Phase Retrieval at Millimetre and Submillimetre Wavelengths using a Gaussian-Beam Formalism.

K.G. Isaak and S. Withington
Cavendish Laboratory, Madingley Rd., Cambridge, U.K.

March 26, 1993

Abstract.

In this paper, we present a technique for recovering the complex mode coefficients of a propagating beam from intensity measurements alone. Such a technique is important for experimentally characterising the behaviour of millimetre and submillimetre-wave optical systems. We describe a phase-retrieval method, based on the Gaussian-mode formalism, which is suitable for analysing the behaviour of cylindrically symmetric systems; the technique can easily be extended to cover asymmetric systems. We validate the basic procedure by recovering the complex mode coefficients of a number of different simulated amplitude distributions.

1 Introduction.

There now exists a complete range of techniques for studying and optimising the behaviour of long-focal-length millimetre and submillimetre-wave optical systems [1]. For example, a procedure exists for determining the truncation loss at any plane in a complicated quasioptical system [2], a technique has been developed for calculating the way in which power is scattered between modes at diffracting apertures [3], and a procedure has been reported for studying the aberrative effects of off-axis mirrors on collimated fields [4]. To a large extent the above techniques can only be used if the complex mode coefficients of the illuminating beam are known, and these can only be determined by evaluating the overlap integrals across some surface in the system for which the amplitude and *phase* of the field are, by some other means, already known. Indeed, the near- and far-field power patterns of corrugated [5], conical [6], and diagonal horns [7] have all been studied in this way. In many cases, however, the complex mode coefficients are difficult to determine, and a direct experimental method is required. For example, the power patterns of planar antennas and horn-lens combinations are particularly difficult to calculate, as are the power patterns of horns that have been modified for close-packed arrays. Even with nominally well-determined overlap integrals, there is often disagreement between the calculated and experimental beam patterns.

At microwave frequencies it is possible to determine the phase of a field experimentally [8] but this is much more difficult at submillimetre wavelengths, and therefore, some way of inferring phase from amplitude measurements alone is required. It should be appreciated that fitting a simple Gaussian to the far-field power pattern of a beam is of limited value, because the Gaussian only gives information about the way in which the scale size of the beam varies as the beam propagates: it gives no information about the way in which the form of the beam varies.

In this paper, we introduce, describe, and demonstrate a phase-retrieval technique that uses the Gaussian-mode formalism. It seems that, because of its modal nature, Gaussian optics is particularly well suited to the phase retrieval process. That is to say, we should be able to measure the intensity distribution at two planes in an optical system and retrieve information abut the phase. In section 2 we give a brief outline of the history and theory of phase retrieval. In section 3 we introduce the numerical techniques and simulations that we have used to test the viability of the technique. Finally in section 4, we describe and discuss the results of a number of simulations based on the beams produced by Bessel, Gaussian and uniformly-illuminated aperture distributions.

2 Theory

The problem of recovering phase from amplitude measurements alone was first addressed by optical and electron microscopists. In the early 1970s, Gerchberg and Saxton [9] [10] suggested a technique for recovering phase from intensity measurements in the image and the diffraction planes of an electron microscope. By making use of the Fourier transform relationship between the two planes, they devised an algorithm to solve iteratively for the phase. A variation to this method was suggested by Misell [11], whereby the phase information was recovered by Fourier transforming between two slightly defocused far-field images. More recently, the technique has been applied to the microwave region for the purposes of antenna metrology [12]. According to Misell's algorithm, one makes a guess at the aperture field and then Fourier transforms this into the far field. The calculated phases are then combined with the real amplitudes to make a composite field which is transformed back to the aperture. The model antenna is then defocused, by introducing an appropriate phase term, and the far-field beam pattern recalculated. By iterating between two slightly defocused intensity distributions, the phase across the aperture can be recovered. As far as large-scale antenna metrology is concerned, the main problem is obtaining access to the far-field beam pattern. In recent times, a number of groups have concentrated on recovering phase from near-field data [13] [14] [15]. The Misell algorithm works reasonably well, although at some level problems occur due to large-scale phase errors. For example, the pointing and focusing of the actual data may differ from that of the model. Lasenby and colleagues [16] have solved this problem by using least-square fits to Zernike polynomials. In fact, they now use an algorithm where both the polynomial coefficients, and the discretised amplitudes and phases, are fitted to the intensity measurements through a least-squares procedure. This is the method preferred by an Italian group who have studied the properties of such minimisation techniques in great detail [17]. Indeed, they have devised a procedure for studying the modal content of laser beams [18], and this is essentially the technique we wish to apply to submillimetre-wave optical systems.

If the beam to be studied is circularly symmetric, the intensity distribution at some distance z from the waist can be expressed as

$$I \propto \left| \sum_n A_n L_n\left(\frac{2r^2}{w(z)^2}\right) exp\left(\frac{-r^2}{w(z)^2}\right) exp\left(j2ntan^{-1}\frac{z}{z_c}\right) \right|^2 , \qquad (1)$$

where A_n are the complex mode coefficients, r is the radial position in the measurement plane, $w(z)$ is the Gaussian beam radius at the measurement plane, and z_c the confocal distance. The same expression can be written in the discretised form

$$I(r,z) \propto \left| \sum_n A_n \psi_n(m,i) \right|^2 \quad (2)$$

where ψ_n is a basis function appropriate to the symmetry properties of the source, and m and i characterise the data plane and data point respectively.

In the first equation, the Gaussian radius characterises the scale size of the beam at a plane, whereas the phase factor characterises the form of the beam. By noting that at large z the phase term tends to $\frac{\pi}{2}$ one can see the analogy between mode propagation and Fourier optics. It would seem that at least two intensity measurements are needed to constrain the mode coefficients. The planes used in the Gerchberg-Saxton algorithm are necessarily the diffraction and image planes, whereas in the Misell algorithm any two slightly defocused far-field planes are used. In the context of Gaussian modes, we can use any number of intensity distributions spaced in any way across phase-slippage space. Indeed, the main advantage of the Gaussian-mode approach is that one simply changes plane by changing a single parameter: the phase slippage. From a practical point of view, the Gaussian-mode approach is straightforward also, because it is possible to get access to any plane simply by inserting a long-focal-length off-axis mirror.

A key feature of the modal approach is that one does not have to calculate Fourier transforms at all, and the derivatives of the error function with respect to the real and imaginary parts of the mode coefficients are easy to calculate. The quantity we wish to minimise is

$$\chi^2 = \sum_m \sum_i \frac{(a_m |E_{mi}|^2 - D_{mi})^2}{\sigma^2_{mi}} \quad (3)$$

where we have included plane-dependent normalisation factors a_m to account for experimental variables and beam-dilution effects. In this equation, E_{mi} is the modelled field at point i on plane m, and D_{mi} is the actual data. As has already been seen, the intensity has a nonlinear dependence on the mode coefficients and hence the problem-solving process is necessarily an iterative one.

3 Numerical considerations.

The algorithm chosen to minimise the above function is based on a method suggested by Levenberg and Marquardt [19]. Close to the solution, χ^2 can be well approximated by the quadratic form. Far from the minimum, however, the second-order approximation to the Taylor expansion of χ^2 is no longer appropriate, and the best one can do is to take a step down the steepest gradient. The elegance of the Levenberg-Marquardt method lies in its ability to start with the latter process, swapping to the so-called inverse Hessian method as the minimum is approached. As minimisation proceeds, one has to determine at what point the process should be terminated. When working with noisy data it would seem meaningful to terminate the process when χ^2 = number of data points.

A problem often encountered in function minimisation is that of local minima. Rigorous studies by Isernia et al. [20] have shown that these can be largely overcome by weighting the chisquared sum by $1/D^2_{mi}$. This scaling has the effect of broadening out the quadratic surface in parameter space along which one moves as part of the minimisation process. This modification may be particularly important when analysing noisy submillimetre-wave data. The weighting means, however, that the best believable

χ^2 is no longer as above, and another convergence criterion must be used. At present, the minimisation process is allowed to continue until the fractional change between successive iterations falls below 0.005. Additional code, based on other nonlinear minimisation techniques [21], has been written to enable us to distinguish between idiosyncrasies in the minimisation procedure and undesirable features of the functional itself.

We consider hypothetical, circularly-symmetric horns as the basis for our simulations. This is merely for simplicity in the first instance. As can be seen from (2), the phase-retrieval approach is completely general and with a change of basis set, from Laguerre to associated Laguerre or Hermite polynomials, one can readily extend the phase-retrieval procedure to asymmetric beams.

Careful consideration must be given to the way in which the intensity distributions are sampled. In some sense, each Gaussian mode can be considered to represent one transverse spatial-frequency component. The intrinsic sampling can then be approximated by the Nyquist sampling rate. As one is sampling a power distribution it is necessary to sample at twice this rate. In the case of the nth Laguerre-Gaussian mode there are $\frac{(2n+1)}{4}$ periods over the extent of the mode, and therefore to sample at the Nyquist power frequency, one needs to sample $(2n+1)$ points over the power pattern. We oversample using 4n data points. It can be shown [22] that the spatial extent of a Laguerre-Gaussian polynomial is given by $w(z)\sqrt{2n+1}$ where $w(z)$ is the Gaussian beam radius at the measurement plane. Preliminary tests indicated that sampling to $w(z)\sqrt{2n}$ produces better convergence , and it is to this limit that we sample. As the beam waist is a variable parameter, there is strictly an infinite number of mode sets that may be used to expand the propagating beam. Usually, the beam waist is set to maximise the power in the lowest-order mode: in the case of a corrugated horn this would correspond to a waist-to-aperture size ratio at the aperture of 0.6435. This value, while minimising the short comings of the propagation of a single Gaussian mode, does not necessarily sample the aperture distribution in an optimal way. Indeed, thermodynamically, this value must be wrong as it does not correctly take into account the number of degrees of freedom in the image. We have started to explore the effect of different mode sets on the efficiency of the retrieval process.

4 Results.

To demonstrate the phase-retrieval method we have produced simulated data based on the beam patterns of Gaussian, Bessel, and uniformly-illuminated aperture distributions—the Bessel-function distribution corresponds exactly to the field produced by a corrugated horn.

In each case, the data planes were oversampled out to $w(z)\sqrt{2n}$ where $w(z)$ and n are the Gaussian beam radius at the measurement plane and the number of modes used in each data set respectively. 4n data points were used to oversample each distribution. To perform a retrieval one requires the data from two planes and an initial set of guesses to the mode coefficients: this determines the position at which one starts on the χ^2 surface. Various initial trial sets have been used: all coefficients equal, all coefficients except the fundamental set to zero, and all coefficients randomly chosen. The later was intended to simulate the random phases used as the starting condition by Misell and others. It was found that in some cases the successful recovery of the mode coefficients was dependent on the trial guess. For all of our simulations we used two far-field planes; a preliminary attempt at finding the optimum data planes failed to produce any conclusive results.

Initially, we attempted to recover phase from distributions simulated using 3 and 5 mode coefficients. The results shown in this paper are based on the recovery of 10 modes. In Figs 1,2, and 3, we show the corrugated-horn, uniformly-illuminated, and truncated-Gaussian aperture distributions that have been simulated and successfully retrieved. In each case, plot (a) shows the power distribution in each of the two far-field data planes, plot (b) shows the recovered and actual intensity distribution in the aperture, and plot (c) shows the actual and recovered phase distribution in the aperture. The plots show that excellent agreement between actual and recovered aperture distributions can be achieved. In these plots, the phase across the aperture is flat. This occurs because the Gaussian beam-mode radius of curvature is set equal to the actual field radius of curvature in the aperture. As a consequence, the mode coefficients do not contribute to the phase. A further issue is that the retrieval process does not constrain the absolute phase. We could use the retrieval process to force the imaginary part of the zeroth-order mode coefficient to zero if desired. Tables 1,2, and 3 further illustrate the success of the technique. The unconstrained phase of the fundamental mode can be clearly seen, as can the constant phase difference between the modes. This phase is simply the phase slippage associated with the dispersion of the propagating modes, and it agrees almost exactly with the expected value. Perhaps one of the most satisfying aspects of these simulations is the dynamic range that can be retrieved. For example, in the case of the corrugated horn, the fundamental and the first higher-order mode span 4 orders of magnitude.

In each case, the minimisation was allowed to continue until the weighted χ^2 between successive iterations was less than 0.005. This loop condition was found to be both under and over stringent. Reconstructions of the power patterns based on the retrieved mode coefficients taken after just a few iterations were often indistinguishable from the power patterns constructed using the final solutions. In contrast, there were a few cases where the reconstructions bore no resemblance to the simulated data set, despite the convergence criterion having, in some senses, being met. A typical retrieval of 10 mode coefficients might take around 60 iterations, although we believe that it should be possible to reduce this number significantly. Finally, the algorithm was also tested without the weighting factor $1/D_{mi}^2$. This resulted in poor convergence properties, supporting the view that weighting χ^2 reduces the problem of local minima.

5 Conclusions.

We have shown that it is possible to recover the complex mode coefficients of a propagating submillimetre-wave beam through intensity measurements alone. The mode coefficients can be recovered with high dynamic range. We are not satisfied with the performance of the Levenberg-Marquardt minimisation technique as applied to this particular phase-retrieval problem, and therefore, we are currently investigating more robust methods. The technique will be extremely useful for characterising the behaviour of millimetre and submillimetre-wave optical systems and components.

References

[1] S. Withington, J.A. Murphy, A. Egan, and R.E. Hills,"On the design of broadband quasioptical systems for submillimetre-wave radio-astronomy receivers," Int. Journal of Infrared and Millimeter Waves, vol. 13, no. 10, pp. 1515-1537, 1992.

[2] J.A. Murphy, A. Egan, and S. Withington, "Truncation in millimetre and submillimetre-wave optical systems," To be published in IEEE Trans. Antennas and Propagat.

[3] J.A. Murphy, S. Withington, and A. Egan, "Mode conversion at diffracting apertures in millimetre and submillimetre-wave optical systems," To be published in IEEE Trans. Microwave Theory Tech.

[4] S. Withington and J.A. Murphy, "Gaussian-mode analysis of "thin" mirrors," Proceedings of the Fourth International Symposium on Space Terahertz Technology, see this issue.

[5] R.J. Wylde, "Millimeter-wave Gaussian beam-mode optics and corrugated feed horns," Proc. Inst. Elec. Eng., vol. 131, pt. H, pp. 258-262, 1984.

[6] J. A. Murphy, "Aperture efficiencies of large axisymmetric reflector antennas fed by conical horns," IEEE Trans. Antennas and Propagat., vol. AP-36, pp. 570-575, 1988.

[7] S. Withington and J. A. Murphy, "Analysis of diagonal horns through Gaussian-Hermite modes," IEEE Trans. Antennas Propagat., vol. AP-40, no. 2, pp. 198-206, 1992.

[8] G. Hygate and J.F Nye, "Measuring microwave fields directly with an optically modulated scatterer," Meas. Sci. Technol., vol 1., pp. 707-709, 1990.

[9] R.W. Gerchberg and W.O. Saxton, "Phase determination from image and diffraction plane pictures in the electron microscope," Optik, vol. 34, pp. 275-284, 1971.

[10] R.W. Gerchberg and W.O. Saxton, "A practical algorithm for the determination of phase from image and diffraction plane pictures," Optik, vol. 35, pp.237-246, 1972.

[11] D.L. Misell, "A method for the solution of the phase problem in electron microscopy," J. Phys. D:Appl.Phys., vol. 6, L6-L9, 1973.

[12] D. Morris, "Phase retrieval in the radio holography of reflector antennas and radio telescopes," IEEE Trans. Antennas and Propagat., vol. AP-33, pp. 749-755, 1985.

[13] T. Isernia, R. Pierri, and G. Leone, "New technique for estimation of farfield from near-zone phaseless data," Electronics Letters, vol. 27, no. 8, pp. 652-654, 1991.

[14] T. Isernia, G. Leone, and R. Pierri, "New approach to antenna testing from near field phaseless data: the cylindrical scanning," IEE Proceedings-H, vol. 139, no.4, pp.363-368, 1992.

[15] A.P Anderson and S. Sali, "New possibilities for phaseless microwave diagnostics. Part1: Error reduction techniques". IEE Proceedings-H, vol.132, no.4, pp291-298, 1986.

[16] A.N. Lasenby, "Phase retrieval using HST images," In Proc. of Workshop on Hubble Space Telescope.

[17] T. Isernia, G. Leone, and R. Pierri,"A quadratic inverse problem: the phase retrieval,"in Inverse Methods in Action, P.C. Sabatier, ed. (Springer-Verlag, Berlin, 1990), p. 285.

[18] A. Cutolo, A. Esposito, T. Isernia, R. Pierri, and L. Zeni,"Characterization of the transverse modes in a laser beam: analysis and application to a Q-switched Nd:YAG laser,"Applied Optics, vol. 31, no. 15, pp. 2722-2733, 1992.

[19] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, "Numerical recipes in C,"Cambridge University Press, 1988.

[20] T. Isernia and R. Pierri, Private communication.

[21] The Numerical Algorithms Group Ltd, 1990.

[22] R.L. Phillips, L.C. Andrews,"Spot size and Divergence for laguerre gaussian beams of any order,"Applied Optics, vol 22, no 5, pp. 643-644, 1983.

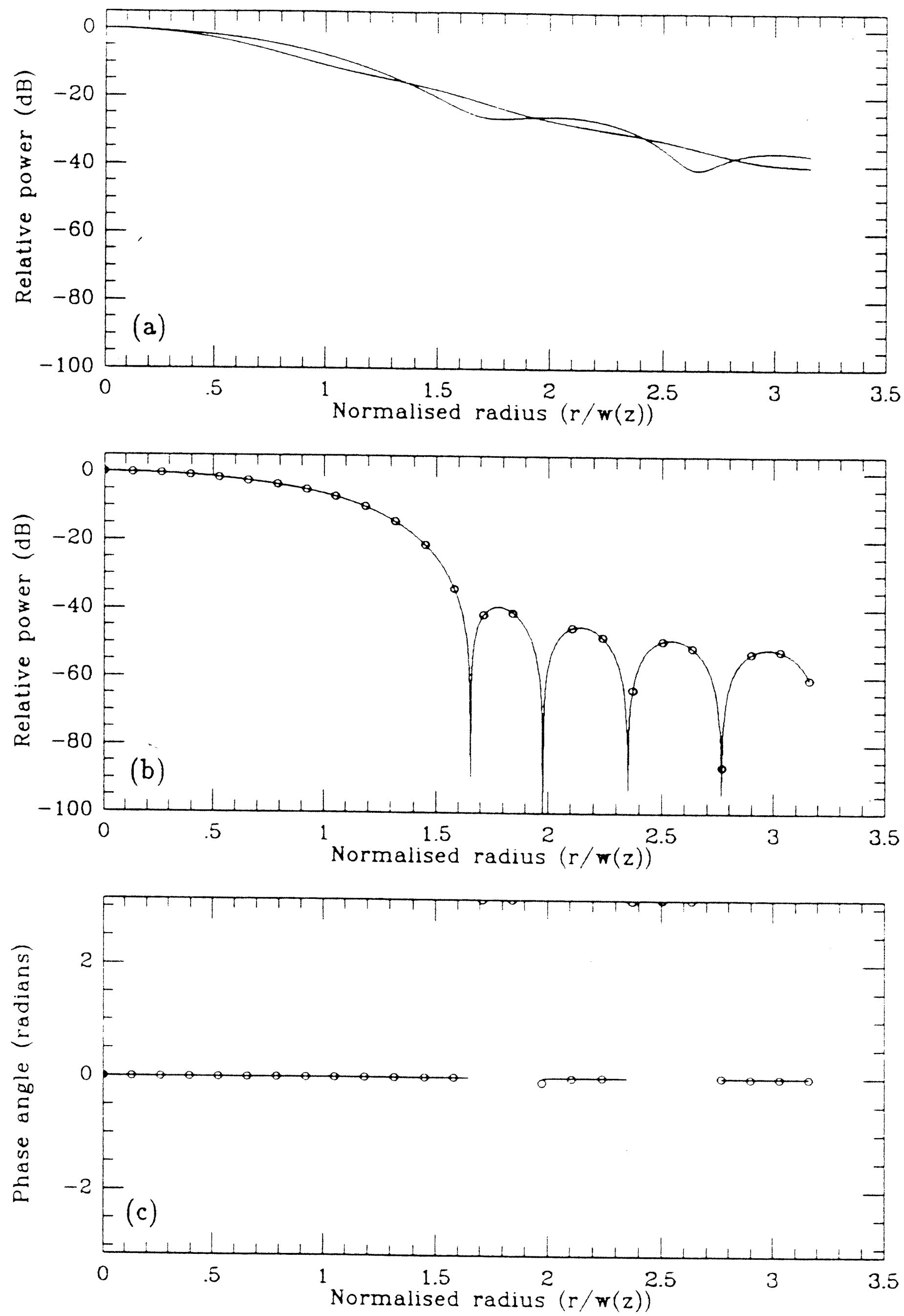

Figure 1: Retrieval of Bessel function aperture distribution. (a)Power pattern of the data planes (b) Simulated (solid line) and retrieved (open circles) power pattern as referenced to the aperture (c) Simulated (solid line) and retrieved (open circles) phase across the aperture.

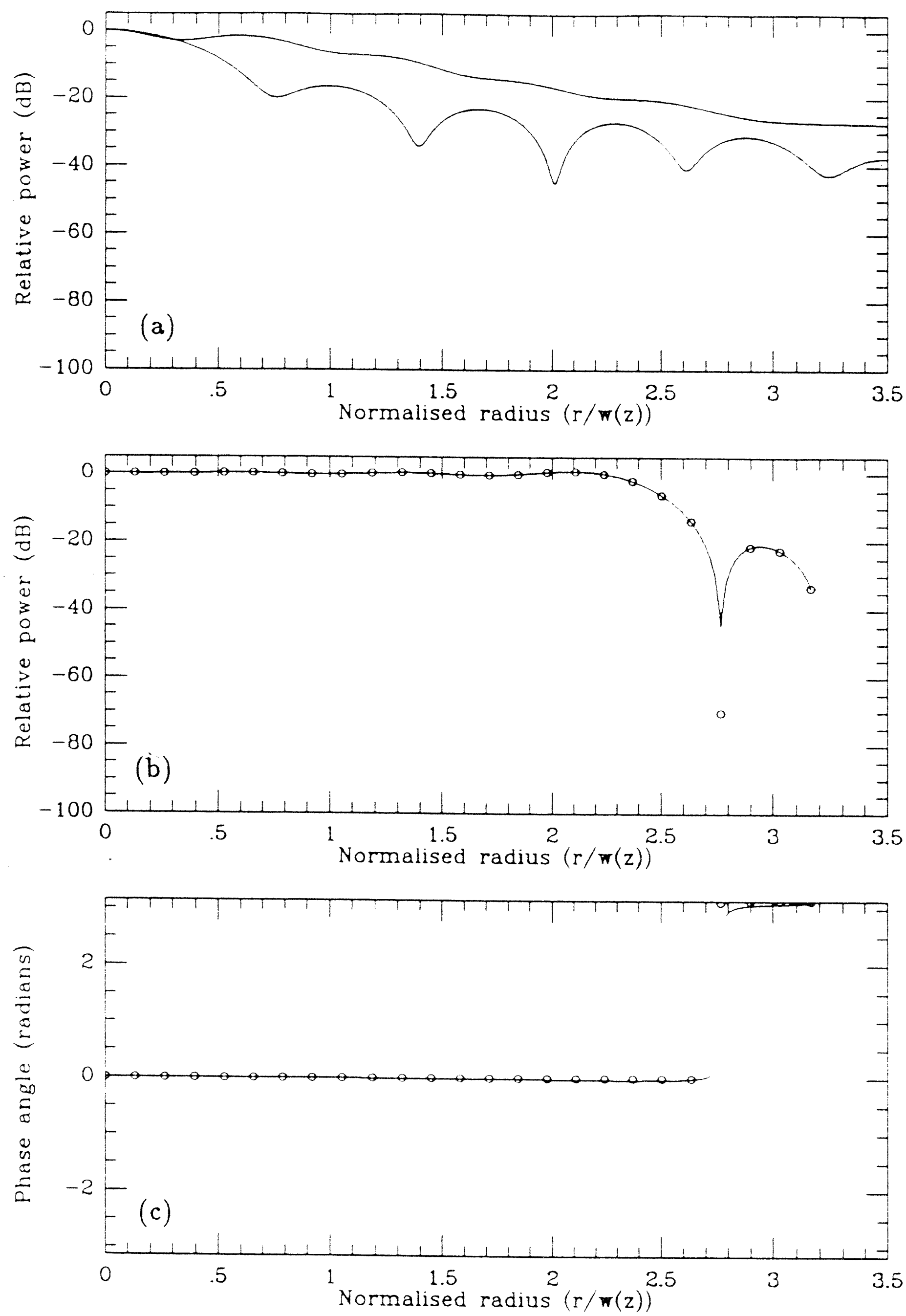

Figure 2: Retrieval of Uniformally illuminated aperture distribution. (a)Power pattern of the data planes (b) Simulated (solid line) and retrieved (open circles) power pattern as referenced to the aperture (c) Simulated (solid line) and retrieved (open circles) phase across the aperture.

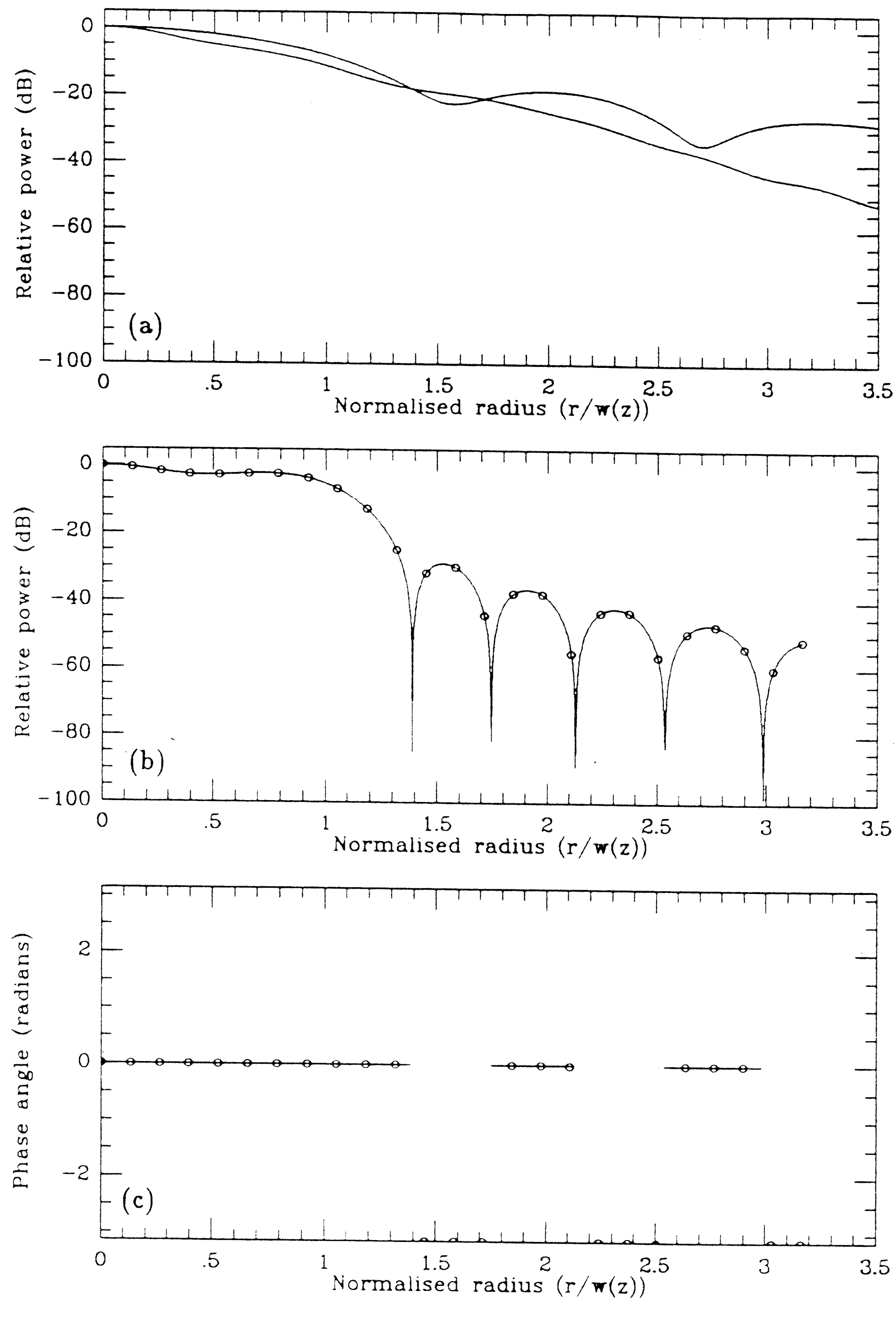

Figure 3: Retrieval of a truncated gaussian aperture distribution. (a)Power pattern of the data planes (b) Simulated (solid line) and retrieved (open circles) power pattern as referenced to the aperture (c) Simulated (solid line) and retrieved (open circles) phase across the aperture.

Table 1: Simulated and retrieved mode coefficients for a corrugated horn aperture distribution.

Simulated mode coefficients.	Corrected mode amplitudes	Retrieved mode amplitudes	Retrieved phase (radians)	Corrected Δphase between modes (radians).
0.99033	0.99059	0.99060	-1.5614	
-3.3966e-05	-3.3975e-05	1.7793e-05	0.5460	1.0342
-0.12050	-0.12054	0.12052	0.6205	-0.0744
-4.3073e-02	-4.3085e-02	4.3098e-02	0.1409	0.4795
1.9586e-02	1.9591e-02	1.9575e-02	2.8026	0.4799
3.4120e-02	3.4130e-02	3.4134e-02	2.3229	0.4797
1.9988e-02	1.9994e-02	2.0001e-02	1.8429	0.4800
1.7311e-04	1.7316e-04	1.7826e-04	1.3584	0.4844
-1.2536e-02	-1.2539e-02	1.2541e-02	-2.2579	0.4747
-1.5204e-02	-1.5298e-02	1.5211e-02	-2.7380	0.4802

note: The corrected amplitudes are renormalised to account for the fact that only 99.95% of the total power is contained within the first 10 modes.The expected phase difference between modes is $\Delta = 2arctan\frac{\pi k_w{}^2 ah^2}{4lh\lambda}$, which for the horn dimensions used is 0.4797 radians.

Table 2: Simulated and retrieved mode coefficients for a uniformly illuminated aperture distribution.

Simulated mode coefficients.	Corrected mode amplitudes	Retrieved mode amplitudes	Retrieved phase (radians)	Corrected Δphase between modes (radians).
0.56518	0.57089	0.57453	-0.4218	
-0.55150	-0.56064	0.56064	2.5314	0.1884
0.47975	0.48771	0.48771	-0.7986	0.1884
-0.28107	-0.28573	0.28573	2.1546	0.1884
-5.9638e-03	-6.0627e-03	6.0649e-03	1.9659	0.1887
0.14504	0.14745	0.14745	-1.3638	0.1881
-1.16756e-02	-1.1869e-02	1.1869e-02	1.5894	0.1883
-9.72321e-02	-9.8844e-02	9.8845e-02	1.4009	0.1885
-2.32536e-02	-2.3639e-02	2.3639e-02	1.2125	0.1884
6.10623e-02	6.2075e-02	6.2075e-02	-2.1175	0.1884

note: The corrected amplitudes are renormalised to account for the fact that only 96.76% of the total power is contained within the first 10 modes. The expected phase difference between modes is $\Delta = 2arctan\frac{\pi k_w{}^2 ah^2}{4lh\lambda}$, which for the horn dimensions used is 0.1884 radians.

Table 3: Simulated and retrieved mode coefficients for a truncated gaussian aperture distribution - truncated at $\frac{1}{e}$ point at the aperture.

Simulated mode coefficients.	Corrected mode amplitudes	Retrieved mode amplitudes	Retrieved phase (radians)	Corrected Δphase between modes (radians).
0.93843	0.95716	0.95716	0.9769	
3.0528e-02	3.1137e-02	3.1127e-02	0.1534	0.8235
-0.19301	-0.19695	0.19696	2.4715	0.8235
-0.14740	-0.15035	0.15034	1.6479	0.8236
-3.8529e-02	-3.9298e-02	3.9282e-02	0.8243	0.8236
4.6776e-02	-4.7689e-02	4.7699e-02	-3.1408	0.8236
8.4895e-02	8.6589e-02	8.6583e-02	2.3188	0.8235
8.1724e-02	8.3355e-02	8.3340e-02	1.4954	0.8231
5.3276e-02	5.4340e-02	5.4331e-02	0.6720	0.8234
1.5863e-02	1.6180e-02	1.6178e-02	-0.1515	0.8235

note: The corrected amplitudes are renormalised to account for the fact that only 96.12% of the total power is contained within the first 10 modes. The expected phase difference between modes is $\Delta = 2arctan\frac{\pi {k_w}^2 \, ah^2}{4lh\lambda}$, which for the horn dimensions used is 0.8236 radians.