Single measurements on probability density functions and their use in non-Gaussian light scattering

1996 J. Phys. A: Math. Gen. 29 7501

(http://iopscience.iop.org/0305-4470/29/23/016)

J. Phys. A: Math. Gen. **29** (1996) 7501–7517. Printed in the UK

Single measurements on probability density functions and their use in non-Gaussian light scattering

T R Watts, K I Hopcraft and T R Faulkner

Department of Theoretical Mechanics, University of Nottingham, University Park, Nottingham, NG7 2RD, UK

Received 3 June 1996, in final form 2 August 1996

Abstract. Some properties of non-Gaussian probability density functions applicable to coherent light scattering experiments are investigated using random walk theory. The values taken by these functions together with the slope of the distribution near the origin can be used to discern the number of particles present, in particular when this number is small. The behaviour of these functions in cases where there are fluctuations in the number of particles present and the scattering cross sections indicate that the conclusions drawn are generic and not sensitive to the details of the model. The technique may be applicable to subject areas other than the scattering of coherent light.

1. Introduction

The concept of a random walk has proved to be of great utility and has been employed with considerable success in diverse areas of mathematics, the physical and biological sciences (see [1] and [2] for example). Rayleigh's study of the statistical properties of sound [3] was an early example of the use of a two-dimensional random walk giving an elegant physical manifestation of the central limit theorem of statistics. His results gained new impetus with the advent of the laser, where the interference of scattered coherent light leads to 'speckle'. The real and imaginary parts of the complex scattered field each have Gaussian distributions, so that the resultant amplitude has a Rayleigh distribution and the resultant intensity is exponentially distributed. The epithet 'Gaussian' is used to describe a random variable having any of these distributions. The random walk has subsequently been used in optics to model the many instances of non-Gaussian statistics that can arise [4, 5].

Pearson [1] and Kluyver [6] studied the problem of random walks as the number N of identical length steps increased. The probability distributions of the resultant distance from the origin were found and the approach to the Gaussian limit could thus be observed. For $N < 7$ the distributions are manifestly non-Gaussian but, for $N \geqslant 7$, the appearance of the intensity distribution is difficult to distinguish from the exponential limit distribution. Nevertheless, by calculating the moments of the distributions, the departures from the Gaussian limit can be quantified for an arbitrary number of steps. For $N = 10$ the second normalized intensity moment is 10% less than its Gaussian value of 2 and the third normalized moment is some 8.5% in excess of its Gaussian value of 6.

Pusey *et al* [7] interpreted many of the results contained in [1] in the context of coherent light scattered from a fixed number and from a fluctuating population of statistically independent but identical scatterers. They demonstrated that for a large number of steps (or

0305-4470/96/237501+17$19.50

particles) the statistics of the scattered light were approximately Gaussian and that departures from the Gaussian limit could be used to infer some properties of the individual steps. This information could, in principle, be extracted from an analysis of the moments of the resultant distribution rather than from the probability density function itself. The moments exhibit enhanced values that are usually in excess of the Gaussian values when the step sizes vary and when N is small. The magnitude of this enhancement factor depend on the details of the step-size distribution or, equivalently, the physical properties of the scatterers. A principal difficulty of using second- and higher-order moments to make quantitative deductions stems from the amount of data required to be certain of having reliable and unbiased estimates for the statistics. Having reasonable estimates of the moments is equivalent to measuring the entire probability density function accurately, especially those events occurring in the tail of the distribution. Recording the frequency of these events can be prohibitive if the random process under observation is significantly non-Gaussian in nature [8]. Another disadvantage inherent in using the moments is that the information about the number fluctuations and the scattering cross sections are difficult to decouple from one another. Although it is possible to achieve this decoupling in some instances [9, 10], prior information about the nature of the scatterers must first be assumed and the configuration for making the measurements requires multiple sensor arrays.

In this paper the ideas first mooted in [7] are revisited but with different purview. It will be shown that it is possible to extract information about the number of scatterers present or, equivalently, the number of steps that have been taken in the random walk, by making in principle just two measurements on the probability density function itself rather than from the moments associated with it. Specifically the information sought is uniquely contained in the enhanced values of the scattered intensity probability distribution function (pdf) and its derivative at or near to the origin. The enhancement is most pronounced when the actual number or effective number of scatters is low, precisely in those circumstances when the pdf has pronounced non-Gaussian character.

Section 2 contains a brief review of the form of the probability density functions that obtain from identical particles, or equivalently for random walks of fixed step length. The principle that the number of steps can be gleaned from $P(0)$ is demonstrated for the case of a fixed number of particles being present. This situation is then generalized in section 3 to the case when the number of particles is fluctuating according to a particular discrete random process. The final generalization is made in section 4 where fluctuations in both the number of steps and in the length of steps is discussed with a view to treating scattering by non-identical but statistically similar particles. A discussion about possible experimental constraints is given and the conclusions are drawn in the final section.

2. Distributions with fixed number of steps and step lengths

Consider a two-dimensional random walk comprising N steps with each step being of fixed length r. The direction of each successive step is independent from the previous one and occurs in an entirely random direction. The resultant of such a walk may be written as

$$R\exp(\mathrm{i}\Phi) = \sum_{j=1}^{N} r\exp(\mathrm{i}\phi_j)$$

where the elemental phases ϕ_j are uniformly distributed over 2π radians. The probability distribution of the resultant phase Φ is also uniformly distributed over 2π radians and it is straightforward to show that the resultant amplitude after N steps has probability distribution given by [1]

$$P_N(R) = R\int_0^\infty uJ_0(uR)(J_0(ur))^N \,\mathrm{d}u$$

where $J_n(x)$ is a Bessel function of order n. Using this, the pdf of the resultant intensity $I = R^2$ is easily found to be

$$P_N(I) = \frac{1}{2}\int_0^\infty uJ_0\left(u\sqrt{I}\right)(J_0(ur))^N \,\mathrm{d}u \tag{2.1}$$

which may be evaluated analytically for certain values of N. When $N = 1$ the pdf is a delta function at $I = r^2$ because the step size is constant. Denoting the mean intensity as $\langle I\rangle = Nr^2$ allows the pdf to be written in terms of the dimensionless variable $x = I/\langle I\rangle$, whereupon the pdfs for $N = 2, 3$ may be shown to be

$$\langle I\rangle P_2(I) = \begin{cases} \dfrac{1}{\pi(x(2-x))^{1/2}} & \text{for } x < 2 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

and

$$\langle I\rangle P_3(I) = \begin{cases} \dfrac{k}{2\pi^2(3x)^{1/4}}K(k) & \text{if } 0 < x < 1/3 \\ \dfrac{1}{2\pi^2(3x)^{1/4}}K(1/k) & \text{if } 1/3 < x < 3 \\ 0 & \text{if } x > 3 \end{cases} \tag{2.3}$$

where $K(k)$ is a complete elliptic integral of the first kind [11] with modulus

$$k = 4\left(\frac{\sqrt{3x}}{\left(3-\sqrt{3x}\right)\left(1+\sqrt{3x}\right)^3}\right)^{1/2}.$$

The curves of these distributions are shown in figure 1(a) and it is clear that both are singular and discontinuous. P_2 is singular at the origin and when $x = 2$, this being the furthest possible extent from the origin after two steps, and for values of x greater than 2 it is zero. P_3 is finite at the origin, has a logarithmic singularity at $x = 1/3$ and a finite jump discontinuity when $x = 3$, after which it is zero. It is apparent that the distributions for two or three particles are very different from the limiting exponential distribution, and therefore the statistics of the scattered light would be significantly non-Gaussian.

The singularities and discontinuities in P_2 and P_3 may seem surprising at first, but it should be realized that the integrated probability over a certain range is always finite. A partial explanation of these discontinuities may be given by making a comparison with random walks in one dimension. For a one-dimensional random walk comprising two steps there are only two possible outcomes, the resultant is either zero or two step lengths, each occurring with equal probability. In two dimensions the distribution is smoother but there are still singularities corresponding to the delta functions in the one-dimensional case. For three steps in a one-dimensional walk there are still only two possible outcomes: a resultant of one-step length and three-step lengths occurring with probability $3/4$ and $1/4$ respectively. Once again these delta functions are smoothed out in two dimensions, there being a singularity at $x = 1/3$ and a discontinuity at $x = 3$ marking the end of the range. For walks having a higher number of steps, the combination of possible outcomes increases

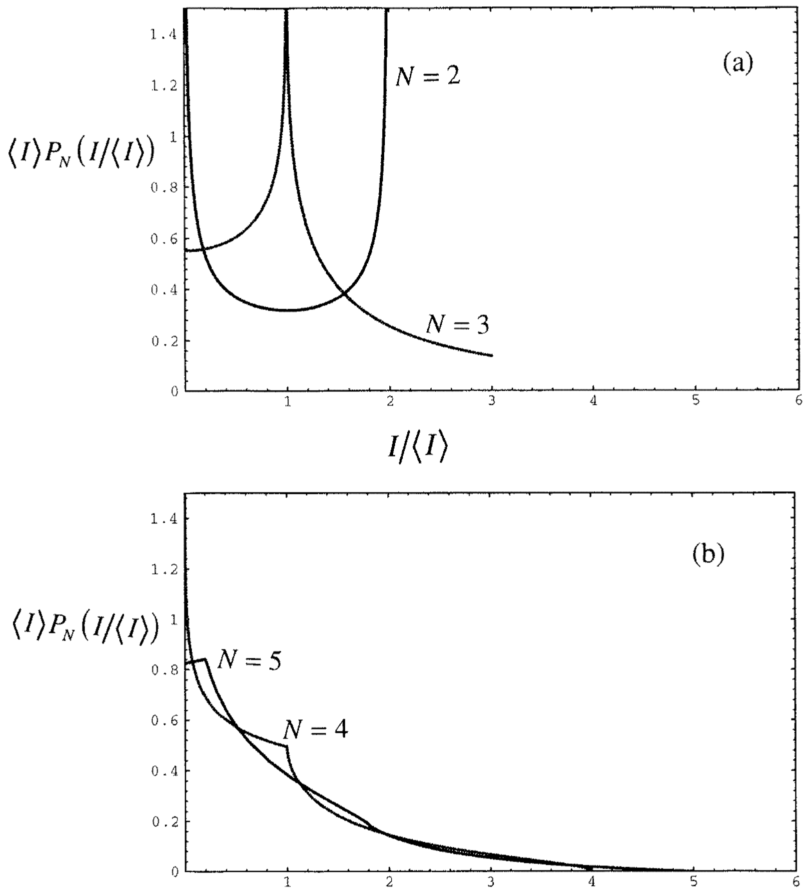

Figure 1. The pdf for random walks having unit step length and uniformly distributed phases with (a) two and three steps and (b) four and five steps.

and so a detailed comparison between the features appearing in one- and two-dimensional walks becomes increasingly difficult to interpret.

For larger values of N the distribution (2.1) is readily calculated using a convergent Fourier–Bessel series [12, 13]. This series is valid for $0 < I < \langle I \rangle$ and is given by

$$\langle I \rangle P_N(I) = \frac{1}{N} \sum_{m=1}^{\infty} J_0(j_m I^{1/2}/Nr)[J_0(j_m/N)]^N/[J_1(j_m)]^2 \tag{2.4}$$

where j_m is the mth positive zero of the Bessel function $J_0(x)$. Taking the first 400 terms in this series is usually sufficient to obtain an accurate solution and figure 1(b) shows the curves for P_4 and P_5. P_4 is infinite at the origin and possesses discontinuities in slope at $x = 1$ and $x = 4$. By contrast P_5 is finite for all x but has discontinuities in its slope at $x = 1/5$ and $x = 5$. Note that near the origin the slope of P_5 is positive, as was that for P_3. Figure 2 shows curves of the logarithm of P_6 and P_7, these being distinguishable only for larger values of the intensity. By way of comparison, the limiting exponential distribution is also depicted. A striking feature in all of these graphs is the behaviour of the distribution at or near the origin as N varies and it is this behaviour that will be exploited.

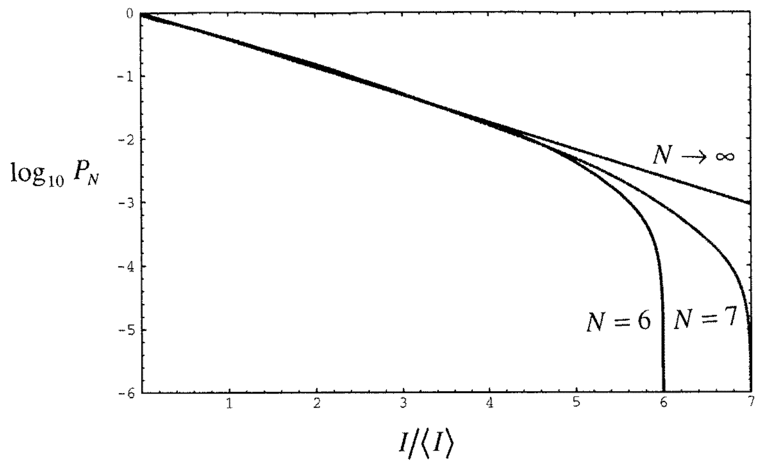

Figure 2. The logarithm of the pdfs for six- and seven-step random walks having unit step length and uniformly distributed phases together with the exponential limit distribution. The difference between the curves is only evident for larger values of the intensity.

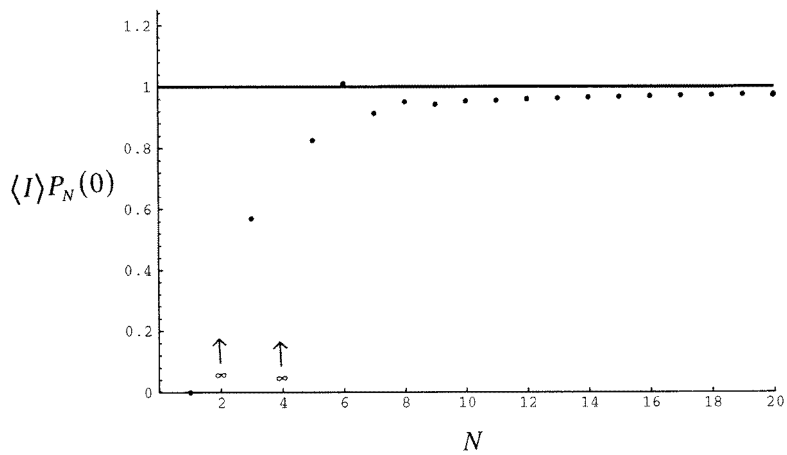

Figure 3. The normalized intensity at the origin as a function of the number of identical steps in the walk for N ranging up to 20.

Figure 3 shows the values of the dimensionless quantity $f = \langle I\rangle P_N(0)$ as a function of N ranging through 1 to 20. For $N = 1$, the value of f is clearly zero. For $N = 2$ and 4 the value of f is infinite. The analytical result (2.3) has been used to calculate f for $N = 3$, and equation (2.4) has been employed for all other values of N up to 20. The trend towards the Gaussian value of unity from below is evident for large N, but f varies greatly for smaller values of N. It will be shown that this variability can, in principle, be used to determine the value of N, but first it is instructive to consider whether this result is maintained when the number of steps in the walk fluctuates. This will be considered in the next section.

3. Distributions with varying number of steps and fixed step lengths

Calculating the pdf for the situation when the number of steps N in the walk fluctuates according to some random process requires a slight generalization to equation (2.1) equivalent to performing an average over the number fluctuation distribution. Proceeding in this way yields an averaged pdf, viz.

$$P_{\overline{N}}(I) = \frac{1}{2}\int_0^\infty u J_0\left(u\sqrt{I}\right) Q(J_0(ur))\,\mathrm{d}u \tag{3.1}$$

where $Q(s) = \sum_{N=0}^{\infty} s^N p(N)$ is the probability generating function and $p(N)$ is the probability that N steps have taken place.

If the number of steps is fluctuating according to a Poisson process [14], then (3.1) can be written as

$$P_{\overline{N}}(I) = \frac{1}{2}\exp(-\overline{N})\int_0^\infty u J_0\left(u\sqrt{I}\right)\exp(\overline{N}J_0(ur))\,\mathrm{d}u. \tag{3.2}$$

Evaluating (3.2) at the origin will always give an infinite result because of the singular contributions from P_2 and P_4. These singularities will be present irrespective of the value taken for $\overline{N}$ and so to circumvent this, (3.2) is evaluated at a point x slightly displaced from the origin where the values of all the probabilities are finite. An arbitrary value $x = 1/100$ is used throughout this section. To calculate the dimensionless quantity $f = P_{\overline{N}}(x)$ requires expansion of the exponential in the integrand of (3.2) after which the resultant pdf can be found through employing the analytical results (2.2, 2.3), the Fourier–Bessel series (2.4) and asymptotic formulae for large values of N. The details for this procedure are contained in the appendix. Figure 4 shows f and g plotted as functions of $\overline{N}$. For small values of $\overline{N}$ there is on average fewer than one particle present and so the pdf must, as it does, tend to zero. For moderate values of $\overline{N}$ there is a slight enhancement above unity, evidently caused by an average over the distributions for two, four and six particles. For larger values of $\overline{N}$ the curve approaches the Gaussian value of 1, but this time from above (cf figure 3). If $\overline{N}$ is to be estimated from this curve an additional measure is required since for each value of f there may correspond two values of $\overline{N}$. A unique estimate for $\overline{N}$ can be found by using the slope of the distribution near the origin (recall that for the case when N was fixed, the derivative could be positive or negative). Defining the dimensionless derivative to be $g = P'_{\overline{N}}(x)$ enables the slope of the distribution near the origin to be determined in a similar fashion to that for f, the results for which are shown in figure 4(b). For small values of $\overline{N}$ the slope of the distribution is small for the same reasons that the value of f is small. For larger values of $\overline{N}$, g is large and negative, this being due to the dominant contributions from the two- and four-particle distributions. As $\overline{N}$ increases, g approaches the asymptotic Gaussian limit of -1.

Another model for number fluctuations known to be of importance in non-Gaussian light scattering [15] is the negative binomial distribution for which

$$p(N) = \binom{N+\alpha-1}{N}\frac{(\overline{N}/\alpha)^N}{(1+(\overline{N}/\alpha))^{N+\alpha}} \qquad \text{for } \alpha > 0. \tag{3.3}$$

This distribution can account for the clustering of particles for small values of α and is asymptotic to the Poisson distribution when α is large. Using (3.3) in (3.1) enables both f and g to be found as before and the results are displayed in figure 5 for the particular choice $\alpha = 2$. Comparing these results with those given in figure 4 shows that the value of f now increases substantially above unity and its asymptotic value is now not given by

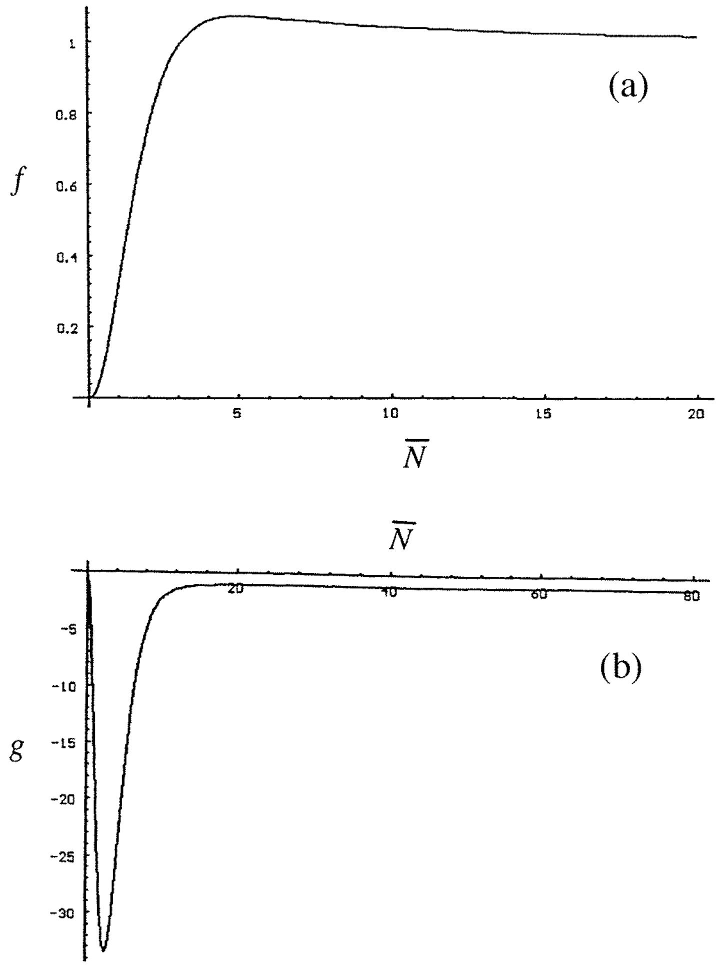

Figure 4. (a) The normalized intensity pdf f near the origin ($x = 1/100$) as a function of the number of identical steps in the walk for N fluctuating according to a Poisson random process. The approach to the Gaussian asymptote of unity occurs for sufficiently large value of N. (b) The normalized derivative of the pdf g evaluated near the origin, at the same value of x.

the Gaussian limit. This is because the large $\overline{N}$ asymptotic limit for the negative binomial distribution depends on the cluster parameter and may be summarized thus

$$p(x) = \begin{cases} \dfrac{\alpha}{\alpha - 1} & \text{if } \alpha > 1 \\ -\ln(x) & \text{if } \alpha = 1 \\ \dfrac{\alpha^{\alpha}\Gamma(1-\alpha)x^{\alpha-1}}{\Gamma(\alpha)} & \text{if } \alpha < 1. \end{cases} \tag{3.4}$$

It would appear that the value of the probability density near the origin when used in conjunction with its derivative can provide a gauge for the number of scattering centres present or, equivalently, the number of steps in the random walk that have taken place. It is pertinent to enquire whether these results are sensitive to the nature of the individual steps

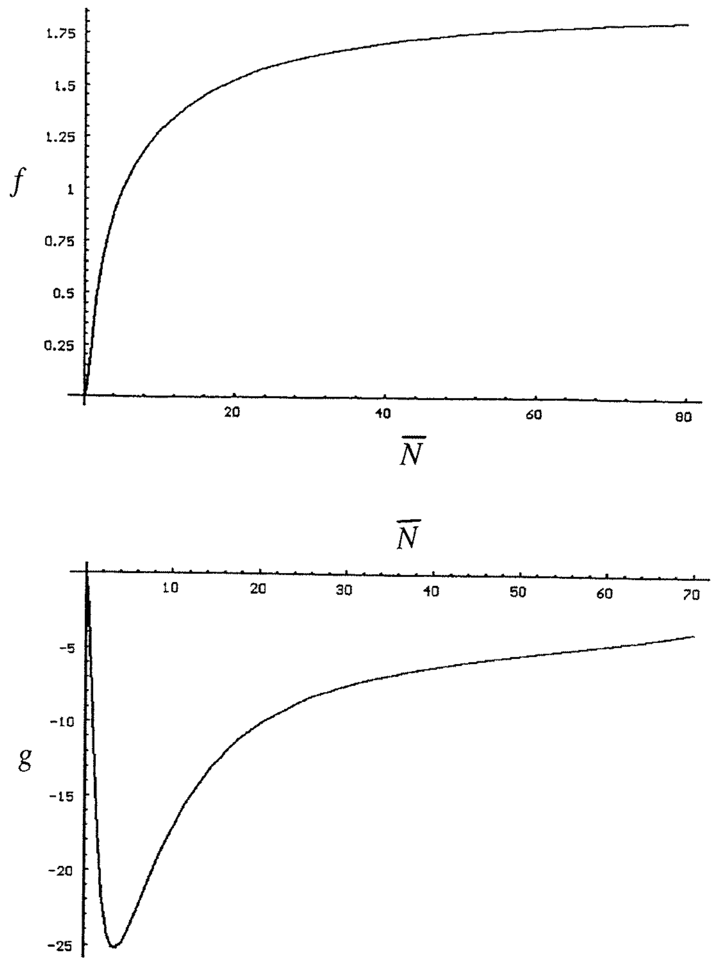

Figure 5. The functions f and g for a random walk with identical steps having negative binomial number fluctuations with cluster parameter $\alpha = 2$. A value of $x = 1/100$ is used for both curves.

comprising the walk and to the distribution describing the number of steps. These questions will be addressed in the following section.

4. Limiting forms of distributions for variable step lengths

The results of the last section indicate that the value of the pdf near the origin contains information about the number of steps that have taken place in a two-dimensional random walk where each of the steps is identical. The relaxation of this constraint to enable consideration of the case when the step lengths are fluctuating is straightforward, requiring only a slight generalization of equation (3.1). Provided that each of the steps is statistically similar, the resultant pdf may be written as:

$$P_{\overline{N}}(I) = \frac{1}{2}\int_0^\infty u J_0\left(u\sqrt{I}\right) Q(\langle J_0(ur)\rangle)\,\mathrm{d}u \tag{4.1}$$

where $\langle\cdots\rangle$ denotes an ensemble average over the distribution for the step lengths.

A model for the distribution of r which enables the calculations to be performed analytically throughout but which is nevertheless relevant to numerous instances in non-Gaussian scattering is the K-distribution [15–16], for which

$$p(r) = \frac{2b}{\Gamma(\nu)}\left(\frac{br}{2}\right)^{\nu} K_{\nu-1}(br) \qquad \text{for } \nu > 0.$$

Inserting this into (4.1) gives

$$P_N(I) = \frac{b^2}{2\Gamma(N\nu)}\left(\frac{I^{1/2}b}{2}\right)^{N\nu-1} K_{N\nu-1}(bI^{1/2}) \tag{4.2}$$

being the pdf of the resultant I formed from a sum of N K-distributed steps. This result shows that the resultant intensity is also K-distributed but with a scaled index, illustrating the stability properties [17] of this distribution. The value and physical interpretation of the index appearing in the K-distribution is rather dependent upon the particular application being studied. These issues are considered in [18, 19]. The mean of the K-distribution is $\langle I\rangle = 4N\nu/b^2$ and this may be used to form the dimensionless quantities

$$f = \lim_{l\to 0}\langle I\rangle P_N(I)$$

$$g = \lim_{l\to 0}\langle I\rangle^2 \frac{\mathrm{d}P_N(I)}{\mathrm{d}I}.$$

It is because r is fluctuating that it is possible to proceed directly to the limit in this case and evaluate the pdf and its slope at the origin. However, for certain values of the index ν, the K-distribution is singular at the origin, in which case f and g must, as before, be determined for some small value of intensity. These calculations are possible to perform using the same technique used in section 3. The details add little to the argument and so it is the idealized results of f and g at the origin that will be discussed here. Provided then that the scaled index $N\nu > 1$, f is easily found with the aid of asymptotic formulae for Bessel functions [11] to be

$$f = \frac{N\nu}{(N\nu - 1)} \tag{4.3}$$

and if $N\nu > 2$, g can be found in a similar fashion:

$$g = \frac{-(N\nu)^2}{(N\nu-1)(N\nu-2)}. \tag{4.4}$$

The different conditions for f and g on the scaled index stem from the logarithmic singularity of the K-distribution at the origin when $\nu = 1$. These simple expressions exhibit many of the salient features appearing in the analyses for fixed step length. Provided that $N\nu > 1$ the value of f is finite and decreases to unity with increasing $N\nu$, whilst the value of g in the same limit is -1. Both these results are symptomatic of the move towards exponential intensity statistics as one would expect from the central limit theorem.

More interesting and significant still is the behaviour of f and g when N is itself fluctuating according to some random process. The type of number fluctuations that yield to an analytical approach include those described by binomial, Poisson and negative binomial distributions, the last two of which have already been analysed. Taken as a group these three are important because they span the range for which the normalized variance is respectively less-than, equal to and in excess of the mean. Taken in combination therefore, the results obtained using these models will be indicative of any instance when number fluctuations occur.

When N fluctuates according to a Poisson process, the resultant pdf at the origin is readily found from (4.1) and (4.3) to be

$$P_{\overline{N}}(0) = \frac{b^2 \exp(-\overline{N})}{4\nu} \sum_{n=1}^{\infty} \frac{\overline{N}^n}{n!(n-1/\nu)}.$$

The contribution for $n = 0$ is omitted since it corresponds to no particles being present. This series may be summed in terms of tabulated functions to give

$$f = \overline{N}\nu \exp(-\overline{N}) \left(1 - \frac{\pi}{\nu} \operatorname{cosec}\left(\frac{\pi}{\nu}\right) L^{1/\nu}_{-1/\nu}(\overline{N})\right)$$

where $L^a_b(x)$ is the Laguerre function [11]. Figure 6 shows f plotted as a function of $\overline{N}$ for different values of the index ν. When $\overline{N}$ is less than unity there are on average very few particles present and so the value of the pdf at the origin must, as it does, fall to zero. For values of $\overline{N}$ in the range 1–10 f peaks before declining to the Gaussian value of 1 when $\overline{N}$ is large. The value of the peak, and to a lesser extent its location, depends on the index of the K-distribution but is typically in excess of 1.5 for those values of ν shown. If $f \lesssim 1$ then one may categorically state that there are fewer than three particles present on average. For $\overline{N} \gtrsim 4$ f is double valued and so an additional measure is required to obtain a unique estimate for $\overline{N}$, this being the derivative g.

The dimensionless derivative can be found in a similar fashion to be:

$$g = \frac{-1}{4}\overline{N}^2 \exp(-\overline{N}) \left\{\nu \operatorname{cosec}\left(\frac{\pi}{\nu}\right) \sec\left(\frac{\pi}{\nu}\right) \left[4\pi \cos\left(\frac{\pi}{\nu}\right) L^{1/\nu}_{-1/\nu}(\overline{N}) - 2\pi L^{2/\nu}_{-2/\nu}(\overline{N}) \right.\right.$$
$$\left.\left. -\nu \sin\left(\frac{2\pi}{\nu}\right)\right]\right\}.$$

Figure 6(b) shows g in the same range for different values of the index ν. The correct asymptotes of zero and -1 are obtained for small and large values of $\overline{N}$ respectively, with significant deviations from these asymptotes when $\overline{N}$ is of order unity.

Similar behaviour is obtained when negative binomial and binomial number fluctuations are considered. The relevant sums for negative binomial number fluctuations are:

$$f = \frac{\overline{N}\nu}{(1+\overline{N}/\alpha)^\alpha} \sum_{N=1}^{\infty} \binom{N+\alpha-1}{N} \left(\frac{\overline{N}}{\alpha+\overline{N}}\right)^N \frac{1}{(N\nu-1)}$$

$$g = \frac{-\overline{N}^2\nu^2}{(1+\overline{N}/\alpha)^\alpha} \sum_{N=1}^{\infty} \binom{N+\alpha-1}{N} \left(\frac{\overline{N}}{\alpha+\overline{N}}\right)^N \frac{1}{(N\nu-1)(N\nu-2)}$$

where α is the cluster parameter. Again, these may be summed in terms of tabulated functions to give

$$f = \frac{\nu\overline{N}}{(1+\overline{N}/\alpha)^\alpha} \left\{1 - {}_2F_1\left(\alpha, \frac{-1}{\nu}, 1-\frac{1}{\nu}, \frac{\overline{N}}{(\overline{N}+\alpha)}\right)\right\}$$

$$g = \frac{-\nu^2\overline{N}^2}{2(1+\overline{N}/\alpha)^\alpha} \left\{2\,{}_2F_1\left(\alpha, \frac{-1}{\nu}, 1-\frac{1}{\nu}, \frac{\overline{N}}{(\overline{N}+\alpha)}\right) - {}_2F_1\left(\alpha, \frac{-2}{\nu}, 1-\frac{2}{\nu}, \frac{\overline{N}}{(\overline{N}+\alpha)}\right) - 1\right\}$$

where ${}_2F_1$ is Gauss' hypergeometric function [20]. A set of curves are shown for f and g as functions of $\overline{N}$ for various values of the K-distribution index ν and cluster parameter. Figure 7 shows f and g for $\alpha = 3, 5$ and 10. The generic form of the curves is similar to that shown for the Poisson fluctuations, but the cluster parameter now plays an important role, leading to a significant enhancement in the peak values of f and g. Moreover the asymptotic limits are no longer given by the Gaussian values, but depend on the cluster

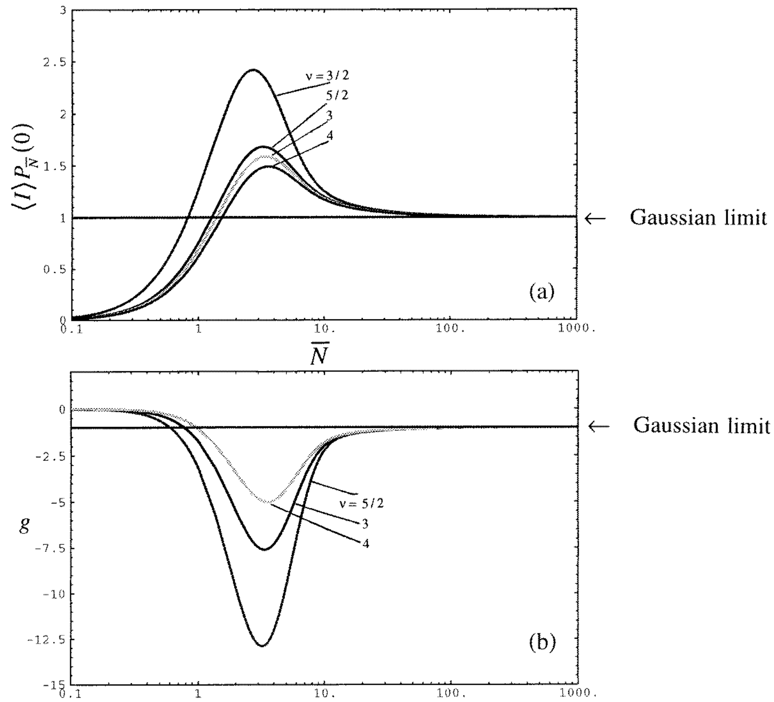

Figure 6. (a) The function f for Poisson number fluctuations for a random walk with fluctuating step lengths. The index ν of the K-distribution that governs the step-size distribution produces the different curves. (b) The function g for Poisson number fluctuations for a random walk with fluctuating step lengths for different values of ν.

parameter. Note that these limits are not attained until $\overline{N}$ is surprisingly large, illustrating the importance of the clustering mechanism. Specifically the asymptotic limits are easily shown to be

$$\lim_{\overline{N}\to\infty} f = \frac{\alpha}{\alpha - 1}$$
$$\lim_{\overline{N}\to\infty} g = \frac{-\alpha^2}{(\alpha - 1)(\alpha - 2)}.$$

These tend to the Gaussian values when $\alpha \to \infty$, as they should, for in this limit the negative binomial distribution is asymptotic to the Poisson distribution whose results have already been discussed.

The last class of number fluctuation to consider is binomial. In this case the relevant sums become:

$$f = \nu\overline{N}(1-p)^{\overline{N}/p} \sum_{N=1}^{\overline{N}/p} \binom{\overline{N}/p}{N} \left(\frac{p}{1-p}\right)^N \frac{1}{N\nu - 1}$$
$$g = -(\nu\overline{N})^2(1-p)^{\overline{N}/p} \sum_{N=1}^{\overline{N}/p} \binom{\overline{N}/p}{N} \left(\frac{p}{1-p}\right)^N \frac{1}{(N\nu - 1)(N\nu - 2)}$$

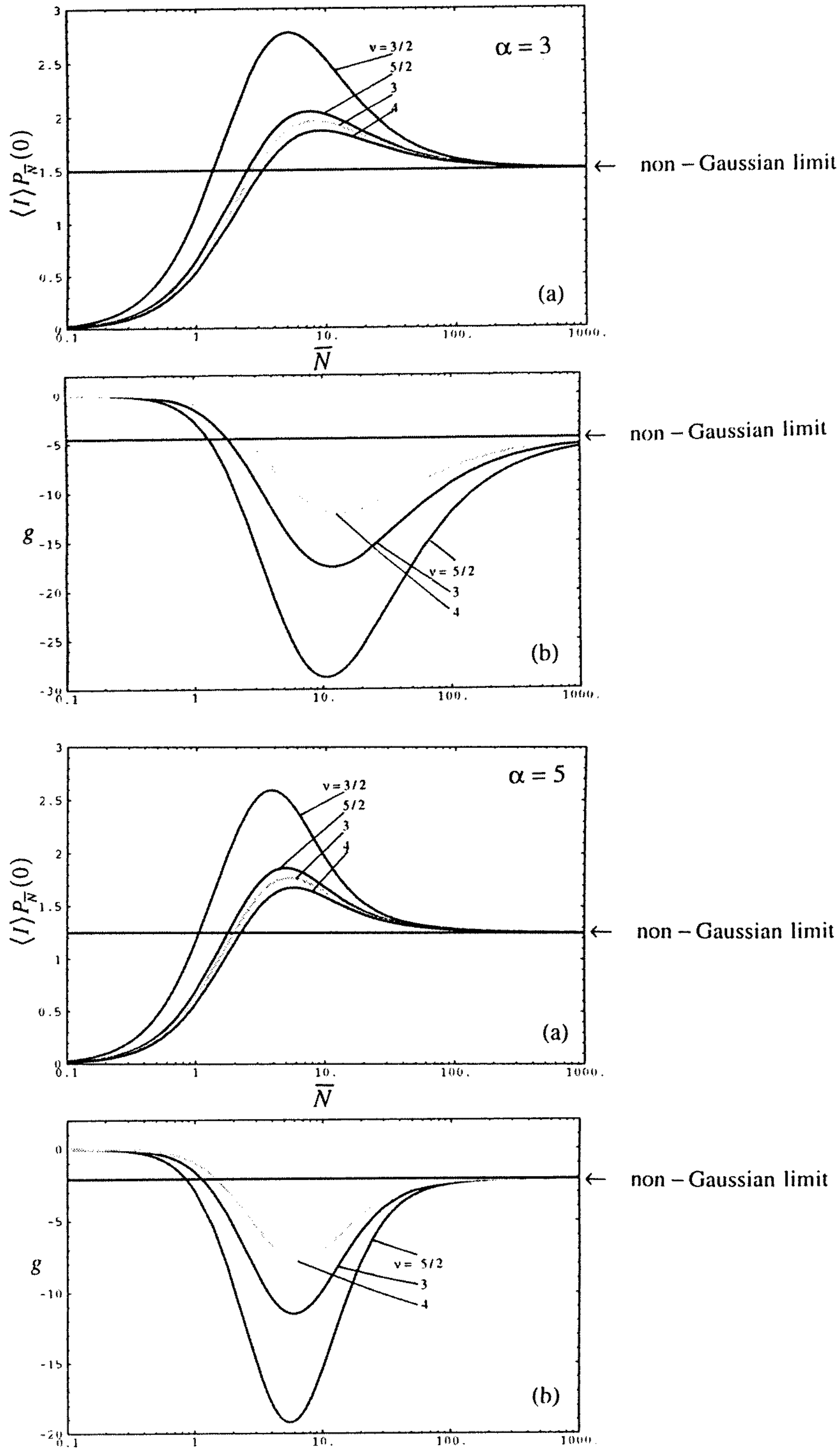

Figure 7. The functions f and g for negative binomial number fluctuations with $\alpha = 3, 5$ and 10. The curves for $\alpha = 10$ are similar to those shown for Poisson fluctuations. For smaller values of α, the absolute maximum values of the functions increase and the curves broaden. The asymptotic limit is no longer the Gaussian one for large values of $\overline{N}$.

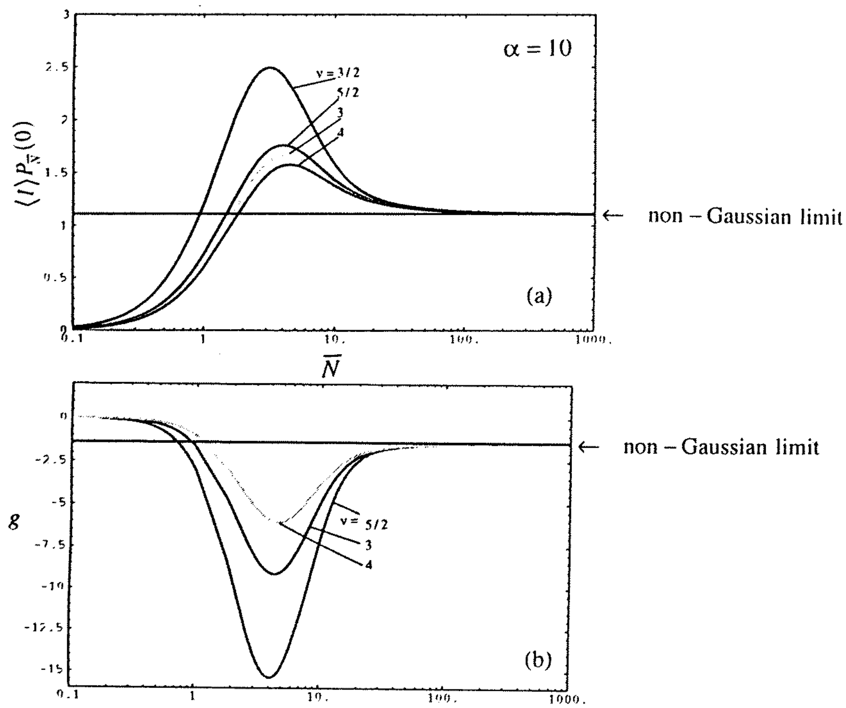

Figure 7. (Continued)

where $p < 1$ is the probability that a group of $\overline{N}$ particles is present. For the purposes of tabulation these sums may be written as

$$f = \nu\overline{N}(1-p)^{\overline{N}/p}\left(1 - {}_2F_1\left(-\frac{\overline{N}}{p}, \frac{-1}{\nu}, 1-\frac{1}{\nu}, \frac{-p}{1-p}\right)\right)$$

$$g = -\frac{1}{2}(\nu\overline{N})^2(1-p)^{\overline{N}/p}\left(2\,{}_2F_1\left(-\frac{\overline{N}}{p}, \frac{-1}{\nu}, 1-\frac{1}{\nu}, \frac{-p}{1-p}\right)\right.$$
$$\left. - {}_2F_1\left(-\frac{\overline{N}}{p}, \frac{-2}{\nu}, 1-\frac{2}{\nu}, \frac{-p}{1-p}\right) - 1\right).$$

Figure 8 shows f and g for two values for p. When p is small the binomial distribution may be approximated by the Poisson distribution and so the results displayed in figure 6 are recovered. For increasing values of p the general form of the curves is as before with the position of the peak occurring for $\overline{N}$ of order unity. Note that for larger values of p the peak value of f is progressively less than is the case for Poisson number fluctuations, but the position of the peak occurs at systematically smaller values of $\overline{N}$. This observation is to be expected because $p = 1$ corresponds to the case in which the binomial distribution becomes deterministic. Hence the curves for fixed N that are analogous to those depicted in figure 3 are recovered. The curves will not be identical because the step lengths were taken to be fixed for the case depicted in figure 3. This therefore serves to illustrate the point that fluctuations in the step size serve to enhance the value of $P(0)$ and its slope at the origin. A reason for this enhancement is that the fluctuations in step size reduce the *effective* number of contributions to the resultant of the random walk and when N is small the values the pdfs take at the origin have been shown to deviate significantly from the

Gaussian value.

The emergent pattern of behaviour shown by these curves is characterized by the variance of the number fluctuations. For distributions whose variance is less than the mean, the curve for f has a broader peak for smaller values of $\overline{N}$ but the peak value is less than the Poissonian one. The results are asymptotic to the Gaussian limit for large values of $\overline{N}$. For number fluctuations whose normalized variance are in excess of the mean, the curve for f is broader for larger values of $\overline{N}$ and the peak value is substantially greater than the corresponding Poisson result. The curves are generally not asymptotic with the Gaussian limit. In general, all the curves exhibit a significant departure from the Gaussian values when $\overline{N}$ is less than ~ 10. The overall behaviour is not that dissimilar from the results for fixed step lengths, which showed a very slight enhancement above the Gaussian values at moderate $\overline{N}$. The variability of the step lengths amplifies this effect without substantially altering the overall morphology of the curves and is therefore likely to be a robust feature of any system where the step lengths fluctuate, irrespective of their governing distribution.

5. Discussion

This paper has examined some of the properties of non-Gaussian pdfs that can arise in coherent optical scattering. These pdfs were obtained using elements of the theory of random walks in the plane. It has been shown that their form near the origin allows the number of steps in the walk, or equivalently the number of scatterers or coherency volumes, to be determined. The results that have been shown would appear to be generic and do not depend, except in detail, on the fluctuating nature of the number or length of the steps comprising the walk.

Insofar as these results are interpreted or used in the context of the scattering of coherent radiation, the measurement of the pdf and its slope at or near to the origin by a real detector would necessarily require a degree of integration by the instrument. This will naturally involve a further averaging of the distributions, leading to an additional smoothing of the curves. It should be stressed, however, that the curves have been produced as the result of two separate averaging processes and, as has already been mentioned, their form is comparatively insensitive to the details of this process. It may be expected therefore that the curves will retain their structure, altering only in the magnitude of the effect.

The use of low-intensity light sources or weak scatterers necessitate the use of photon-counting techniques which again introduces an averaging process. Photon-counting causes additional fluctuations to occur, for there is a finite probability of measuring no photons even from a source of constant intensity. The photocount probability distribution $P(n)$ for detecting n photons can be obtained from the Mandel relationship [21] and depends on the efficiency η of the detector and the sample time t:

$$P_N(n) = \frac{(\eta t)^n}{n!} \int_0^\infty I^n P_N(I) \exp(-\eta t I)\, \mathrm{d}I.$$

The analogous measure would be to determine $P_N(0)$, or for n small, as a function of the expected number of scatterers N that are present. This quantity is related to the characteristic function of the intensity pdf and so may always be calculated. For the K-distribution model discussed in section 4, the photocount distribution has the advantage that it can be found exactly in terms of tabulated Whittaker functions [16]:

$$P_N(n) = \left(\frac{Q}{\overline{n}}\right)^{N(\nu+1)/2} \frac{\Gamma(Q+n)}{\Gamma(Q)} \exp\left(\frac{Q}{2\overline{n}}\right) W_{-(Q/2+n),(Q-1)/2}\left(\frac{Q}{\overline{n}}\right)$$

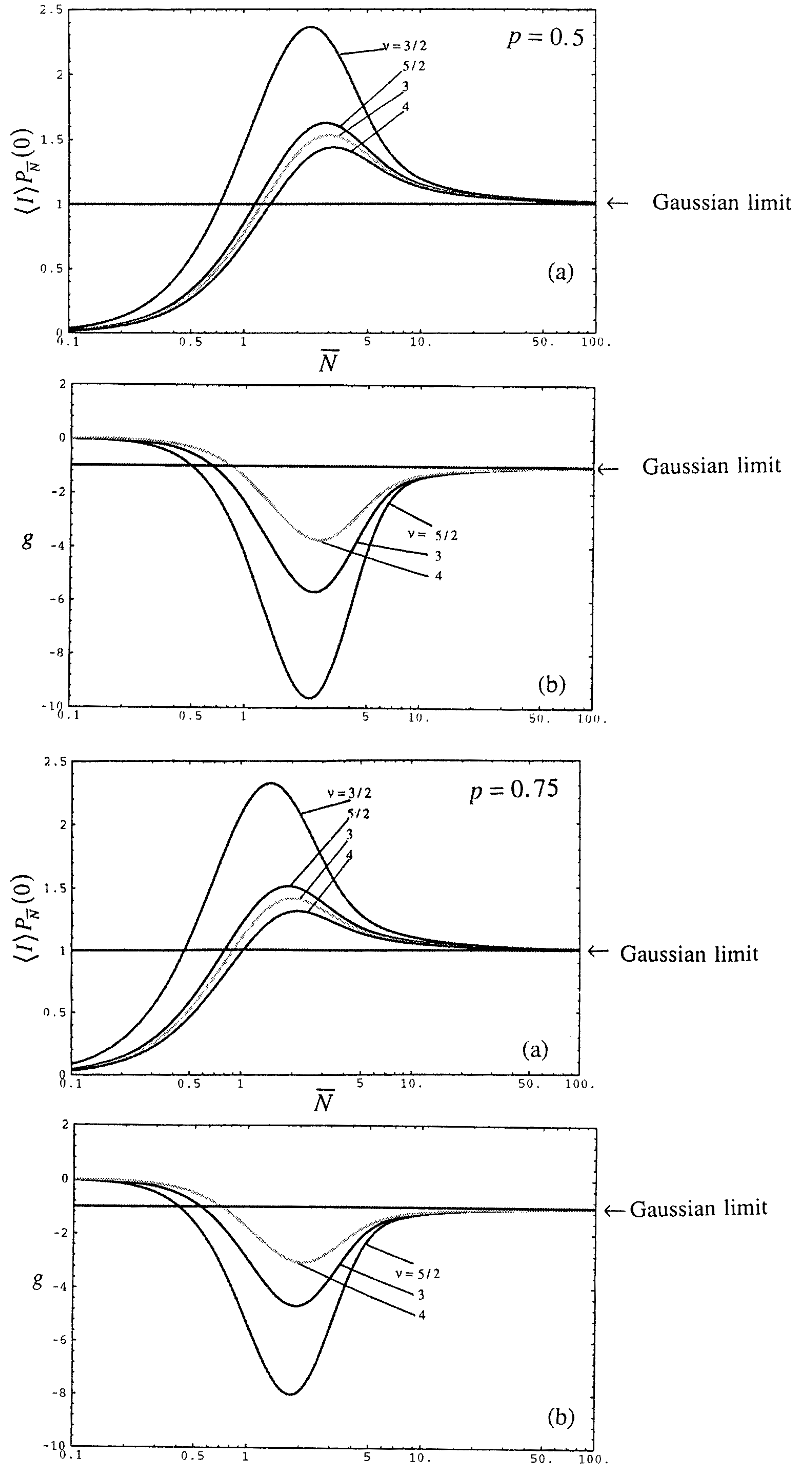

Figure 8. The functions f and g for binomial number fluctuations with $p = 0.75$ and 0.5 being the probability that $\overline{N}$ particles are present. The absolute maximum value of the curves and the breadth of the peak both increase with decreasing certainty that the particles are present. The asymptotic limit is once again the Gaussian one for large values of $\overline{N}$.

where $\bar{n} = 4\eta t N(1+\nu)/b^2$ is the mean number of photocounts, this being proportional to the sample time, and $Q = N(\nu+1)$. The photocount density function is likely to be more sensitive to the details of the distributions used for the random walk and these matters will be considered elsewhere.

The work described in this paper may be of consequence to any coherent optical detection system working at or near to the diffraction limit for focused light in those circumstances when a measure for the number of scattering objects is required. These instances may range from low concentrations of contaminant in solution to unresolved clusters of objects. The work may also find an application to areas other than light scattering, such as low-dimensional systems in solid state physics and in population dynamics, where the concept of the random walk has proved to be of value.

This work can be generalized to consider random walks in three or more dimensions with comparative ease. It is well known that the continuum limit of a random walk is equivalent to obtaining a solution to the diffusion equation. The introduction of clustering in the number of steps and variability in the step lengths modifies the equivalent diffusion process. Such a mechanism may find an application to modelling impurities in semiconductors among others.

Acknowledgments

T R Watts is supported by a University of Nottingham Postgraduate Scholarship. We would like to thank Professor E Jakeman and Dr J Walker for illuminating discussions.

Appendix

The method by which the curves appearing in section 3 are obtained warrants some discussion. They are formed using the analytical results (2.2, 2.3), the Fourier–Bessel series (2.4) and an asymptotic expansion of (2.1) that is employed for sufficiently large values of N. This latter expansion is readily found from an expansion of the Bessel function appearing in the integrand of (2.1) viz.

$$P_N(I) = \frac{1}{2}\int_0^\infty u J_0\left(u\sqrt{I}\right)\exp(-(ur)^2 N/4)\left(1 - \frac{3N(ur)^4}{64}\right) + \cdots \,\mathrm{d}u$$

whereupon integrating term by term yields the result

$$P_N(I) = \frac{1}{\langle I\rangle}\exp(-I/\langle I\rangle)\left(1 - \frac{3}{2N}\left(\left(\frac{I}{\langle I\rangle} - 2\right)^2 - 2\right) + O\left(\frac{1}{N^2}\right) + \cdots\right) \tag{A1}$$

where $\langle I\rangle = Nr^2$. This pdf may be used as it stands when N is fixed, but must be averaged over the appropriate distribution $p(N)$ when N is fluctuating. This requires evaluating

$$P_{\overline{N}}(I) = \sum_{n=1}^{\infty} p(N) P_N(I). \tag{A2}$$

In this sum exact values of $P_N(I)$ are used for $N = 2, 3$ (from equations (2.2) and (2.3)), the Bessel expansions (from (3.4)) for $N = 4$–100 and (A1) for $N > 100$. In order to use (A1), sums having the form:

$$\sum_{N=1}^{\infty} \frac{p(N)}{N^k} \qquad k = 1, 2, \ldots \tag{A3}$$

must be determined, the details for which depend upon the particular distribution being considered. For Poisson number fluctuations the sum is expressible in terms of generalized hypergeometric functions [15]:

$$\exp(-\overline{N}) \sum_{N=1}^{\infty} \frac{\overline{N}^N}{N!N^k} = \overline{N} \exp(-\overline{N})\, {}_{k+1}F_{k+1}\left(\begin{matrix} 1, \ldots 1; \\ \\ 2, \ldots 2, \end{matrix} \quad \overline{N}\right).$$

Following this procedure $P_{\overline{N}}(I)$ can be found to any desired accuracy. However, since values near the origin are required, the first few terms are sufficient.

If the number fluctuations follow a negative binomial distribution, the sums (A3) appearing in the average are also expressible in terms of the hypergeometric functions giving

$$\sum_{N=1}^{\infty} \binom{N+\alpha-1}{N} \frac{(\overline{N}/\alpha)^N}{N^k(1+(\overline{N}/\alpha))^{N+\alpha}}$$
$$= \alpha \left(\frac{\alpha}{\alpha+\overline{N}}\right)^{\alpha} \left(\frac{\overline{N}}{\alpha+\overline{N}}\right)_{(k+2)} F_{(k+1)}\left(\begin{matrix} 1, \ldots, 1, \alpha+1; \\ \\ 2, \ldots, 2 \end{matrix} \quad \frac{\overline{N}}{\alpha+\overline{N}}\right).$$

References

[1] Pearson K 1906 *A mathematical theory of random migration* Drapers' Company Research Memoirs, Biometric Series III (London: Dulan)

[2] Chandrasekhar S 1943 Stochastic problems in physics and astronomy *Rev. Mod. Phys.* **15** 1–89

[3] Rayleigh J W S 1880 *Phil. Mag.* **10** 75

[4] Pusey P 1977 Statistical properties of scattered radiation *Photon Correlation and Velocimetry* ed H Z Cummins and G R Pike (New York: Plenum) pp 45–141

[5] Jakeman E and Tough R 1988 Non-Gaussian noise processes *Adv. Phys.* **37** 471

[6] Kluyver J C 1905 A local probability problem *Proc. Koninklijke Akademie van Vetenschappen te Amsterdam* **14** 341–50

[7] Pusey P N, Schaefer D W and Koppel D E 1974 Single-interval statistics of light scattererd by identical scatterers *J. Phys. A: Math. Gen.* **7** 530–40

[8] Tapster P R, Weeks A R, Pusey P N and Jakeman E 1989 Analysis of probability-density functions for laser scintillations in a turbulent atmosphere *J. Opt. Soc. Am.* A **6** 782–85

[9] Hopcraft K I, Green D and Smith P R 1995 Non-Gaussian fluctuations of light scattered from particles and their characterisation *Proc. Progress in Electromagnetics Research Symposium '95* (Seattle: Washington) p 1080

[10] Bates A P, Hopcraft K I and Jakeman E 1995 Intensity correlations of polarised light scattered from ellipsoidal particles *Proc. Progress in Electromagnetics Research Symposium '95* (Seattle: Washington) p 78

[11] Abramowitz M and Stegun I 1972 *Handbook of Mathematical Functions* 9th edn (New York: Dover)

[12] Merzbacher E, Feagin J M and Wu T-H 1977 Superposition of the radiation from *N* independent sources and the problem of random flights *Am. J. Phys.* **45** 964–69

[13] Bennet W R 1947 Distribution of the sum of randomly phased components *Q. Appl. Math.* **5** 303

[14] Evans M, Hastings N and Peacock B 1993 *Statistical Distributions* 2nd edn (New York: Wiley)

[15] Jakeman E 1980 On the statistics of K-distributed noise *J. Phys. A: Math. Gen.* **13** 31–48

[16] Jakeman E and Pusey P N 1976 A model for non-Rayleigh sea echo *Proc. IEEE* **AP-24** 806–14

[17] Gnedenko B D 1968 *The Theory of Probability* 4th edn (New York: Chelsea)

[18] Jakeman E and Pusey P 1978 Significance of K-distributions in scattering experiments *Phys. Rev. Lett.* **40** 546

[19] Barakat R1986 Weak-scatterer generalization of the K-density function with application to laser scattering in atmospheric turbulence *J. Opt. Soc. Am.* A **3** 401

[20] Erdélyi A 1954 *Tables of Integral Transforms* vol II (New York: McGraw-Hill)

[21] Mandel L and Wolf E 1995 *Optical coherence and quantum optics* (Cambridge: Cambridge Univesity Press)