PhD University of Edinburgh 1992

Phase-Only Optical Information Processing

D. J. Potter

Chapter 3
The Weak Phase Approximation

Chapter one introduced the Zernike phase filtering operations which result in an image intensity proportional to the object phase. Two equations commonly appear in the explanation of the filter operation - expansion of the exponential $e^{if(x)}$, equation 2.2 and the expression for the resulting spectrum 1+iF(ν), equation 2.3. These equations are deemed valid when the object f(x) has a relatively small phase retardance, and together summarise what is commonly referred to as the weak phase approximation.

The first section of this chapter suggests a possible redefinition of what is meant by the `weak' phase approximation in the light of the Bessel function analysis of chapter two. If the object phase retardance f(x) is represented as a Fourier series then, it shall be shown, a useful parameter in the analysis of the spectrum is the maximum size of any Fourier coefficient a_j, which in this thesis is termed the `Zernike Limit'. A refined calculation of this limit, encompassing a number of spectral and image quality parameters as useful aids, is presented in section two. Section three briefly investigates likely causes of poor image quality and obtains some unexpected results from the Bessel function program described in Chapter two.

1 Classical Explanation

The classical phase contrast filtering operations introduced by Zernike were explained at the beginning of chapter two with the assistance of a Taylor series expansion of the object so that

$$e^{i\,f(x)} \cong 1 + if(x) \qquad (1)$$

The Fourier Transform G(ν) of the phase object was found to be

$$G(\nu) \cong \delta(\nu) + iF(\nu) \qquad (2)$$

where F(ν) is the Fourier Transform of f(x). This line of reasoning clearly leads to an explanation of the phase contrast imaging process, but it will now be shown that an approach from a different perspective can lead to a more complete understanding of the process.

As explained in the previous chapter, each Fourier coefficient produces a whole comb of spectral orders in the frequency plane which convolves with the combs from every other coefficient a_j, where

$$f(x) = \sum_{j=1}^{N} a_j \cos(jx + \Phi_j) \qquad (3)$$

With regard to producing a spectrum linear with F(ν), only the primary orders of each comb, having a complex amplitude of $J_{\pm1}(a_j)$, are required. Should the size of coefficients a_j be so large as to cause appreciable ghost spectral orders ($J_k(a_j) : k \geq 2$) to appear then the resulting convolutions with these orders effectively destroys this linearity. Therefore, if convolutions with ghost orders are assumed to be the principle cause of spectrum (and

consequently image) nonlinearity an image intensity linear in f(x) should occur after a Zernike phase contrast operation if each and every a_j is less than a certain limit - the `Zernike' limit.

Two consequences of this conjecture are immediately apparent:

1. f(x) may be as large as allowed by equation 3.3 and thus may not be *small*.
2. The upper limit to a_j is independent of the number of terms in the series describing f(x).

Care must be taken with this conjecture - in particular convolutions occur between the primary orders of each comb and the zero orders of every other comb (or the resultant unit comb) as illustrated in figure 3.6. Also, even though the convolutions which do occur may alter the *amplitude spectrum* by only a small percentage, the phase will also be altered slightly. The complex amplitude at each frequency then has a phase which is roughly `i' times Φ_j, plus or minus a small error term due to the convolution with other Bessel combs. The introduction of this phase error, which is likely to be random in nature, is akin to viewing the spectrum through a (small retardance) random phase plate. It is well known that image formation can withstand relatively large errors in the amplitude spectrum a_j but is highly sensitive to the phase spectrum Φ_j. Therefore, any calculation to determine the Zernike limit to a_j based on amplitude errors alone is likely to result in an absolute maximum figure.

1.1 Linearity and Non-Weak Phase Objects

The image intensity I(x) after phase contrast filtering at the zero of frequency may no longer obey the relation

$$I(x) \cong 1 \pm 2f(x) \qquad (4)$$

due to the large possible size of f(x) allowed by equation 3.3 but this is not to say that the intensity distribution may no longer be linearly proportional to f(x). Recall the microscopists' view of image formation (chapter1, section 1.1.4). As the object phase retardance increases light which would have gone into the undiffracted wave to form the constant background field of the image may instead be channeled into the diffracted wave, and makes up the (larger) image field describing the form of the object f(x). This idea is illustrated in figure 3.1.

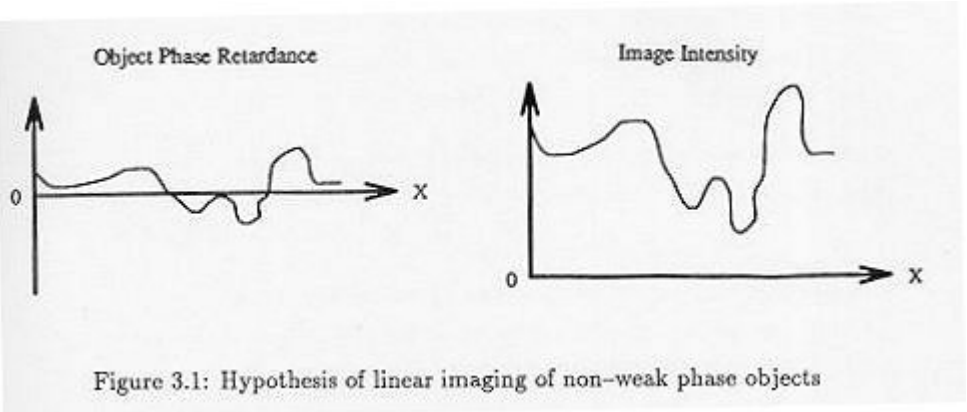

Figure 3.1: Hypothesis of linear imaging of non-weak phase objects

Figure 3.1: Hypothesis of linear imaging of non-weak phase objects

The principle of conservation of energy requires that $\int_{-\infty}^{+\infty} I(x)\,dx$=Constant for both images shown, and that in a loss free system[1] the constant is the same for both images.

The upper (Zernike) limit to the size of a_j may be determined from close inspection of the Bessel function graphs, or preferably by calculation. As the argument x of $J_k(x)$ increases, so the Bessel functions increase in accordance with their order, which is to say that $J_k(x)$ is always greater than $J_{k+1}(x)$, $k \geq 1$, until the functions start to decrease again at relatively large x. Therefore it is only necessary to find that value x at which $J_{+2}(x)$ becomes significant in order to determine the Zernike limit. Choosing a significant value of $J_{+2}(x)$ to be 0.03, it is found J_{+2}

(x) ≥ 0.03 when x=0.5. In order to be doubly sure of not stepping into the region where $J_{+2}(x)$ becomes significant it would be prudent to define the Zernike limit to be 0.4.[2]

To summarise,

1. **Classically, the weak phase approximation refers to a Taylor expansion of the function describing the complex transmission of the object plane when it contains a *weak* phase object of phase retardance f(x). Classically, this should not exceed [(λ)/10] to be called weak [7, page 490].**

2. **From a convolution viewpoint, it is perhaps more appropriate that the weak phase referred to is that of the Fourier coefficients describing f(x) and not f(x) itself.**

3. **If ghost order convolutions dominate the formation of spectral (and therefore image) non-linearities, this would imply linear imaging of the phase f(x) is possible if no Fourier coefficient of f(x) exceeds a finite limit, here termed the Zernike limit. At this stage, it is thought such a limit should not exceed the value of 0.4 radians.**

4. **As a consequence of (3) above, f(x) need no longer be small and may exceed the conventional value of [(λ)/10].**

5. **The above estimate is obtained from considerations of errors in the amplitude spectrum of f(x) and is therefore likely to be an absolute maximum.**

These conjectures will be tested against results obtained by computer simulation in section 3.3. Recalling the essence of the previous chapters, convolution is at heart the fundamental process at work in determining the properties of a phase object spectrum - the Taylor expansion result is completely contained within its framework but of course the opposite is not true. As such, conjectures based on investigation into convolution effects should be given more weight over the rather limited predictions of the Taylor expansion.

1.2 A Qualification

It shall now be argued that the *range* of Fourier coefficients a_j affects the accuracy with which the image intensity distribution approximates the object phase f(x).

This result can best be demonstrated by considering the convolution of the N=2 comb with the first unit comb. So as to reduce the amount of algebra to a minimum, an observation frequency γ = 1 is taken again for which equation 3.41 of Chapter 2 shows the complex amplitude after the first convolution stage to include the terms

$$\begin{aligned} \ldots \; + \quad & J_{+1}(a_1)\, J_0(a_2)\, i^{+1} \\ + \quad & J_{+3}(a_2)\, J_{-1}(a_2)\, i^{-2} + J_{-1}(a_1)\, J_{+1}(a_2)\, i^{0} \\ + \quad & J_{-3}(a_1)\, J_{+2}(a_2)\, i^{-1} + J_{-5}(a_1)\, J_{+3}(a_2)\, i^{-2} + \ldots \end{aligned} \qquad (5)$$

As mentioned in Chapter 2, the first of these terms contains the information on the spectrum of f(x) and the remaining terms are all non-linear functions of the ideal amplitude [$(a_1)/2$] and phase ϕ_1 (all ϕ_j set to zero in the above expression for compactness) at the primary order location γ = +1. It is pertinent to ask the question `How do these additional cross-product terms perturb the spectrum from the ideal *linear* term $J_{+1}(a_1)\, J_0(a_2)\, i^{+1}$? '

Considering the non-linear terms to be *noise* terms, the noise to signal ratio for this frequency may be defined as

$$\text{Noise} \quad \sum \sum^{\prime} J_{m_1}(a_1)\, J_{m_2}(a_2)$$

$$\overline{\text{Signal}} = \frac{m_1 \quad m_2}{J_{+1}(a_1)\, J_0(a_2)} \tag{6}$$

for any set values of a_1 and a_2. If the phases of each Fourier coefficient are taken into account, as well as the $i^{m_1+m_2}$ terms, then the noise function is actually a phasor addition. However, increasing a_j alone serves to increase the modulus of each component phasor so the modulus of the final complex number is also increased. For the purpose of illustration only, it is deemed acceptable to define the noise to signal ratio as above. The gradient $[\,d/dx]J_n(a_j)$ is always greater than zero for $n \geq 1$ in the range of a_j concerned with here, and as δa_j is positive the numerator of the noise to signal ratio must increase as the range of a_j increases. At the same time, the denominator must decrease as $[\,d/dx]J_0(a_j)$ is negative. Thus the noise to signal ratio must always increase if the range of coefficients increases.

The ideal spectrum of this simple phase object would only have spectral orders at primary order locations $\gamma = \pm 1$, $\gamma = \pm 2$ as well as the DC order. Having considered the noise to signal ratio at $\gamma = +1$, what of the N=2 primary order location $\gamma = 2$? As a_j increases not only do the ghost spectral orders increase but the linearity of $J_1(a_j)$ with a_j decreases, and it is primarily this effect which will decrease the denominator of the noise to signal ratio at $\gamma = 3$. If the noise to signal ratio of this unit comb has been increased, it would be surprising if a further convolution with this comb (to find the spectrum of a three frequency phase object) would result in a decreased noise to signal ratio of the next unit comb as a whole.[3] Further, the above calculation could be performed for the $\gamma = 3$ frequency position to show that increasing a_3 increases the noise to signal ratio there and so on for any number of convolutions. This argument suggests that to obtain an image of intensity linearly proportional to the object phase, not only must a_j lie below the Zernike limit but (from consideration of equation 3.6) the spectral distribution of a_j should be as flat as possible.

2 Numerical Simulation

A FORTRAN computer program was written to test the conjectures of this chapter. Two types of Fourier coefficient sets were used in the simulations, one having a_j and phase Φ_j selected at random and the other with a_j set to follow a Gaussian distribution with Φ_j again being randomly selected. The maximum value of the Fourier coefficients was incremented in steps of 0.02 and, in order to obtain statistically valid results, 50 random data sets (and therefore phase objects), were made for each set of limits to a_j.

For the first run of the program (using random a_j) twelve coefficients were chosen to make up each phase object. The minimum limit to the size of a_j was set to be 0.02, and the maximum allowed to vary in steps of 0.02 up until the Zernike limit of 0.4 was reached. The second run used a Gaussian distribution of a_j and the number of Fourier coefficients `N' selected so that 95% of the energy of the Gaussian was contained within $-N < j < +N$. The coefficients were defined as

$$a_j = a_{max} \times e^{-[\,1/2][(a_j)/(\sigma)])^2} \tag{7}$$

and σ was chosen to be 16, requiring 28 Fourier coefficients a_j to fulfil the 95% energy condition. Typical computed Gaussian amplitude spectra (where deviations from the smooth shape show up best) for three values of a_{max} are shown in figure 3.3.

Upon Fourier Transformation, the amplitude and phase of each spectrum was recorded and compared with that of the ideal spectrum of f(x). As the image intensity function may resemble f(x) in form but be scaled and DC offset, it is necessary upon re-transformation to find a scaling parameter which resulted in the best fitting of object phase f(x) to the shape of the image intensity. Appendix six provides a brief summary of the scaling

procedure and some helpful information on the fine details of the simulation such as selecting object wavelengths to eliminate numerical phase errors in the FFT routine.

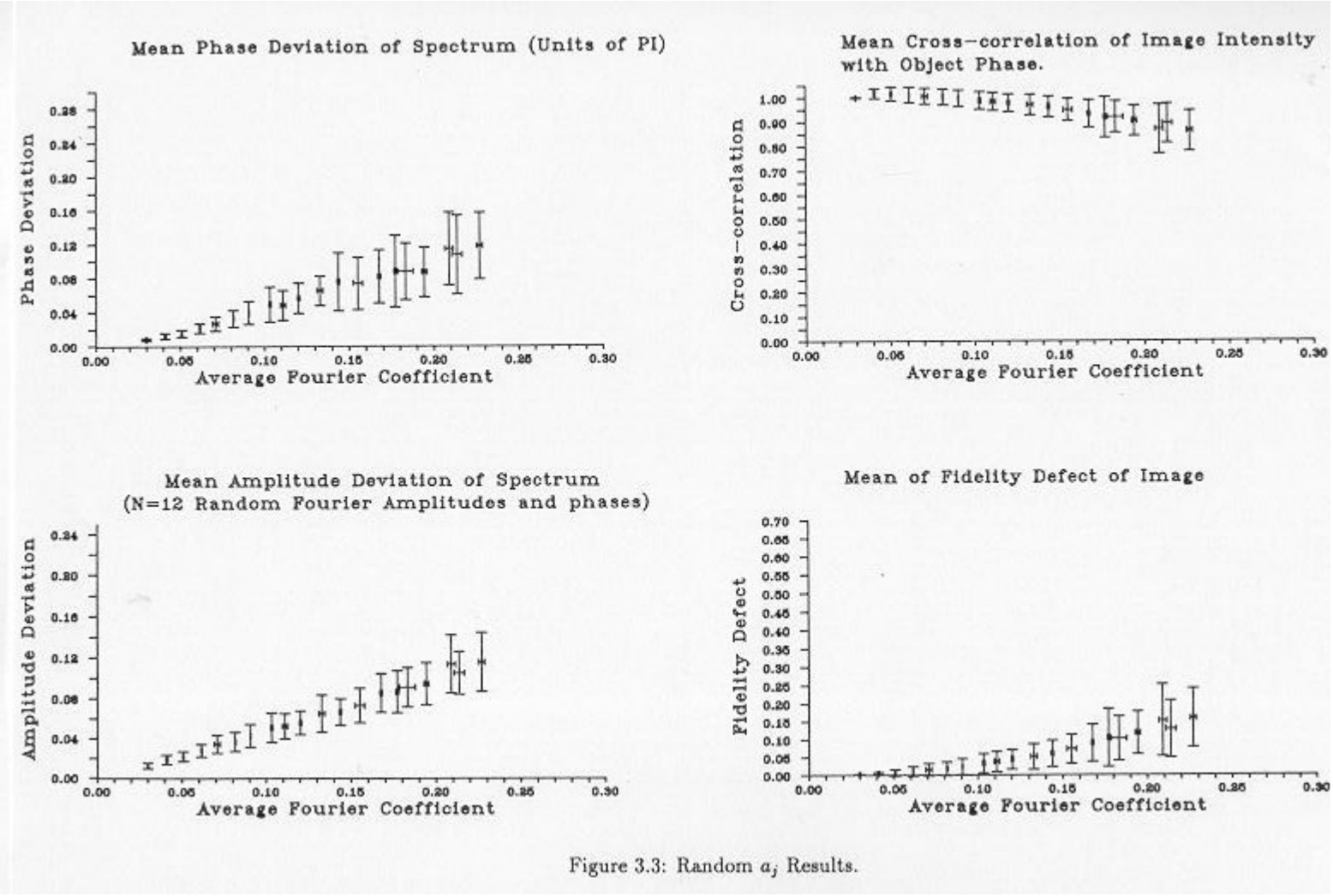

Figure 3.3: Random a_j Results.

Figure 3.2: Typical Amplitude Spectra: Gaussian a_j (Zero Order Suppressed)

3.1 Spectrum Quality

For each spectrum in the simulation, the mean spectrum amplitude and phase deviations, D_{amp} and D_{Φ}, were calculated. If F_j represents the amplitude and θ_j the phase of the measured spectrum (in units of π) then

$$D_{amp} = \frac{1}{N \; \bar{a}_j} \sum_{j=1}^{N} \left| \left(\frac{a_j}{2}\right) - F_j \right| \qquad (8)$$

$$D_{\Phi} = \frac{1}{N} \sum_{j=1}^{N} |\theta_j - \Phi_j - 0.5| \qquad (9)$$

where $[\bar{a}_j)]$ is the mean of the amplitudes a_j and N is the total number of coefficients. Equation 3.2 shows the object spectrum in frequency space is multiplied by `i', so that for a direct comparison between object phase and numerically recovered spectral phase data it is first necessary to subtract [(π)/2] radians from Φ_j. The $[\bar{a}_j)]$ term in the denominator of D_{amp} is merely a weighting, the relative errors of each frequency being weighted by the

importance of that frequency compared to the mean of all amplitudes a_j, so that each term in the sum is really

$$\frac{\left|\left(\frac{a_j}{2}\right) - F_j\right|}{\left(\frac{a_j}{2}\right)} \times \frac{\left(\frac{a_j}{2}\right)}{-a_j} \tag{10}$$

The merit of such a weighting may be argued but as the simulations are primarily intended to check on image linearity, this is not a great point of contention.

The mean and standard deviation of D_{amp} and D_{Φ} (which are themselves mean values) are calculated for the 50 data sets found for each upper limit to a_j. In order to obtain a parameter characterising the *range* of a_j, the Fourier coefficients are sorted into size order and a least squares algorithm computes the best gradient m so that

$$b_I = m \times I \tag{11}$$

where the b_I are the size ordered a_j. This is a slightly more accurate method of estimating the range of a_j than taking the parameter $[(a_{max}-a_{min})/N]$, especially so in the case where a_j obey a Gaussian distribution.

3.2 Ultrasonic Analogy

Strong similarities exist between the spectra of phase objects and the spectra recorded from Doppler velocimetry measurements using ultrasound. As such, simulations of ultrasonic spectra [26], [27] have been analysed in an almost identical manner to the analysis of phase object spectra in this chapter and this area is briefly reviewed here.

In Doppler velocimetry, a highly directional, very narrow bandwidth sound pulse is reflected off various surfaces (scatterers) within the medium under study and the time taken to receive the reflection is taken as proportional to the distance of the reflector from the source. If the reflector is moving then a frequency shift will occur which can be detected and used to obtain information about the velocity profile within the medium. The analogy with phase object spectra comes about due to the interaction of the reflected signal from every scatterer with the signal from every other scatterer. Due to the velocity distribution within the medium (typically blood), the sound waves reflected from each scatterer have a slightly different frequency. Further, a reflected wave may scatter again and again off many scattering centres before reaching the detector.

The signal received at the detector is the Fourier integral of the signals from these reflected waves. Suppose n scattering particles exist within the medium. The *ideal* Doppler spectrum would then consist of n single peaks separated from the source frequency by known amounts $\delta\nu_i$ according to the formula

$$\frac{\delta\nu_i}{\nu_s} = \frac{2\,v_i}{c} \tag{12}$$

where ν_s and c are the source frequency and speed of ultrasound in the medium.[4] The reflected waves from each scatterer interfere and, due to the particle velocity distribution, result in a measured Doppler time signal [26] of

$$y(t) = \sum_{j=1}^{M} a_j \cos(2\pi f_j t + \Phi_j) \tag{13}$$

where the a_m random Rayleigh variables[5] and Φ_m are uniformly distributed on the range $-\pi,+\pi$. M is the number of bins into which the frequency range of the spectrum has been divided - essentially the spectrum resolution. In the case of a phase object every sine wave component describing the phase retardance creates not just one but a

whole spectrum of frequencies in the Fourier plane. Further, each of these frequencies interacts with the frequency combs produced by every other sinusoid of the object. In ultrasonics the word ascribed to these unwanted frequencies is *speckle* and its suppression is desirable in that only the underlying first order spectrum (from the initially scattered waves) contains useful information. This is analogous to the effect of the ghost orders on the image of a phase object studied in this chapter.

In particular, the simulations studied here compare the spectrum with that of the ideal spectrum which is known exactly. In Doppler ultrasonics the time average Fourier Transform of equation 3.13 is used as the goal for which all speckle reduction techniques should aim to reach, and equations identical in form to equation 3.8 of this chapter seek to quantify the variation of the spectrum from the ideal; the only difference being that in the ultrasonic case the a_j and Φ_j naturally vary with time whereas in this chapter they are artificially varied to simulate different phase objects.

3.3 Image Quality

Before any quantitative observations are made it is also necessary to somehow quantify the quality of the resulting image. Many possible measures of image quality exist but only two have been selected for this simulation - the cross-correlation of object phase with (scaled) image intensity, and the *fidelity defect* [7, pages 171-175] of the image as compared to the object. Cross-correlation is a familiar tool used to quantify differences between two functions, though the fidelity defect may not be. Briefly, if g(x) represents the image intensity of an object with amplitude transmittance f(x), the fidelity of the image is defined as

$$\frac{\int |f(x) - g(x)|^2 \, dx}{\int |f(x)|^2 \, dx} \qquad (14)$$

and as such is a measure of the mean difference between both functions. In the context of this chapter, f(x) is not the object transmittance but the object phase which it is hoped to produce an image of after an intermediate spatial filtering process. In order to gain a single figure for each image quality parameter, only the central ordinate of the cross-correlation was calculated, which reduces to

$$\frac{\int f(x)\, g(x) \, dx}{\int f(x)^2 \, dx} \qquad (15)$$

Figure 3.3 shows both the mean amplitude and phase deviations together with the cross-correlation and fidelity defect results obtained from the first run of the program, and figure 3.4 for the second run. Graphs measuring the same parameter are drawn to the same scale for both runs for the program so that visual comparison may readily be made.

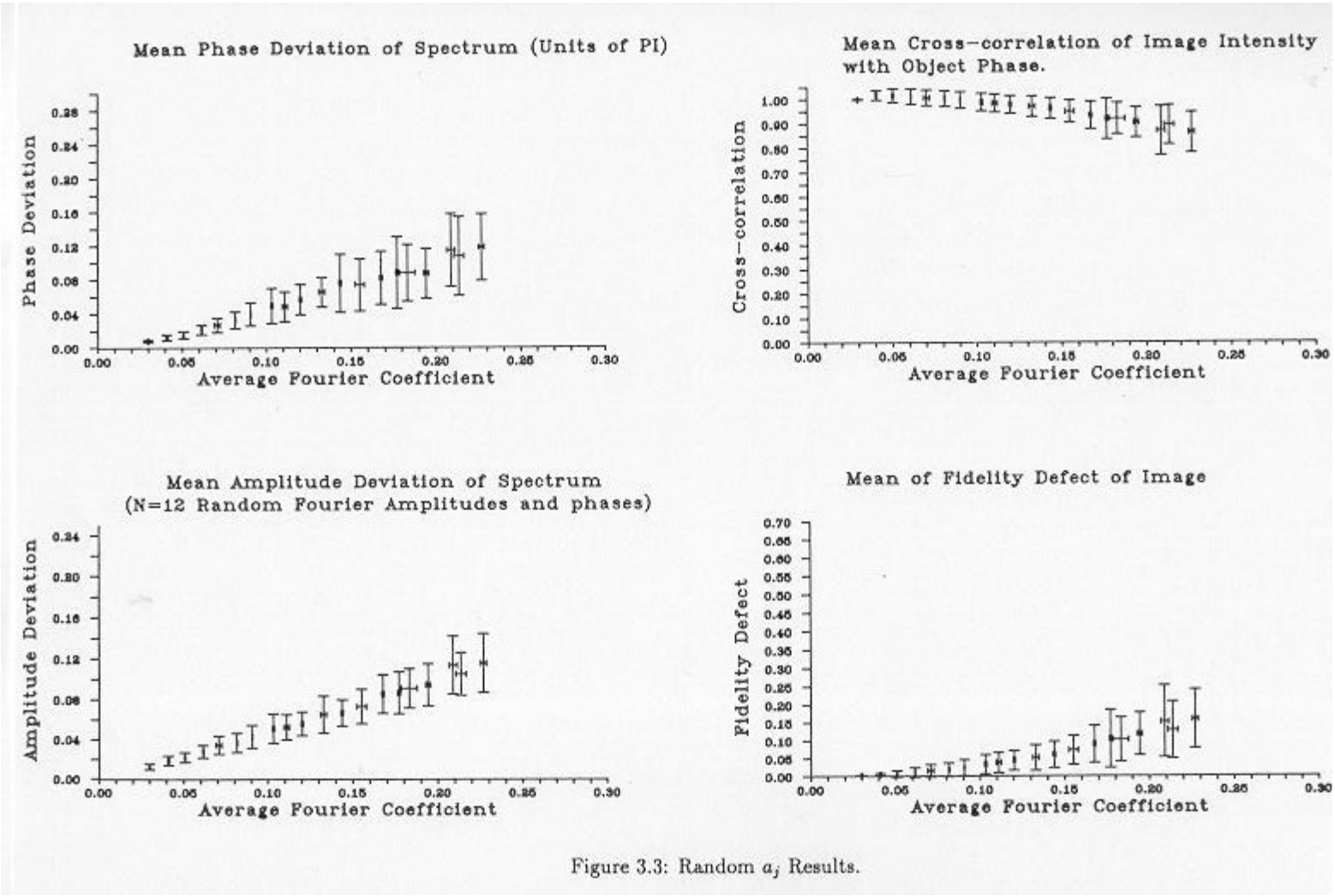

Figure 3.3: Random a_j Results.

Figure 3.3: Random a_j Results.

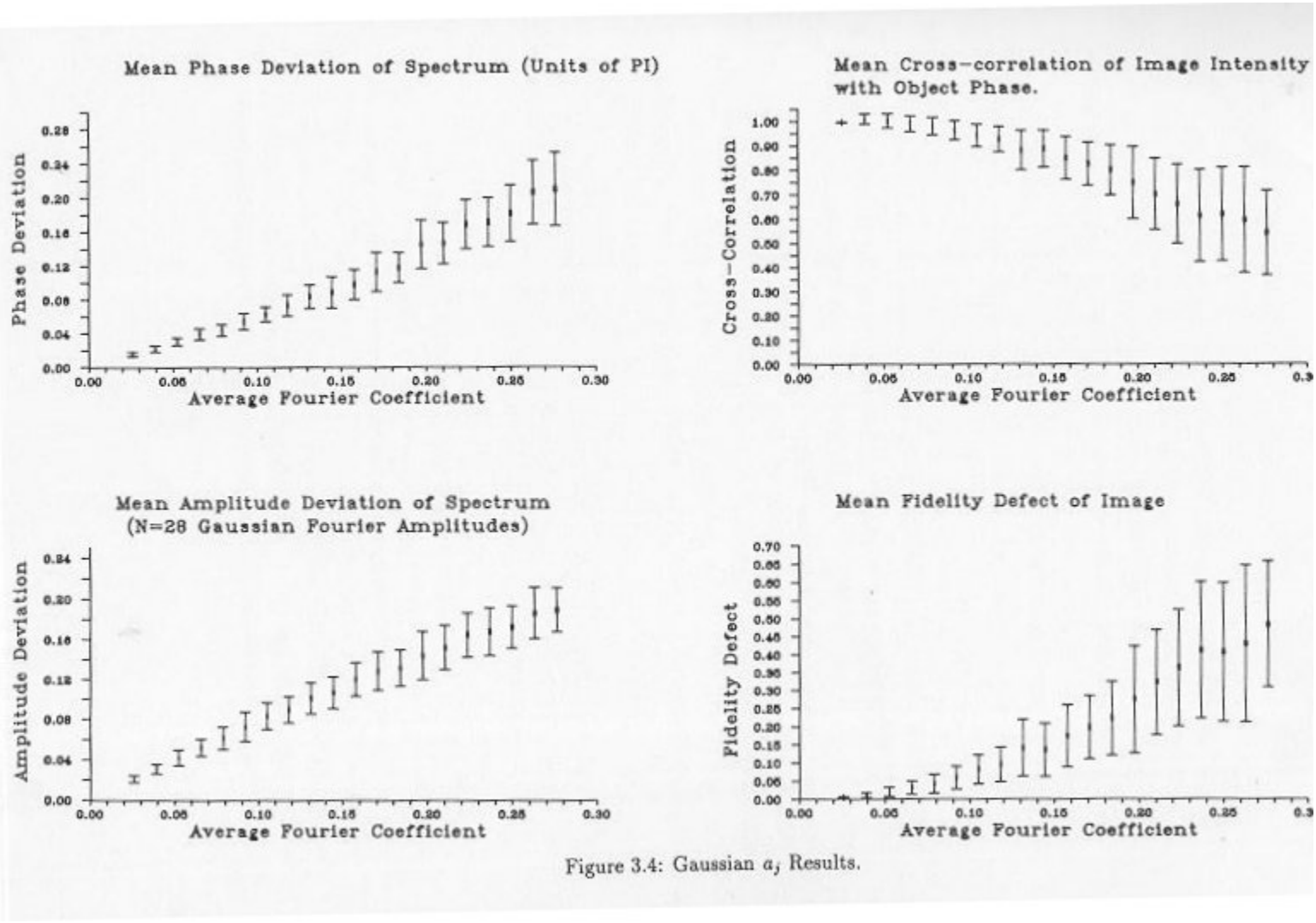

Figure 3.4: Gaussian a_j Results.

Figure 3.4: Gaussian a_j Results.

3.4 Observations

Firstly it is clear that both spectral integrity and image linearity deteriorate as the mean amplitude of Fourier coefficient increases. This implies one of two things - either the ghost orders are still making a significant contribution to the spectrum or convolutions with the primary orders cause more harm than previously imagined. In either case one must conclude that detrimental convolution effects are still occuring. The conjecture that the spectrum deviation increases as the range of Fourier coefficients increases is verified in this simulation even at extremely small values of a_{max}.

Tables 3.1 and 3.2 contain the results of three regions of interest from the first and second run of the program respectively. Examining the image quality results obtained from the first run (random a_j) observe that the mean cross-correlation between object phase and image intensity falls to ≅ 88±8% when a_{max}=0.4 (corresponding to $[\bar{(a_j)}] \cong 0.2$). It is the opinion of the author that this value lies in the lower range of what might be considered an acceptable figure for this parameter. Reducing the value of a_{max} to 0.2, it is observed, results in a much improved cross-correlation and fidelity defect which are better still when a_{max} falls to 0.1.

For a given a_{max}, the second run shows that objects with spectra obeying a Gaussian distribution of amplitude generally suffer more severe spectrum and image quality degradation than do objects with random spectral amplitude profiles. For the random spectral profile the higher frequencies will, on average, contain as much energy as the lower frequencies and one might expect the detrimental convolution effects to cause roughly equal amounts of spectral deviance from the ideal over the whole frequency range. However, by its very nature the

Gaussian amplitude distribution has less energy in the higher frequencies where convolution effects might cause relatively more harm. As it is the high spatial frequency components which are generally responsible for the image acuity, one might argue that the Gaussian spectral profile should be more sensitive to the convolution process than the random spectral profile. For the Gaussian profile, figure 3.4 shows that the increase of image fidelity defect is almost exponential in nature, and the fall off of cross-correlation as a_{max} increases is rapid indeed. As for the first run however, an upper limit of $a_{max} \cong 0.1$ is observed to produce excellent cross-correlation and fidelity defect results.

a_{max}	$[\bar{(a_j)}]$	Cross-correlation	Fidelity Defect
0.4	0.2	0.88±0.08	0.13±0.07
0.2	0.1	0.98±0.03	0.03±0.02
0.1	0.06	1.01±0.03	0.01±0.01

Table 1: Summary results for randomly selected a_j

a_{max}	$[\bar{(a_j)}]$	Cross-correlation	Fidelity Defect
0.4	0.26	59.0±21%	0.42±0.2
0.2	0.13	88.0±8%	0.14±0.08
0.1	0.07	98.9±3%	0.03±0.02

Table 2: Summary results for Gaussian a_j

Figure 3.5 shows the agreement between object phase and (scaled) image intensity for the 12 randomly selected Fourier coefficients at low values of a_j. The trace showing least modulation is in each case that showing the phase modulation and the trace with lesser modulation represents the resulting image intensity. Observe that even at very large a_j there remains close agreement between object phase structure and image intensity structure with the exception of regions of greatest phase retardance. However, these areas are of considerable interest. For small values of a_j, regions of greatest phase retardance in the object (the dips) are accurately represented in the structure of the image intensity I(x) i.e. [d/dx] I(x) > 0 when [d/dx](phase) > 0.

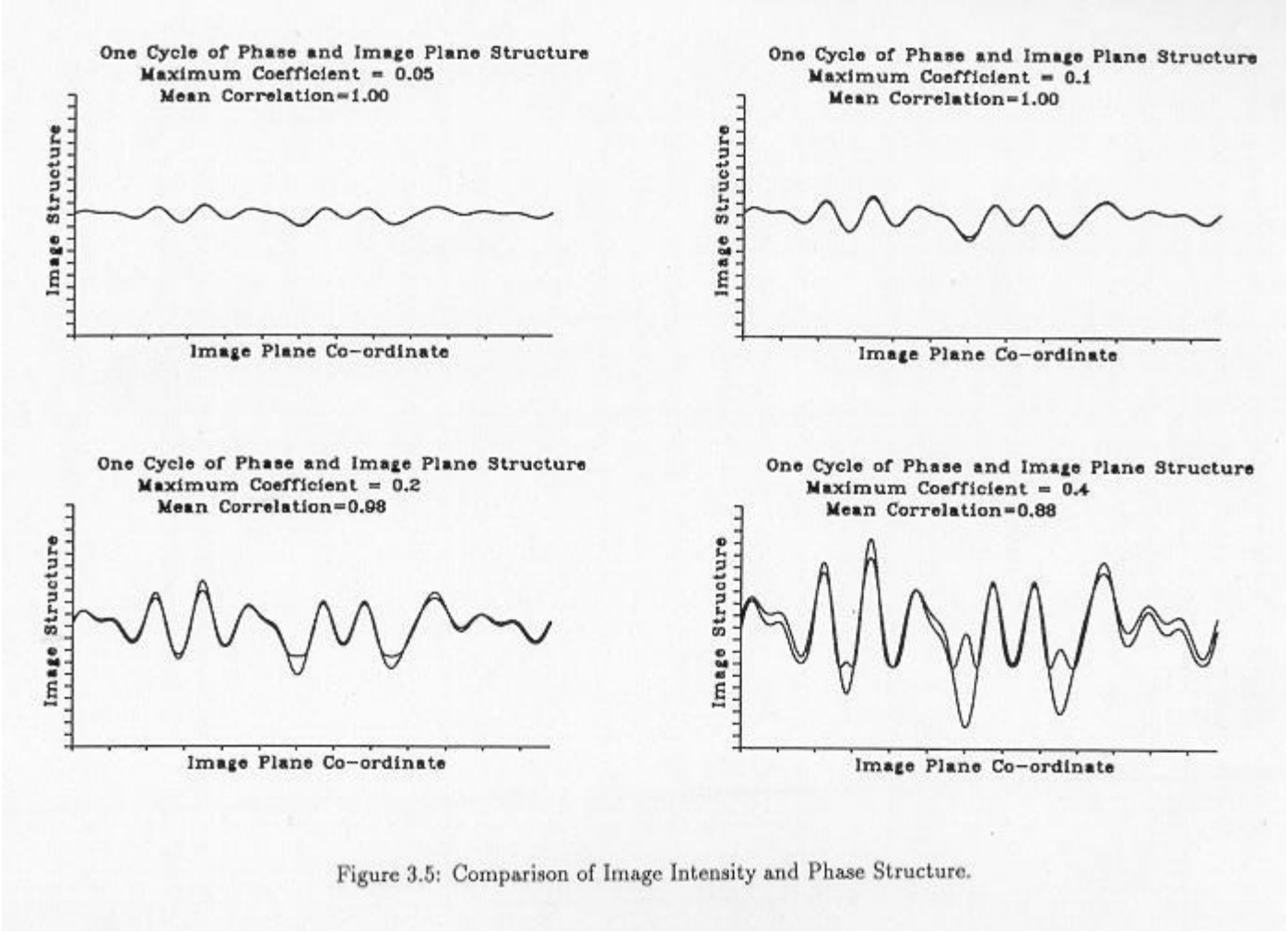

Figure 3.5: Comparison of Image Intensity and Phase Structure.

Figure 3.5: Comparison of Image Intensity and Phase Structure.

However, as the size of a_j increases, these areas are represented by dips which are not as pronounced as those of the phase structure until at large a_j there is actually a reversal of the dip so that $[d/dx]I(x) < 0$. This most definitely occurs when the maximum Fourier coefficient a_j has a value of 0.4, and it should be noted from figure 3.6 that the typical maximum depth of modulation of those phase objects with $a_j=0.4$ is close to $[(\lambda)/2]$ where classically the `weak' phase approximation is thought to break down.

The simulation results may be summarised as follows.

1. **Computer simulations using a_j where ghost orders are thought negligible have shown considerable spectral and image degradation still occurs as a_j increases.**

2. **Two very different models of phase objects suggest a Zernike limit of order 0.1. Typically the resulting object has a modulation depth of $[(\lambda)/10]$ - the classical upper limit for image linearity.**

3. **Figures 3.5 and 3.6 show, however, that with the exception of regions of very high phase retardance a high degree of similarity may be found between the phase retardance and image structure even at modulation depths as high as $[(\lambda)/3]$.**

Although this analysis has been conducted in one dimension, it is expected that studies in two dimensions would yield analogous results and that the conjectures of this chapter are generally applicable in 2-D also.

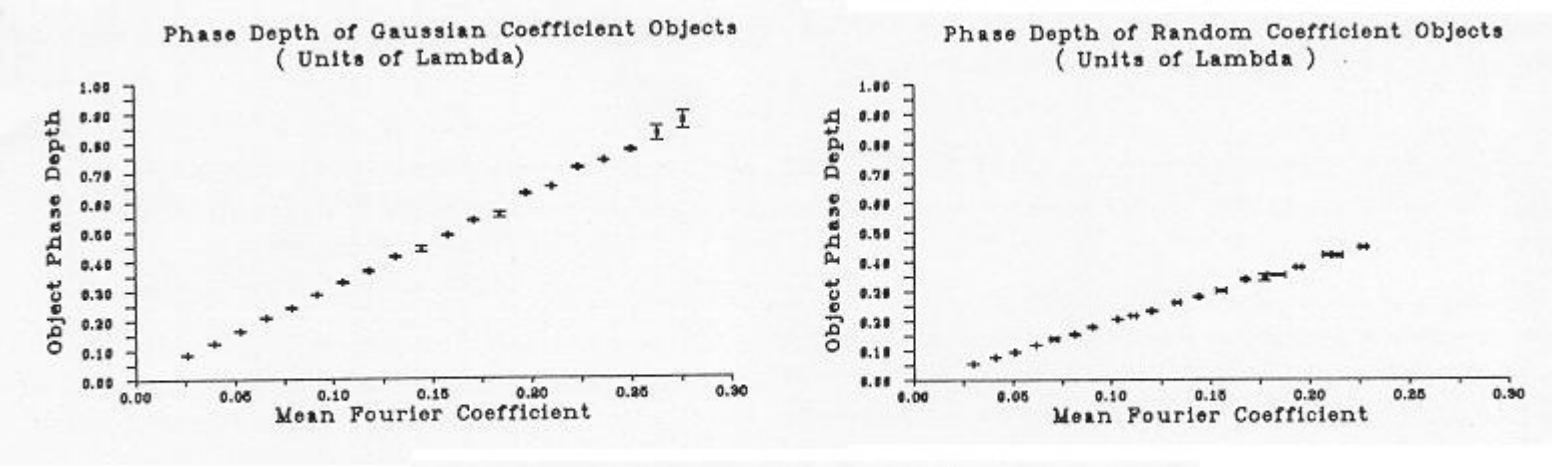

Figure 3.6: Depth of Phase Modulation: Random and Gaussian a_j

Figure 3.6: Depth of Phase Modulation: Random and Gaussian a_j

3 Reduced Zernike Limit

The essential results of this chapter have been stated in the last section. It remains, however, to explain the discrepancy between the expected Zernike Limit and the practical limit as deduced via simulations. In this final section therefore, several of the most likely explanations are put forward. The suggestions are tested by means of the Bessel function convolution program discussed at length in chapter two, which very quickly points to the principle cause of spectrum deterioration (and hence image quality also), even at very low values of the Fourier coefficients a_j.

With regard to spectrum deviations from the ideal, four likely explanations suggest themselves:

1. **Firstly, the value a_{max}=0.4 was chosen on the basis that the size of the ghost orders ($J_{|m|>1}$) was negligible in comparison with $J_1(a_j)$, therefore their effect on the spectrum would also be small.**
   This condition may not be wholly fulfilled.

2. **Secondly, convolutions involving the primary orders were deemed to be less harmful than convolutions with the more numerous ghost orders.**
   This assumption may be in error.

3. **Third, the slight non-linearity of $J_{+1}(a_j)$ for each a_j may result in more significant image degradation than previously thought, even though the spectrum deviation is small.**

4. **Finally, this value of a_{max} was postulated to be an upper limit which phase error effects on the image quality may require a lowering of.**

The FORTRAN program `TRUESPIKE' was written to perform the multiple convolutions of Bessel function combs as described in chapter 3. The unique usefulness of this program is in the ability of the user to artificially reduce Bessel orders higher than the prime orders for each comb ($J_m(a_j)$ where $|m| \geq 2$) to zero. Thus the effects of ghost orders on the final spectrum are instantly determinable. Also, any non-linear rendering of the a_j by the prime orders $J_{\pm1}(a_j)$ can be eliminated by setting the amplitude of these orders to be $[(a_j)/2]$ as in the ideal case. (The phase of the spectrum, it may be remembered, is not altered at the prime order of any comb except for a constant complex multiplier of `i'.) Finally, at each stage of the convolution the phase of any δ-function may be set to that of the ideal spectrum to eliminate phase errors.

The Effects of Ghost Orders

In the program, it was decided to use a Gaussian type spectrum for the simple reason that deviations from the ideal amplitude show up as deviations from the smooth Gaussian profile. The Fourier coefficients a_j and their associated phases Φ_j were selected by the same subroutine which produced the numerical simulations of section two, and thus the spectrums produced may be compared directly with those of figure 3.2 of section 3.

The spectrums calculated using two values of a_{max} are shown in figure 3.8a. The plots are identical whether produced by an FFT routine numerically or by the Bessel convolution program utilising for each frequency Bessel function orders up until the 6'th. Figure 3.8b shows the computed spectrum where all Bessel orders higher than the first have been set to zero.
It is observed that no appreciable difference between the spectrums can be detected for the values of a_{max} used. Therefore the ghost orders *do not* contribute to spectral deviations and the first hypothesis to explain the phenomenon is not the right one.

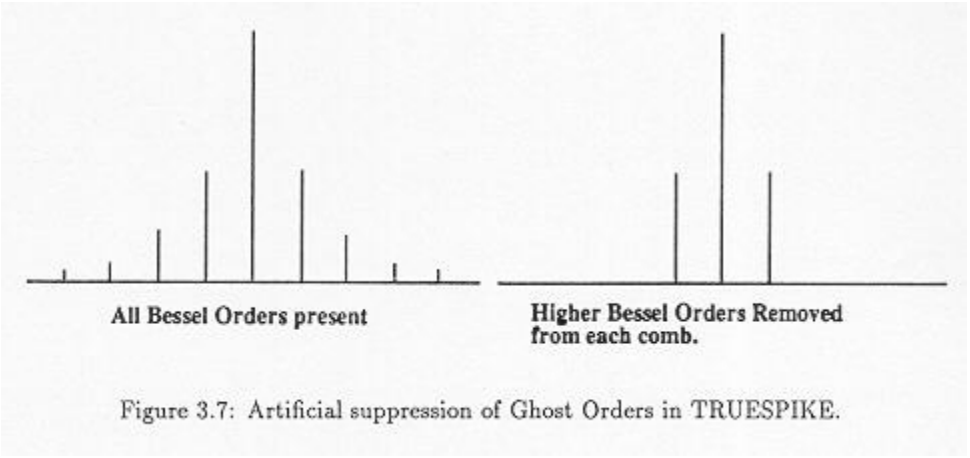

Figure 3.7: Artificial suppression of Ghost Orders in TRUESPIKE.

Figure 3.7: Artificial suppression of Ghost Orders in TRUESPIKE.

Though not plotted, calculations with a_{max} as high as the initial Zernike limit of 0.4 have shown no detrimental ghost order effects either. One can conclude that

1. **The value of a_{max} as determined from considerations of energy in comb orders higher than the prime orders is not at fault, in that simulations verify these orders do not affect the spectrum even at this limiting value.**

2. **Though useful in pointing to the right neighbourhood of the Zernike Limit, another effect must be responsible for the observed spectrum deviations.**

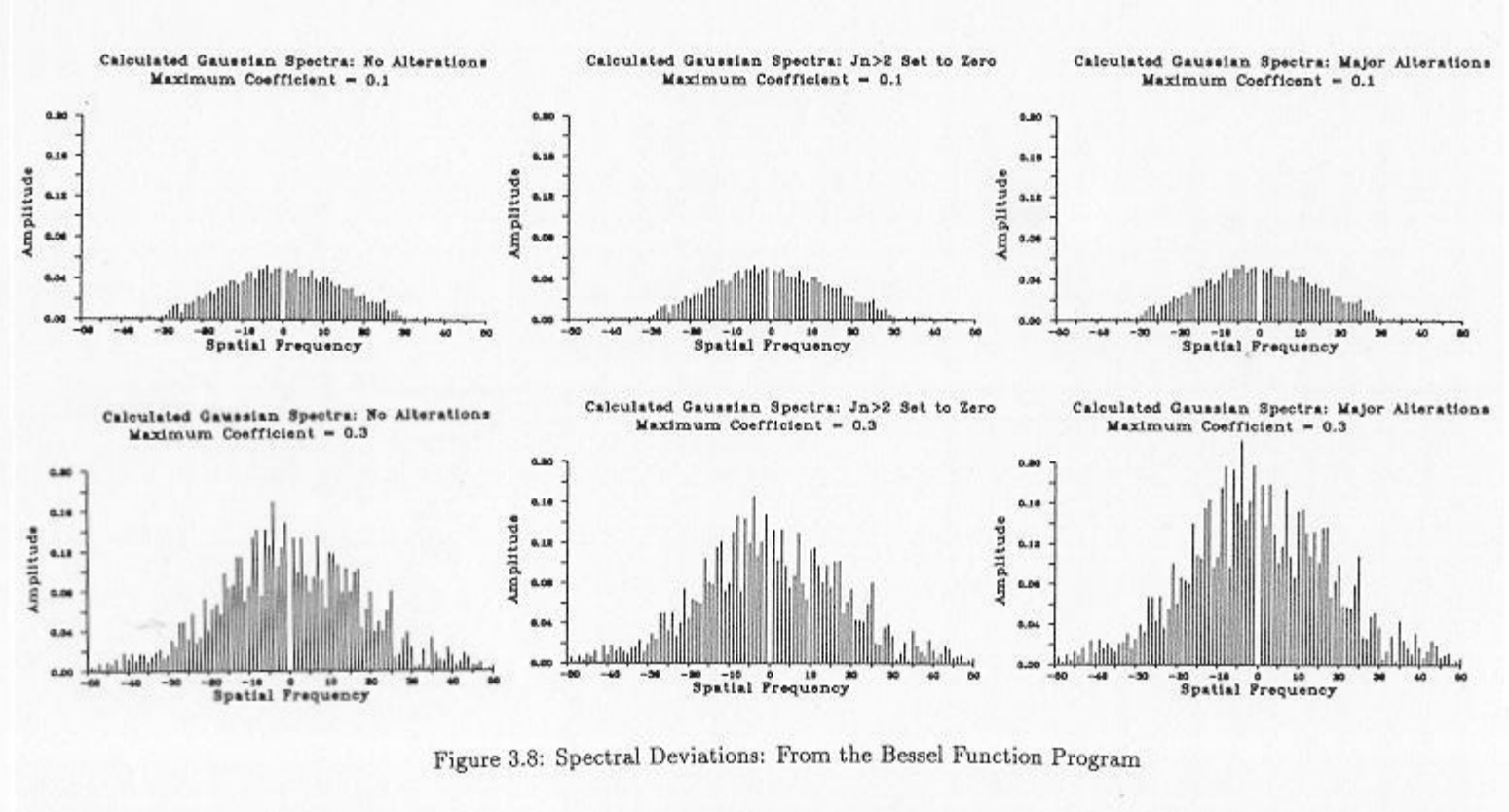

Figure 3.8: Spectral Deviations: From the Bessel Function Program

One might then suspect the linearity of $J_{\pm1}(a_j)$ to be the cause of such large spectral deviations. The error, although small, may propagate and become amplified through the many further convolution stages which result in the final spectrum. To check whether this occurs figure 3.8c shows the spectrum resulting from replacing the primary orders of each comb with the ideal *linear* values of $[(a_j)/2]$. The zero order Bessel function at the origin of each comb was also set to unity, so as not to undo the effects of the change to the prime orders. (As might be expected, removal of the ghost orders causes no difference to any resulting spectrum as their effect is negligible.) It is again observed that no appreciable difference is observed in the spectral profile after this operation, other than a slight overall increase in magnitude of the spectra resulting from the operation on J_0. Therefore,

1. **Non-linear effects present in the prime orders of each comb cannot be held responsible for the degradation of the spectral integrity.**

Convolution

The principle convolution at each stage of the spectrum formation process centres the Bessel comb on the frequency origin and gives rise to the correct positioning of the prime orders of that comb. Unfortunately further convolutions must also occur with every other δ-function of the previous unit comb and it is only these which can explain the spectral error. One such convolution is depicted in figure 3.9 where the N=2 comb sits on top of the first (negative) order δ-function of the unit comb. The resulting amplitude at $\gamma = +1$ is then

$$A(\gamma = +1) = J_{+1}(a_1)\, i^{+1}\, e^{i\Phi_1} + J_{-1}(a_1)\, i^{-1}\, e^{-i\Phi_1} \times J_{+1}(a_2)\, i^{+1}\, e^{i\Phi_2} \qquad (16)$$

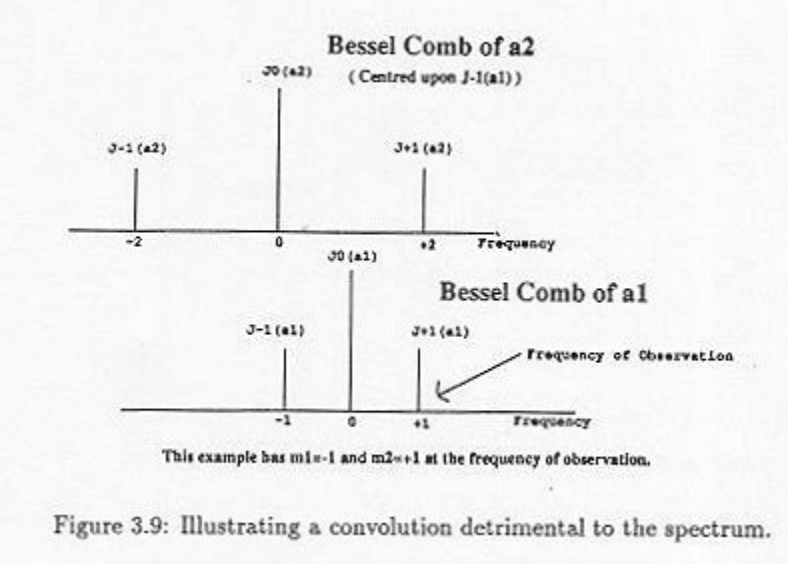

Figure 3.9: Illustrating a convolution detrimental to the spectrum.

Figure 3.9: Illustrating a convolution detrimental to the spectrum.

The second line of this equation is the error term, and an estimation of its importance can be gained as follows: pick any value for $J_{+1}(a_1)$ and a similar value for $J_{+1}(a_2)$. Let `R' denote the ratio $[(J_{+1}(a_1))/(J_{+1}(a_2))]$ and `X' denote the specific value of $J_{+1}(a_1)$, so that the above expression reduces to

$$A(\gamma = +1) = X\, e^{i(\Phi_1 + [(\pi)/2])} - RX^2 e^{i(\Phi_2 - \Phi_1)} \qquad (17)$$

As discussed previously, the error term causes a small variation in the amplitude and phase of the spectrum at $\gamma = +1$, the amount of which is dependent both on R and the phases Φ_j. This is illustrated in figure 3.10. This error cannot be eliminated as this would require elimination of every comb prime order pair also, but is observed to obey an X^2 law. As such, it may be expected to increase dramatically as X (or $J_1(a_j)$) increases.

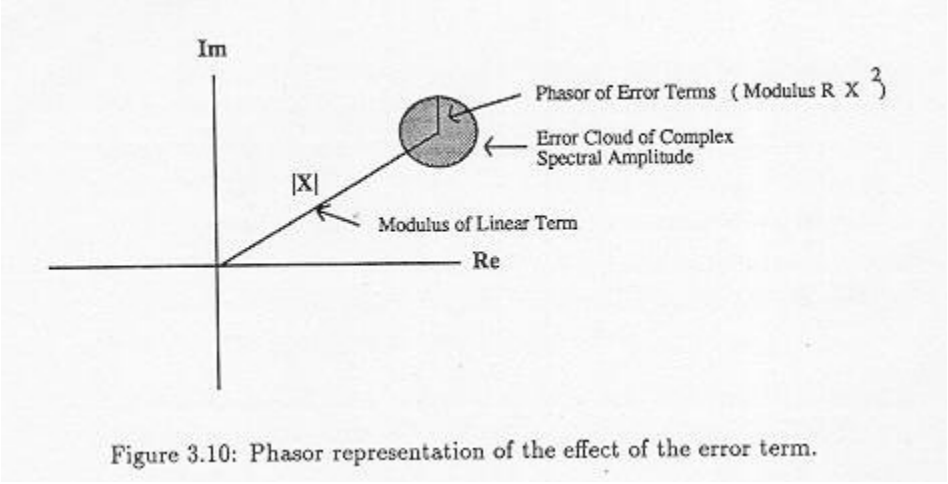

Figure 3.10: Phasor representation of the effect of the error term.

Figure 3.10: Phasor representation of the effect of the error term.

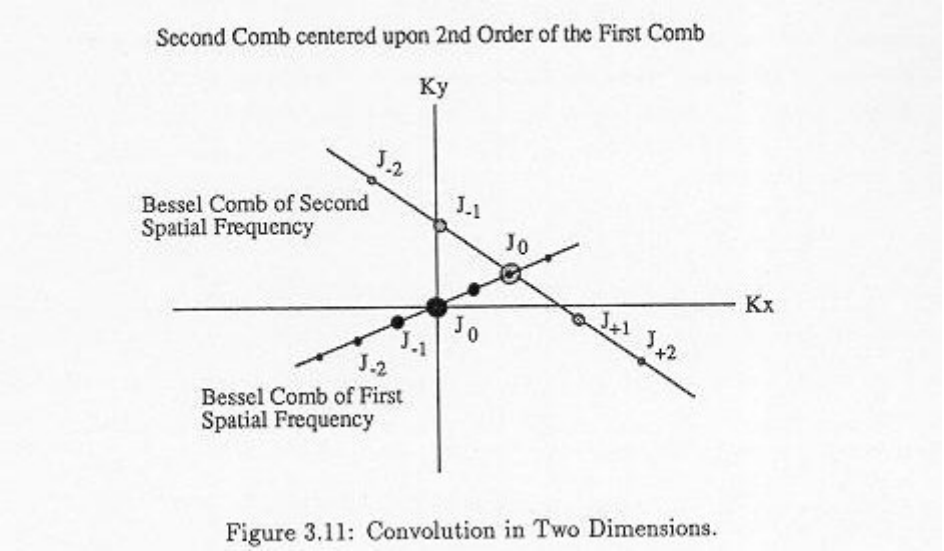

Figure 3.11: Convolution in Two Dimensions.

Figure 3.11: Convolution in Two Dimensions.

3.1 Summary

1. **The Taylor expansion of an exponential allows one to place an upper limit on the size of phase retardance which can be imaged linearly by the Zernike phase contrast operation, this conventionally being [(λ)/10]. It has been shown that even if the phase object is extremely `weak' spectrum deviations still occur and consequently harm the linearity of the image intensity with object phase. Figures 3.3 through to 3.6 show a progressive worsening of spectral integrity and image linearity as the object phase approaches the limit of what is generally considered a weak phase object.**

2. **Thoughtful application of a program which mimics the process at work in the formation of a phase object spectrum has thus allowed perhaps the two most obvious explanations of spectral deviations to be discounted. Purely numerical techniques of computation such as the FFT could not have done this. It is concluded from the observations in this section that spectral deviations of significant proportion occur due to the fundamental process at work in forming the spectrum of even a very weak phase object, which is convolution.**

Footnotes:

[1]**In any real optical system light is diffracted out of the capture zone of each lens, but the constant in question still approximates the total light energy falling on the extent of the object plane.**

[2]**As f(x) denotes a phase retardance, one may convert its angular size to the `fraction of a wavelength' form by division by 2π. Hence, a Zernike Limit of 0.4 requires each component sinusoid to have a depth of modulation not exceeding 2× [0.4/(2π)] = [(λ)/(4π)].**

[3]**For the very special set of phase objects which have a binary retardance however, this is exactly what happens as shall be shown in chapter four.**

[4]**The factor of 2 arises because the scatterer sees a Doppler shifted incident wave oscillating with *that* frequency and emits a wave which is again Doppler shifted due to the translational motion of the scatterer.**

[5]**A Raleigh variable obeys the distribution law $p(x)=[x/(\alpha)]e^{-[(x^2)/(2\alpha)]}$**

File translated from TEX by TTH, version 3.03.
On 27 Oct 2001, 23:41.