT LOOP2
ENDLP1   CMP.L    A1,D2
        BGT LOOP1
        UNLK     A6
        RTS

TA2D_TA2D  RTS          (Pascal Initialisation)
        END
```

## References

- 1 H. Bartelt, B. Wirtitzer, "Shift-invariant imaging of photon-limited data using bispectral analysis" *Opt. Commun.*, **53**, 13-16 (1985).
- 2 B. Wirtitzer, "Bispectral analysis at low light levels and astronomical speckle masking", *J.Opt.Soc.Am.*, A **2**, 14 (1985).
- 3 K.H. Hofmann, G. Weigelt, "Speckle masking observation of the central object in the giant H 2 region NGC 3603", *Astron. Astrophys.*, **167**, L15-L16 (1986).
- 4 A. Labeyrie, "Attainment of diffraction-limited resolution in large telescopes by Fourier analysing speckle patterns in star images", *Astron. Astrophys.*, **6**, 85-87 (1970).
- 5 J.E. Baldwin, C.A. Haniff, C.D. Mackay, P.J. Warner, "Closure phase in high-resolution optical imaging", *Nature*, **320**, 595-597 (1987).
- 6 J.W. Goodman, "Statistical Optics", Wiley, (1985).
- 7 G. Bedard, "Dead-time corrections to the statistical distribution of photoelectrons", *Proc. Phys. Soc.*, **90**, 131-141 (1967).
- 8 R.F. Chang, V. Korenman, C.O. Alley, "Correlations in light from a laser at threshold", *Phys. Rev.*, **178**, 612-621 (1969).
- 9 M.T. Koshiha, "Observational neutrino astrophysics", *Physics Today*, **40**, 38-42 (1987).
- 10 A. Boksenberg, in "Auxiliary instrumentation for large telescopes", ESO Conference Geneva.

- 11 A. Blazit, "Comptage de photons bidimensionnel et applications astrophysiques", University of Nice Thesis, (1987).
- 12 O.H.W. Siegmund, M. Lampton, J. Bixler S. Chakrabarti, J. Vallerger, S. Bowyer, R.F. Malina, "Wedge and strip image readout systems for photon counting detectors in space astronomy", O.S.A Topical meeting on quantum limited imaging and image processing, Hawaii (1986).
- 13 T. Gonsiorowski, "Variable threshold discrimination in a photon-imaging detector", Applied Optics, **23**, 1060-1065 (1984).
- 14 J.S. Morgan, "Speckle imaging with the MAMA detector", ESO Conference and Workshop Proceedings (High-Resolution imaging by interferometry), 381-391 (1988).
- 15 J.G. Timothy, J.S. Morgan, "Imaging by time-tagging photons with the Multi-anode Microchannel Array detector system", Proc. SPIE, **627**, 654-659 (1986).
- 16 A. Blazit, L. Koechlin, J.L. Oneto, "Image processing techniques in astronomy", (D. Reidel, Holland, 1976).
- 17 H. Bartelt, A.W. Lohmann, B. Wirnitzer, "Phase and amplitude recovery from bispectra", Applied Optics, **23**, 3121-3129 (1984).
- 18 G.R. Ayers, M.J. Northcott, J.C. Dainty, "Knox-Thompson and triple correlation imaging through atmospheric turbulence", J.Opt.Soc.Am., A **5**, 964-985 (1988).
- 19 G.R. Ayers, "Correlation techniques for imaging through atmospheric turbulence", University of London Thesis, (Sept 1988).

- 20 J.C. Dainty, A.H. Greenaway, "The signal to noise ratio in speckle interferometry", High angular resolution stellar interferometry, Proc. of Colloquium No. 50, International Astronomical Union. (1978).
- 21 A.W. Lohmann, G. Weigelt, B. Wirtzner, "Speckle masking in astronomy: triple correlation theory and applications", App. Opt., **22**, 4028-4037 (1983).
- 22 A.W. Lohmann, B. Wirtzner, "Triple correlations", Proc. IEEE., **72**, 889-901 (1984).
- 23 T. Matsuoka, T.J. Ulrych, "Phase estimation using the bispectrum", Proc. IEEE. **72**, 1403-1411 (1984).
- 24 J.W. Goodman, J.F. Belsher, "Fundamental limitations in linear invariant restoration of atmospherically degraded images", SPIE Proc., **75**, 141-154 (1976).
- 25 J.C. Dainty, A.H. Greenaway, "Estimation of spatial power spectra in speckle interferometry", J.Opt.Soc.Am., **69**, 786-790 (1978).
- 26 A.W. Lohmann, "Recovery of an object, moving at random in front of a fixed background", Optik, **73**, 174-180 (1986).
- 27 P. Horowitz, W. Hill, "The Art Of Electronics", Cambridge University Press, 437-450 (1988).
- 28 D.L. Fried, "Statistics of a geometric representation of wavefront distortion" J.Opt.Soc.Am., **55**, 1427-1435 (1965).
- 29 D.L. Fried, "Optical resolution through a randomly inhomogeneous medium for very long and very short exposures", J.Opt.Soc.Am., **56**, 1372-1379 (1966).

- 30 A.J. Ticknor, R.L. Easton, H.H. Barrett, "Two-dimensional Radon-Fourier transformer", *Opt. Eng.*, **24**, 82-85 (1985).
- 31 A.A. Canas, "The acousto-optic triple product processor and its applications", *Opt. & Quantum Electronics*, **19**, 79-82 (1987).
- 32 N. Ohyama, S. Ohki, S. Inoue, J. Tsujiuchi, T. Honda, "Discrete radon transform in a continuous space", *J.Opt.Soc.Am.*, A **4**, 318-324 (1987).
- 33 G. Weigelt, B. Wirmitzer, "Image reconstruction by the speckle-masking method", *Opt. Lett.*, **8** 389 (1983).
- 34 P. Nisenson, R. Stachnik, C. Papaliolios, P. Horowitz, "Data recording and processing for speckle image reconstruction", *Proc. SPIE*, **243**, 88-94 (1980).
- 35 J.C. Dainty, B.M. Levine, B.J. Brames, K.A. O'Donnell, "Measurements of the wavelength dependence and other properties of stellar scintillation at Mauna Kea Hawaii", *App. Opt.*, **21**, 1196-1200 (1982).
- 36 K.A. O'Donnell, J.C. Dainty, "Space-time analysis of photon-limited stellar speckle interferometry", *J.Opt.Soc.Am.*, A **70**, 1354-1360 (1980).
- 37 J.G. Walker, "Optimum exposure time and filter bandwidth in speckle interferometry", *IAU Colloquium, High angular resolution stellar speckle interferometry*, Maryland., (1978).
- 38 C.R. Lynds, S.P. Worden, J.W. Harvey, "Digital image reconstruction applied to Alpha Orionis", *Astrophys.*, **207**, 174-180 (1976).
- 39 R.H.T Bates, F.M. Cady, "Towards true imaging by wideband speckle interferometry", *Opt. Commun.*, **32**, 365-369 (1980).
- 40 J.G. Walker, "The phase retrieval problem: A solution based on zero location by exponential apodisation", *Optica Acta.*, **28**, 735-738 (1981).

- 41 J.G. Walker, "Computer simulations of a method for object reconstruction from stellar speckle interferometry data", *Appl. Opt.*, **21**, 3132-3137 (1982).
- 42 B.L. McGlamery, "Image restoration techniques applied to astronomical photography", *Astronomical Use Of Television-Type Image Sensors*, NASA SP-256(NTIS N71-28522), 167-192 (1970).
- 43 L. Mertz, "Phase estimation with few photons", *App. Opt.*, **23**, 1638-1640 (1984).
- 44 M.S. Scivier, T.J. Hall, M.A. Fiddy, "Phase unwrapping using the complex zeros of a bandlimited function and the presence of ambiguities in two dimensions", *Opt. Acta.*, **31**, 619-623 (1984).
- 45 G.J.M. Aitken, R. Houtman, R. Johnson, J.M. Pochet, "Direct phase gradient measurement for speckle image reconstruction", *Appl. Opt.*, **24**, 2926-2930 (1986).
- 46 K.T. Knox, B.J. Thompson, "Recovery of images from atmospherically degraded short-exposure images", *Astrophys.*, **193**, 45-48 (1974).
- 47 R.C. Jennison, "A phase sensitive interferometer technique for the measurement of the Fourier transforms of spatial brightness distributions of small angular extent", *Mon. Not. R. Astr.*, **118**, 276-284 (1958).
- 48 C.A. Haniff, C.D. Mackay, D.J. Titterton, D. Sivia, J.E. Baldwin, P.J. Warner, "The first images from optical aperture synthesis", *Nature*, **328**, 694-695 (1987).
- 49 A.H. Greenaway, "Terrestrial Optical Aperture Synthesis Technique (TOAST)", *Opt. Comm.*, **58**, 149-154.
- 50 L. Mertz, "Transformations In Optics", Wiley, New York, (1965).

- 51 C. Roddier, F. Roddier, in "Image formation from coherence functions in astronomy", D. Reidel, Amsterdam, 175-185 (1979).
- 52 P. Nisenson, M. Karovska, L. Koechlin, C. Papaliolios, C. Standley, "High Angular resolution observations of SN1987A", ESO Conference and Workshop Proceedings (High-Resolution imaging by interferometry), 381-391 (1988).
- 53 G.R. Ayers, M.J. Northcott, J.C. Dainty, "Knox-Thompson and triple correlation imaging through atmospheric turbulence" J.Opt.Soc.Am., 5, 986-992 (1988).